Interfaces + Recursion = Language?

Studies in Generative Grammar 89

Editors

Henk van Riemsdijk
Jan Koster
Harry van der Hulst

Mouton de Gruyter
Berlin · New York

Interfaces + Recursion = Language?

Chomsky's Minimalism and the View from Syntax-Semantics

edited by

Uli Sauerland
Hans-Martin Gärtner

Mouton de Gruyter
Berlin · New York

Mouton de Gruyter (formerly Mouton, The Hague)
is a Division of Walter de Gruyter GmbH & Co. KG, Berlin.

The series Studies in Generative Grammar was formerly published by
Foris Publications Holland.

♾ Printed on acid-free paper which falls within the guidelines
of the ANSI to ensure permanence and durability.

Library of Congress Cataloging-in-Publication Data

Interfaces + recursion = language? : Chomsky's minimalism and the view from syntax-semantics / edited by Uli Sauerland and Hans-Martin Gärtner.
p. cm. − (Studies in generative grammar ; 89)
Includes bibliographical references and index.
ISBN-13: 978-3-11-018872-1 (cloth : alk. paper)
1. Minimalist theory (Linguistics) 2. Grammar, Comparative and general − Syntax. 3. Semantics. I. Sauerland, Uli. II. Gärtner, Hans-Martin. III. Title: Interfaces plus recursion equals language?
P158.28.I58 2007
415−dc22

2007013045

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

ISBN 978-3-11-018872-1
ISSN 0167-4331

Cover design: Sigurd Wendland, Berlin.
Printed in Germany.

Preface

Uli Sauerland and Hans-Martin Gärtner

Human language is a phenomenon of immense richness: It provides finely nuanced means of expression that underlie the formation of culture and society; it is subject to subtle, unexpected constraints like syntactic islands and cross-over phenomena; different mutually-unintelligible individual languages are numerous; and the descriptions of individual languages occupy thousands of pages. Recent work in linguistics, however, has tried to argue that despite all appearances to the contrary, the human biological capacity for language may be reducible to a small inventory of core cognitive competencies. The most radical version of this view has emerged from the Minimalist Program (Chomsky 1995, 2005): The claim that language consists of only the ability to generate recursive structures by a computational mechanism. On this view, all other properties of language must result from the interaction at the interfaces of that mechanism and other mental systems not exclusively devoted to language. Since language could then be described as the simplest recursive system satisfying the requirements of the interfaces, one can speak of the Minimalist Equation: Interfaces + Recursion = Language.

The question whether all the richness of language can be reduced to this minimalist equation has already inspired several fruitful lines of research that led to important new results (see the review in Hauser et al. 2002). While a full assessment of the minimalist equation will require evidence from many different areas of inquiry, this volume focuses especially on the perspective of syntax and semantics. Within the minimalist architecture, this places our concern with the core computational mechanism and the (LF-)interface where recursive structures are fed to interpretation. Specific questions that the papers address are: What kind of recursive structures can the core generator form? How can we determine what the simplest recursive system is? How can properties of language that used to be ascribed to the recursive generator be reduced to interface properties? What effects do syntactic operations have on semantic interpretation? To what extent do models of semantic interpretation support the LF-interface conditions postulated by minimalist syntax?

The symposium "Interfaces + Language = Recursion? The View from Syntax-Semantics" was organized in March 2005 jointly by the Zentrum

für allgemeine Sprachwissenschaft (Center for General Linguistics) and the Berlin Brandenburgische Akademie der Wissenschaften (Berlin-Brandenburg Academy of Sciences) on the occasion of Noam Chomsky's visit to Berlin. In addition to the lecture by Noam Chomsky, six other leading scholars in syntax and semantics presented current work of their own that relates to the minimalist equation "Interfaces + Language = Recursion". Furthermore, Günther Grewendorf, Dieter Wunderlich, and Angela Friederici presented brief comments on the talk by Noam Chomsky. Over 350 linguists from Germany and many other countries attended the event and contributed to the discussion, that this book emerged from. This volume contains papers based on all seven paper presentations that were given at the symposium.

We would like to thank all the people who contributed to the success of symposium and the creation of this volume. Many people from the Center for General Linguistics and Berlin-Brandenburg Academy of Sciences were involved in the various aspects organizing the symposium and without all this help it would have been impossible to create such an event. Especially, we thank Manfred Bierwisch for his active support at all stages of the symposium. The creation of this book has been a pleasure thanks to all the wonderful people who we could work with. We especially thank Paul David Doherty for his help with typestting, Julia Köhler for help with the index, the series editor Henk van Riemsdijk for his support, and Ursula Kleinhenz of Mouton de Gruyter for her advice at various stages of this enterprise. Finally, we thank the six anonymous colleagues who reviewed the papers appearing in this volume for their valuable contributions.

References

Chomsky, Noam
1995 *The Minimalist Program.* MIT Press, Cambridge, Mass.
2005 Three factors in language design. *Linguistic Inquiry*, 36: 1–22.

Hauser, Marc D., Noam Chomsky, and Tecumseh Fitch
2002 The faculty of language: what is it, who has it, and how did it evolve? *Science*, 298: 1569–1579.

Table of Contents

Preface v

Approaching UG from Below 1
Noam Chomsky

The Subject-In-Situ Generalization Revisited 31
Artemis Alexiadou and Elena Anagnostopoulou

Towards a Relativized Concept of Cyclic Linearization 61
Gereon Müller

Strategies of Subject Extraction 115
Luigi Rizzi and Ur Shlonsky

Some Remarks on Locality Conditions and Minimalist Grammars 161
Hans-Martin Gärtner and Jens Michaelis

Flat Binding: Binding without Sequences 197
Uli Sauerland

The Grammar of Focus Interpretation 255
Sigrid Beck

List of Contributors 281

Index of Languages 283

Index of Subjects 285

Approaching UG from Below

Noam Chomsky

The problem problem that has virtually defined the serious study of language since its ancient origins, if only implicitly, is to identify the specific nature of this distinctive human possession. Within the "biolinguistic perspective" that began to take shape fifty years ago, the concern is transmuted into the effort to determine the genetic endowment of the faculty of language FL, understood to be a "cognitive organ," in this case virtually shared among humans and in crucial respects unique to them, hence a kind of species property. So construed, language is I-language, a state of FL, and universal grammar (UG) is reinterpreted as the theory of the initial state of FL. The term "biolinguistics" itself was coined in 1974 by Massimo Piattelli-Palmerini as the topic for an international conference he organized for the Royaumont Institute in Paris and MIT, bringing together evolutionary biologists, neuroscientists, linguists, and others concerned with language and biology, one of many such initiatives, before and since, which sought to explore the extent to which apparent principles of language are unique to this cognitive system, one of "the basic questions to be asked from the biological point of view," as discussed there, and crucial for the study of development of language in the individual and its evolution in the species.[1]

Within the biolinguistic framework, methodological considerations of simplicity, elegance, etc., can often be reframed as empirical theses concerning organic systems generally. For example, Morris Halle's classical argument against postulating a linguistic level of structuralist phonemics was that it required unmotivated redundancy of rules, taken to be a violation of natural methodological assumptions. Similarly conclusions about ordering and cyclicity of phonological and syntactic rule systems from the 1950s were justified on the methodological grounds that they reduce descriptive complexity and eliminate stipulations. In such cases, the issues can be recast as metaphysical rather than epistemological: Is that how the world works? The issues can then be subjected to comparative analysis and related to principles of biology more generally, and perhaps even more fundamental principles about

[1] Piattelli-Palmerini, ed., *A Debate on Bio-Linguistics*, Endicott House, Dedham Mass, May 20-21, 1974.

the natural world; clearly a step forward, if feasible. Such options become open, in principle at least, if the inquiry is taken to be the study of a real object, a biological organ, comparable to the visual or immune systems, the systems of motor organization and planning, and many other subcomponents of the organism that interact to yield the full complexity of thought and action, abstracted for special investigation because of their apparent internal integrity and special properties. From the earliest days there have been efforts to explore closer links between general biology and the biolinguistic perspective. Insofar as methodological arguments in linguistics can be reframed as empirical ones about general operative principles, the analogies may become more substantive.

At the time of the 1974 discussions, it seemed that FL must be rich, highly structured, and substantially unique. In particular, that conclusion was drawn from considerations of language acquisition. The only plausible idea seemed to be that the process is a form of theory construction. Somehow, the child reflexively categorizes certain sensory data as linguistic, not a trivial achievement in itself, and then uses the constructed linguistic experience as evidence for a theory that generates an infinite variety of expressions, each of which contains the information about sound, meaning, and structure that is relevant for the myriad varieties of language use. It was well understood that construction of theories must be guided by what Charles Sanders Peirce had called an "abductive principle" that "puts a limit upon admissible hypotheses," so that the mind is capable of "imagining correct theories of some kind" while discarding infinitely many others consistent with the evidence. Peirce was considering theory construction in the sciences, but the same general observation holds for growth/acquisition of language.[2] In this case, it appeared that the format that limits admissible hypotheses must be highly restrictive, given the empirical facts of acquisition and convergence. The conclusions about the specificity and richness of the language faculty seemed to follow directly. Plainly such conclusions pose serious problems for dealing with the diversity of languages: the well-known tension between descriptive and explanatory adequacy. The conclusions also raise barriers to inquiry into how the faculty might have evolved, since any property specific to language calls for an evolutionary explanation. These matters were discussed repeatedly, and inconclusively, at the 1974 conference.

The crystallization of the Principles & Parameters program a few years later suggested ways to reduce the tension between descriptive and explanatory

[2]Though not, to be sure, the notion of convergence to the correct theory, as in Peirce's concerns.

adequacy, as is familiar. It also removed a major conceptual barrier to the study of evolution of language. With the divorce of principles of language from acquisition, now understood to be a matter of parameter setting, it no longer follows that the format of UG that "limits admissible hypotheses" must be rich and highly structured to satisfy the empirical conditions of rapid convergence on generative systems of the kind required to determine meaning and external manifestation. That might turn out to be the case, but it is no longer an apparent conceptual necessity.

The P&P approach largely emerged from intensive study of a range of languages, but it was also suggested by major developments in general biology, specifically François Jacob's account of how slight changes in the hierarchy and timing of regulatory mechanisms might yield great superficial differences – a butterfly or an elephant, and so on. The model seemed natural for language as well: slight changes in parameter settings might yield superficial variety, through interaction of invariant principles with parameter choices (Chomsky 1980, p.67). The P&P framework also made it possible to pursue more seriously the recognition, from the earliest days of generative grammar, that acquisition of language involves not only a few years of experience and millions of years of evolution, but also "principles of neural organization that may be even more deeply grounded in physical law" (Chomsky 1965, p.59). Again, somewhat parallel developments were proceeding in general biology, now sometimes called the "evo-devo revolution."[3]

Evidently, development of language in the individual must involve three factors: (1) genetic endowment, which sets limits on the attainable languages, thereby making language acquisition possible; (2) external data, converted to the experience that selects one or another language within a narrow range; (3) principles not specific to FL. Some of the third factor principles have the flavor of the constraints that enter into all facets of growth and evolution, and that are now being explored intensively in the "evo-devo revolution."[4] Among these are principles of efficient computation, which would be expected to be of particular significance for generative systems such as I-language. Insofar as the third factor can be shown to be operative in the design of FL, explanation can proceed "beyond explanatory adequacy" in the technical sense, raising new

[3]Sean Carroll (2005). More generally, see Gould (2002, Part II).

[4]There are other third factor elements as well, among them properties of the human brain that determine what cognitive systems can exist, though too little is yet known about these to draw specific conclusions about the design of FL. It also might turn out that general cognitive principles that enter into language acquisition pose conditions on FL design. On the role of such principles in acquisition, see particularly Yang (2002).

questions: not only asking what mechanisms suffice to determine I-language from data available, but why these mechanisms should exist, and whether they are real or just dispensable descriptive technology. The task of accounting for the evolution of language would also be correspondingly eased, for the same reasons that hold for inquiry into evolution generally: the less attributed to genetic information (in our case, the topic of UG) for determining the development of an organism, the more feasible the study of its evolution.

Recent inquiry into these questions in the case of language has come to be called "the minimalist program" MP, but there has been so much misunderstanding, even within professional circles, that it is perhaps worth reiterating that it is a *program*, not a *theory*, and a program that is both traditional in its general flavor and pretty much theory-neutral, insofar as the biolinguistic framework is adopted. Traditional efforts to identify what is distinctive to FL have implicitly abstracted from third factor effects (and from generative processes as well, for the most part). And whatever one's beliefs about design of language may be, the questions of the research program arise. It may also be worth mentioning that the program can only be pursued, whatever theoretical framework one adopts, insofar as some descriptive account of the phenomena to be explained is reasonably unproblematic, often not the case of course, as expected with any system of at least apparent intricacy.

Throughout the modern history of generative grammar, the problem of determining the character of FL has been approached "from top down": How much must be attributed to UG to account for language acquisition? The MP seeks to approach the problem "from bottom up": How little can be attributed to UG while still accounting for the variety of I-languages attained, relying on third factor principles? The two approaches should, of course, converge, and should interact in the course of pursuing a common goal.

One useful way to approach the problem from below is to entertain the strong minimalist thesis SMT, which holds that FL is "perfectly designed." The first task would then be to formulate SMT coherently. The next would be to determine how close it is to true. Naturally, neither task is well-defined a priori, and each is sure to be modified in the course of inquiry. There are various ways to construe SMT, and any specific choice allows various paths that might be followed to investigate its reach. I would like to review where I think we stand after a few years of serious engagement with these issues adopting some choices that seem reasonable though certainly not logically necessary, to suggest a few refinements, and to indicate some of the manifold problems that arise in seeking to close the gap between SMT and the true

nature of FL.[5] UG is what remains when the gap has been reduced to the minimum, when all third factor effects have been identified. UG consists of the mechanisms specific to FL, arising somehow in the course of evolution of language.

An I-language is a computational system that generates infinitely many internal expressions, each of which can be regarded as an array of instructions to the interface systems, sensorimotor (SM) and conceptual-intentional (CI). To the extent that third factor conditions function, the language will be efficiently designed to satisfy conditions imposed at the interface; one can imagine more radical theses, to which I will briefly return. We can regard an account of some linguistic phenomena as *principled* insofar as it derives them by efficient computation satisfying interface conditions. We can therefore formulate SMT as the thesis that all phenomena of language have a principled account in this sense, that language is a perfect solution to interface conditions, the conditions it must at least partially satisfy if it is to be usable at all.

In its most elementary form, a generative system is based on an operation that takes structures already formed and combines them into a new structure.[6] Call it *Merge*. Operating without bounds, Merge yields a discrete infinity of structured expressions. Hence Merge, and the condition that it can apply without bound, fall within UG.

A Merge-based system will be compositional in general character: the interpretation of larger units at the interfaces will depend on the interpretation of their parts, a familiar observation in the study of every aspect of language. If the system is computationally efficient, once the interpretation of small units is determined it will not be modified by later operations – the general property of strict cyclicity that has repeatedly been found. Operations will also typically yield nested rather than crossing dependencies, also a familiar observation (and where crossing dependencies are found, it is commonly, and plausibly, taken to be the result of more complex processes). Thus in "The men who John V_1 V_2 ...," agreement universally holds between *John* and V_1 and between *the men* and V_2, not conversely. There is no obvious reason for this in terms of communicative or parsing efficiency; as is well-known, dependencies quickly overflow memory, so that language use adopts various methods that give it a

[5] I will assume here the general framework of my "On Phases," (to appear); and sources cited there.

[6] There are more complex possibilities, some familiar: e.g., phrase structure grammars, which fall within Emil Post's version of recursive function theory, abandoned for well-known reasons, empirical and conceptual. Another complication beyond pure Merge is adding the principles of associativity and ordering, suppressing hierarchy and yielding sequences.

paratactic flavor. But these familiar properties are an automatic consequence of generation relying on Merge with appropriate compositional conditions. One task of MP is to clarify and test these general ideas, and place them in a broader setting.

A Merge-based system of derivation involves parallel operations. Thus if X and Y are merged, each has to be available, possibly constructed by (sometimes) iterated Merge. The process has a loose resemblance to early theories of generalized transformations, abandoned in the early 1960s for good reasons, now resurrected in a far simpler form for better reasons. But a generative system involves no temporal dimension. In this respect, generation of expressions is similar to other recursive processes such as construction of formal proofs. Intuitively, the proof "begins" with axioms and each line is added to earlier lines by rules of inference or additional axioms. But this implies no temporal ordering. It is simply a description of the structural properties of the geometrical object "proof." The actual construction of a proof may well begin with its last line, involve independently generated lemmas, etc. The choice of axioms might come last. The same is true of generation vs production of an expression, a familiar competence-performance distinction. But even if one were to take the intuitive interpretation literally, generation of an expression is not strictly "bottom-up," because of the parallelism of operations. A strict "bottom-up" interpretation is, for example, compatible in principle with the assumption that in performance, the first XP (say a noun phrase) is produced or perceived first, even if later merged into some ultimately embedded expression (as internal or external argument, for example). Or many other assumptions about use of language.[7] In addition to Merge applicable without bounds, UG must at least provide atomic elements, lexical items LI, each a structured array of properties (*features*) to which Merge and other operations apply to form expressions. These features contain information relevant to the way their arrangements are interpreted at the interfaces: all information insofar as I-language satisfies the Inclusiveness Condition, a natural principle of efficient computation.[8] A particular language is identified at least by valuation of parameters and selection from the store of features

[7]E.g., perception models based on Bradley Pritchett's "bottom-up" theta-attachment model. For discussion, and exploration of new ideas and empirical results highly relevant to considerations here, see Reinhart (2006).

[8]The condition is radically violated in the mapping to the SM interface, even more so in strong versions of Distributed Morphology that take all phonological features of LIs to be inserted in this mapping. It is also violated in standard versions of mapping to CI, but that raises non-trivial questions about the architecture of cognitive systems, difficult to examine because of limited information about their language-independent nature.

made available by UG, and a listing of combinations of these features in LIs (the lexicon), satisfying further conditions that we put aside here. There is substantial evidence that human LIs are crucially distinct from the symbolic elements of other animals at both interfaces. At the CI interface, they lack the kind of relation to mind-independent elements that appears to be a general property of animal communication systems; something similar is taken for granted for phonological elements. If so, there is no *reference*-like relation for human language, hence no semantics in the technical sense of Frege, Peirce, Tarski, Carnap, Quine, and others, or contemporary "externalist" theorists of reference. The reasons have been discussed elsewhere and I will put them aside here, but if so, these are further genetically determined components of FL (or the conceptual resources on which it draws), and a problem to be addressed in study of evolution of language (or of the pre-linguistic conceptual resources available to humans).

In addition to such properties as these, UG must contain the principles that map external data to linguistic experience, providing the basis for language acquisition. The extent to which these properties and their organizing principles are unique to FL could be clarified by comparative studies, but there is little doubt that it is substantial.

The conclusion that Merge falls within UG holds whether such recursive generation is unique to FL or is appropriated from other systems. If the latter, there still must be a genetic instruction to use Merge to form structured linguistic expressions satisfying the interface conditions. Nonetheless, it is interesting to ask whether this operation is language-specific. We know that it is not. The classic illustration is "the mathematical capacity," which troubled Alfred Russel Wallace 125 years ago because it "is wholly unexplained by the theory of natural selection, and must be due to some altogether distinct cause," if only because it remained unused. One possibility is that it is derivative from language. If the lexicon is reduced to a single element, then Merge can yield arithmetic in various ways. Speculations about the origin of the mathematical capacity as an abstraction from linguistic operations are familiar, as are criticisms, including apparent dissociation with lesions and diversity of localization. The significance of such phenomena, however, is far from clear. They relate to use of the capacity, not its possession; to performance, not competence. For similar reasons, dissociations do not show that the capacity to read is not parasitic on the language faculty, as Luigi Rizzi points out.

Suppose the single item in the lexicon is a complex object, say some visual array. Then Merge will yield a discrete infinity of visual patterns, but this is simply a special case of arithmetic and tells us nothing about recursion in

the visual system. The same would be true if we add a recursive operation – another instance of Merge – to form an infinite lexicon on the model of some actual (if rather elementary) lexical rules of natural language, say an infinite array of visual patterns as "lexical items." Again that introduces nothing new, beyond FL. Similar questions might be asked about the planning systems investigated by George Miller and associates 45 years ago. If these and other cases fall under the same general rubric, then unbounded Merge is not only a genetically determined property of language, but also unique to it. Either way, it falls within UG as one of the organizing principles of recursive generation of expressions.

Merge(X_1,...,X_n) = Z, some new object. In the simplest case, $n = 2$, and there is evidence that this may be the only case (Richard Kayne's "unambiguous paths"). Let us assume so. Suppose X and Y are merged. Evidently, efficient computation will leave X and Y unchanged (the No-Tampering Condition NTC). We therefore assume that NTC holds unless empirical evidence requires a departure from SMT in this regard, hence increasing the complexity of UG. Accordingly, we can take Merge(X, Y) = {X, Y}. Notice that NTC entails nothing about whether X and Y can be modified after Merge.[9]

Suppose X is merged to Y (introducing the asymmetry only for expository reasons). Trivially, either X is external to Y or is part of Y: external and internal Merge, respectively; EM and IM (Move). In the latter case, X is not only a part of Y but necessarily a *term* of Y in the technical sense. Without further complication, Merge cannot create objects in which some object O is shared by the merged elements X, Y. It has been argued that such objects exist. If so, that is a departure from SMT, hence a complication of UG.[10]

If an element Z (lexical or constructed) enters into further computations, then some information about it is relevant to this option: at the very least, a property that states that Z can be merged, but presumably more, it is commonly assumed.[11] The optimal assumption is that this information is provided by a designated minimal element of Z, a lexical item W (Z itself, if it is an LI), which is detectable by a simple algorithm; the *label* of Z, the head projected in X-bar theories – possibly a dispensable notion, as discussed below. The label W of Z enters into EM in selection in various ways as well

[9]Let us put aside here the question whether in addition to "set-Merge" there is also an operation "pair-Merge," as discussed in my "Beyond Explanatory Adequacy," (Chomsky 2004).

[10]See Citko (2005). Also Svenonius (2005), on "banyan trees." Citko argues that parallel Merge is "predicted" as IM is, but that is not quite accurate. It requires new operations and conditions on what counts as a copy, hence additional properties of UG.

[11]Actually, a delicate question, having to do with interpretation of deviance at the interface.

as into interpretation of Z. Since W contains all information relevant to further computation involving Z, W is also necessarily the *probe* that selects a *goal* in any internal modification of Z. Minimal search conditions limit the goal of the probe to its complement, the smallest searchable domain. It is impossible, for example, for IM to move a SPEC of W (or one of its terms) to a higher SPEC position, or for an agreement relation to be established between W and an element within its SPEC. Or conversely, unless the SPEC is itself a head, an option barred under narrower phase-theoretic conceptions of the role of the probe in controlling operations.

Restricted to heads (probes), c-command reduces to minimal search. The standard broader notion can be defined in terms of dominance and sisterhood, and a choice has to be made between immediate dominance and some higher projection. But it is not clear that this extension beyond minimal search – a natural computational principle – is necessary. There seems no clear independent reason to believe that sisterhood is a relevant relation. Furthermore, to capture the intended asymmetry, both for EM and IM, choice of projection (labeling) must also be introduced into the definition. For such reasons, the broader notion departs from SMT, on reasonable assumptions, and requires empirical motivation. It has been assumed to be relevant to binding theory, hence presumably to the CI interface, but that may be unnecessary,[12] leaving scopal relations as possible instances of c-command in the broader sense. I know of no other evidence that it plays a role in narrow syntax or the mapping to CI. At the the SM interface side, the idea that the broader notion of c-command determines linearization is the core principle of Kayne's LCA and the very fruitful work it has inspired, and if the foregoing is correct, LCA can plausibly be interpreted as part of the mapping to the SM interface. That requires some device to deal with ordering of merged LIs, either (as in Kayne's work) a further elaboration of Merge and c-command to allow non-branching nodes, or some other departure from SMT, non-trivial it appears. Fukui and Takano review other stipulations that seem necessary, and argue in favor of a head parameter (for which they cite additional evidence, bearing on linear ordering in narrow syntax but not broader c-command). They do note one residue of LCA that is unaccounted for by a head parameter: the near universal SPEC-H ordering – which is narrowed to subject-H ordering unless second-Merge (hence SPEC) is banned within complex VPs and other such structures, a conclusion that is by no means obvious.[13]

[12] See Chomsky (to appear), and below. See also Hasegawa (*2005*), analyzing reflexivation in terms of multiple-agree.

[13] Fukui and Takano (1998). The head-parameter approach they adopt is developed in earlier

It is, however, not clear that the SPEC-H residue would qualify as support for LCA, because of the array of stipulations required to yield the result, some just reviewed. One might want to explore other directions. To mention one, it would not be implausible to seek a parsing account for properties of ordering; often justified (e.g., rightward displacement of complex phrases). One thought that might be pursued, for example, is that for a Pritchett-style parser based on theta-attachment (see note 7), if the external argument precedes the theta-assigner α, then all theta roles are available in the immediate projection of α when it is reached in linear search, simplifying the operations. Similar ideas might extend to the left periphery, on Rizzi-style assumption that the head carries the information about the status of the SPEC. Speculations aside, one general conclusion that seems clear is that LCA involves significant departures from SMT, and therefore must be supported by sufficient empirical evidence.

NTC has always been assumed without comment for EM: there is, for example, no proposal that if V and NP are merged to form VP, then V is merged inside NP. Under SMT, it should hold for IM as well. Assuming so, then an application of IM yields two *copies* of X.[14] There is no rule of formation of copies or remerge, as has sometimes been supposed; just IM applying in the optimal way, satisfying NTC. Repeated IM yields many copies. There must be a procedure to distinguish copies from independent repetitions; that is easily stated with a proper notion of cyclicity in terms of *phases*, to which we return: all and only repetitions formed by IM within a phase are copies.

In a well-designed FL, lacking arbitrary stipulations, both EM and IM should be permitted, and the two kinds of Merge should be expected to yield different interface properties. That is obviously true at the SM interface – the ubiquitous property of "displacement" – and appears to be true at CI as well. The two types of Merge correlate well with the duality of semantics that has been studied from various points of view over the years. EM yields generalized argument structure, and IM all other semantic properties: discourse-related and scopal properties. The correlation is close, and might turn out to be perfect if enough were understood. If so, the conclusions so far conform to SMT.

It also follows that it was a mistake – mine in particular – to suppose that displacement is an "imperfection" of language that has to be assigned to UG

work of Naoki Fukui, and Fukui and Mamoru Saito. Their primary example is head-final Japanese. For indirect but intriguing empirical evidence supporting LCA for Japanese, see Kayne (2004).

[14]Traces, indices, etc., are barred by NTC and Inclusiveness. Hence carry a considerable empirical burden.

or somehow explained in terms of its special functions. On the contrary, its absence would have to be accounted for by a UG stipulation barring IM. It therefore follows that some form of transformational grammar – by now a radically stripped-down version of early proposals – essentially "comes free." If some other device is developed to account for the empirical phenomena of displacement and associated interpretations, it will require a stipulation barring IM and further stipulation of the additional mechanisms, therefore facing a considerable empirical burden.

Unless an element Z is an isolated element (an interjection, or frozen expression), hence of no interest here, its label W must have a feature indicating that Z can be merged. Under NTC, merge will always be to the edge of Z, so we can call this an *edge feature* EF of W.[15] If EF is always deleted when satisfied, then all expressions will be of the form LI-complement; in intuitive terms, they branch unidirectionally. If EF is not deletable, then the elements of expressions can have indefinitely many specifiers (*complement* and *specifier* mean nothing more in this framework than *first-merged* and *later-merged*). Variation among LIs with regard to deletability of EF would be a departure from SMT, so we assume that for all LIs, one or the other property holds. Empirical evidence reveals that SPECs exist, that is, that EF is undeletable. That leaves the question why. SM interface conditions seem to be irrelevant, so we have to look to third factor effects and CI conditions. Both support the choice of undeletability. Only that choice permits IM, which comes free, so if expressive potential is to be used, EF must be undeletable. As for CI, undeletability provides for the duality of semantics.[16]

The property of unbounded Merge reduces to the statement that LIs have EF. The property has to be stated somehow, and this seems an optimal way. So far, then, the only syntactic properties of UG are that it contains Merge and LIs with undeletable EF, and that expressions generated must satisfy interface conditions – in a principled way, insofar as SMT holds.

CI clearly permits interpretation of quantification in some manner. Language should provide such a device if expressive potential is to be adequately utilized. There are various logically equivalent devices, among them variable-

[15]There are several interpretations of "merge to the edge," including a version of *tucking-in* in Norvin Richards's sense. I will put the matter aside here.

[16]As an uninterpretable feature, EF cannot reach the interface, so presumably deletion of EF is an automatic part of the operations of transfer. Note that the same cannot be assumed for the standard uninterpretable features, which can be deleted only when certain structural conditions are satisfied, and will crash the derivation otherwise. A reviewer points out that if EF is always deleted when satisfied, then one case of IM is permitted: self-merge of an LI, which is enough to yield the basic elements of arithmetic.

free logics. The most familiar notation is operator-variable constructions. But that device virtually comes free, given EM and IM expressing the duality of semantics at CI – which may be why it is the most commonly used formal device, and the easiest to learn. In the simplest case, the copy merged to the edge by IM is the operator taking scope over the copy that had previously been merged by EM, the latter understood as the variable; the full structure of the two copies provides the interpretation as a restricted variable, hence yields the options for reconstruction along lines that have been pursued very productively in recent years. These considerations take us a step towards establishing the A/A′-distinction as a property of language with a principled explanation in terms of SMT.

In the cases just mentioned, the apparent optimization of design is relative to the CI interface. That raises the question whether the examples are idiosyncratic in this respect or whether the property holds generally. If the latter, then the relation of the generative procedure to the interfaces is asymmetrical, CI taking precedence: optimization is primarily to the CI interface. The question can be approached on empirical grounds, from various directions. One is along the lines just illustrated: by investigating language design. The ways language deals with IM provide additional evidence of priority of the CI interface. As noted, NTC requires that all copies should be retained under IM: the initial copy is introduced by EM, and all others are introduced by IM. At the CI interface the conclusion is correct, at least to good approximation, as illustrated by reconstruction. It is, however, radically false at the SM interface, where all copies other than the final occurrence generated are deleted, with fairly systematic exceptions not relevant here. Here conditions of computational efficiency and of ease of communication are in conflict. Computational efficiency yields the universally attested facts: only the final position of IM is pronounced, dispensing with the need for multiple applications of the generally complex and language-variable operations of morphology-phonology (and whatever else may be part of the mapping to the SM interface).[17] But that leads to comprehension problems. For perception, major problems, familiar from parsing programs, are to locate the "gaps" associated with the element that is pronounced, problems that would largely be overcome if all occurrences were pronounced. The conflict between computational efficiency and ease of

[17]Note that the issue does not arise in the mapping to CI if, as generally assumed (and plausibly so, on "poverty of stimulus" grounds), it is universal, hence in effect instantaneous and costless. It must be the highest copy that remains or there will be no detectable evidence that IM applied overtly. The observations here refer to overt movement, but they generalize if we adopt the approach to overt/covert movement by Jon Nissenbaum (2000).

communication appears to be resolved, universally, in favor of computational efficiency to satisfy the semantic (CI) interface, lending further support to speculations about its primacy in language design.

There are other well-known cases where language design is dysfunctional for language use: island phenomena for example, which require circumlocution or special devices (e.g., resort to otherwise-barred resumptive pronouns) to allow expression of simple thoughts. Insofar as island phenomena can be reduced to design efficiency, they would lend further support to theses about primacy of the CI interface.[18]

The question can be approached from other directions too. Perhaps relevant are discoveries about sign languages in recent years, which provide substantial evidence that externalization of language is at least partially modality-independent. Among these are striking cases of invention of sign languages by deaf children exposed to no signing and by a community of deaf people who spontaneously developed a sign language. In the known cases, sign languages are structurally very similar to spoken languages, when the modality itself does not require differences. They also are reported to follow the same developmental patterns from the babbling stage to full competence. They are distinguished sharply from the gestural systems of the signers, even when the same gesture is used both iconically and symbolically, as Laura Petitto has shown. She and her colleagues have also studied children raised in bimodal (signing-speaking) homes, and have found no preferences or basic differences. Her own conclusion is that even "sensitivity to phonetic-syllabic contrasts is a fundamentally linguistic (not acoustic) process and part of the baby's biological endowment," and that the same holds at higher levels of structure. Imaging studies lend further support to the hypothesis that "there exists tissue in the human brain dedicated to a *function* of human language *structure* independent of speech and sound," in her words. Studies of brain damage among signers have led to similar conclusions, as has comparative work by Tecumseh Fitch and Marc Hauser indicating, they suggest, that the sensorimotor systems of earlier hominids were recruited for language but perhaps with little special adaptation. Similar conclusions about the primacy of the semantic interface

[18]Of interest in this connection is the investigation of interaction of syntactic structure and derivation with principles that facilitate communication, typically neo-Gricean, involving some form of "reference-set computation" (Reinhart 2006); see Chierchia (2004). A question that might be pursued is the extent to which these inquiries presuppose a pragmatic environment based on trust and intent to communicate effectively, as contrasted with one based on intent to deceive and mislead (or others). If the presupposition turns out to play a role, the ideas developed might be reinterpreted within interpretive components of thought, external to language strictly speaking, using its mechanisms in one rather than another way.

have been advanced by prominent evolutionary biologists. The ideas trace back to the cognitive revolution of the 17th century, which in many ways foreshadows developments from the 1950s.[19]

Generation of expressions to satisfy the semantic interface yields a "language of thought." If the assumption of asymmetry is correct, then the earliest stage of language would have been just that: a language of thought, used internally. It has been argued that an independent language of thought must be postulated. I think there are reasons for skepticism, but that would take us too far afield.

These considerations provide a very simple thesis about a core part of the evolution of language, one that has to be assumed at a minimum, so it would seem, by any approach that satisfies the basic empirical requirement of accounting for the fact that the outcome of this process is the shared human property UG. At the minimum, some rewiring of the brain, presumably a small mutation or a by-product of some other change, provided Merge and undeletable EF (unbounded Merge), yielding an infinite range of expressions constituted of LIs (perhaps already available in part at least as conceptual atoms of CI systems), and permitting explosive growth of the capacities of thought, previously restricted to the elementary schemata but now open to elaboration without bounds: perhaps schemata that allowed interpretation of events in terms of categorization by some property (hence predication, once Merge is available), actor-action schemata, and a few others that might well have earlier primate origins. Such change takes place in an individual, not a group. The individual so endowed would have the ability to think, plan, interpret, and so on in new ways, yielding selectional advantages transmitted to offspring, taking over the small breeding group from which we are, it seems, all descended. At some stage modes of externalization were contrived. Insofar as third factor conditions operate, UG would be optimized relative to the CI interface, and the mappings to the SM interface would be the "best possible" way of satisfying the externalization conditions. Any more complex account of the evolution of language would require independent evidence, not easy to come by; and some account is needed for any complication of UG that resists principled explanation. A common assumption of paleoanthropology is that emergence of language led to the "great leap forward" exhibited in the archaeological record very recently, and the spread of humans all over the

[19]Petitto (2005). Work of Hauser and Fitch cited in Hauser, et al. (2002). See my "Three Factors in Language Design," (2005), and "Some simple evo-devo theses: how true might they be for language?," in *Evolution of Human Language: the Morris Symposium*, SUNY at Stony Brook, October 2005, forthcoming.

world shortly after, all within an eye-blink in evolutionary time.

Various considerations, then, seem to converge rather plausibly on the conclusion that language may be optimized relative to the CI interface, with mapping to the SM interface an ancillary procedure, and complex to the extent that the SM interface has no prior adaptation to these needs. Insofar as SMT holds, generation of structures mapped to CI will be optimal for the CI interface and common to languages apart from parametric and lexical choices (phenomena that require explanation), while phonology, morphology, and whatever else is involved in externalization might be variable and complex and subject to large-scale historical accident, satisfying the linking condition in ways that are as good as possible. That is not a bad first approximation to what the study of language seems to yield.

A more radical conception of the FL-CI interface relation, developed by Wolfram Hinzen, is that "certain empirical properties of thought contents" derive from the structures generated optimally by FL: we are, for example, led to postulate propositions as "intermediate entities between what's in the head and what's out there in the physical universe" on the basis of the role of CP in syntactic generation and hence mapping to CI, so that we can "deflate" these mysterious entities "into the notion of a CP"; and the same with other postulated entities of thought. Thus optimally designed FL "provides *forms* that a possible human structured meaning may have, leaving a residue of non-structured meanings (concepts), a substantive amount of which we share with other animals that lack syntax (or at least do not use it, or do not use it for the purposes of language)." These forms are natural objects "that we can study as such, even though we see them, somewhat miraculously, systematically condition properties of linguistic meaning that we can empirically attest," a novel approach to what has been called "naturalization of meaning." It is "as if syntax carved the path interpretation must blindly follow" (quoting Juan Uriagereka) (Hinzen 2006, p.179, p.235, p.250). One might extend similar ideas to duality of semantics and other notions of the theory of meaning. From this perspective, propositions and other postulated entities of thought go the way of reference, eliminated from the theory of mind and language. The primacy of CI is reduced, though satisfaction of CI conditions cannot be entirely eliminated: CI must have some range of resources that can exploit the properties of generated expressions, along with whatever is involved in use of language to reason, refer, seek to communicate perspicuously, and other mental acts. SMT and the concept of principled explanation would be correspondingly simplified.

Returning to the main track, what further properties of language would

SMT suggest? One is a case of Occam's razor: linguistic levels should not be multiplied beyond necessity, taking this now to be a principle of nature, not methodology, much as Galileo insisted and a driving theme in the natural sciences ever since. We are assuming that FL provides at least instructions for the CI and SM interfaces, the former having priority (perhaps near-tautologically, insofar as the more radical thesis can be sustained). But postulation of any linguistic levels beyond that departs from SMT, and requires justification. Others are postulated in familiar conceptions of language. Thus in versions of EST (the "Y-model"), three internal levels are postulated, each with its specific properties: D-structure, S-structure, and LF.[20] Reliance on Merge as the sole operation dispenses with D- and S-structure, in fact, renders them unformulable (the same with any other notions of underlying and surface structure). It has to be shown that nothing is lost (or better, that something is gained) by this simplification. That appears to be true. If so, we are left only with the internal level LF.

As noted, Merge yields compositional/cyclic properties of the kind that have repeatedly been found. Optimally, there should be only a single cycle of operations. EST postulated five separate cycles: X-bar theory projecting D-structure, overt operations yielding S-structure, covert operations yielding LF, and compositional mappings to the SM and CI interfaces. With the elimination of D- and S-structure, what remains are three cycles: the narrow-syntactic operation Merge (now with overt and covert operations intermingled), and the mappings to the interfaces. As noted earlier, optimal computation requires some version of strict cyclicity. That will follow if at certain stages of generation by repeated Merge, the syntactic object constructed is sent to the two interfaces by an operation Transfer, and what has been transferred is no longer accessible to later mappings to the interfaces (the phase-impenetrability condition PIC). Call such stages *phases*. Optimally, they should be the same for both subcases of Transfer, so until shown otherwise, we assume so (the mapping to the SM interface is sometimes called "Spell-Out"). LF is now eliminated, and there is only a single cycle of operations. The cyclic character of the mappings to the interfaces is largely captured, but not completely: there may be – and almost certainly are – phase-internal compositional operations within the mappings to the interfaces. And with phases in place, the problem of distinguishing copies from repetitions is resolved, since all copies are formed by IM at the phase level, hence identifiable for Transfer (the same observation

[20] I am using LF in the sense defined in EST: the output of narrow-syntactic operations and the input to the mapping to CI. Commonly the term has come to be used to refer to the output of that mapping, so that some other term is needed for LF, within this model or others like it.

extends to successive-cyclic movement). Whatever phases are, it is clear that PIC is restricted to the complement of the phase head, since specifiers of the phase label/head P can be raised in later phases, P as well.

Still keeping to SMT, all operations are driven by labels. Since at least some operations are restricted (by definition) to the phase head, the simplest assumption would be that all operations apply there. But there must be a departure from this assumption: syntactic objects cannot expand unless EM applies at every stage of derivation. The simplest conclusion, then, seems to be that operations other than EM all apply at the phase level. There is considerable evidence to support that conclusion, which I will adopt here. It follows at once that CP must be a phase, since A′-movement is to its edge, a specifier of the label/probe C (a cover term for the elements of Rizzi's left periphery). If the conclusions about duality of semantics and the IM-EM distinction are correct, then C must be the locus of A′-movement to capture scopal properties, so that the phasal character of CP may follow directly from CI-interface conditions and optimal computation, hence SMT. The general line of argument seems clear enough and plausible, though there are gaps to fill.

Phases should be as small as possible, to maximize the effects of strict cyclicity, hence computational efficiency. Let's adopt the (fairly conventional) assumption that verbal phrases are of the form v-VP, where v can be v*, the functional category that heads verb phrases with full argument structure, unlike unaccusatives and passives. Possibly the functional category v determines the verbal character of the root R that is its complement, along lines discussed by Alec Marantz (1997), in which case verbal phrases are of the form v-RP. Problems arise if phases are associated with every operation of Merge – e.g., with VP (or RP). One reason is that at VP, information is not available as to whether the complement of V will be spelled out in situ or raised by IM, or what its structural Case will ultimately be (so that crash at both interfaces is inevitable). Whether similar conclusions hold at the CI level depends on murky questions as to how argument structure is assigned. For example, can additional material (subject, PP, etc.) determine the semantic relation of V-NP (or R-NP)? Take, say, "(the teacher) left the class (with a problem to solve)," "the class left." Under the most natural mechanisms of argument assignment, it is not obvious that the semantic relation of "leave" and the two nominal phrases is determined at the V-NP level. And there are approaches to far more intricate cases for which the assumption appears to be radically wrong.[21] If VP is not transferable to CI, then for unaccusative/passive (and probably many

[21] For review, analysis, and sources, see Marantz (2005).

other structures), the smallest domain within which the V-object relation can be assigned its semantic (theta) role is above vP (in fact CP, with TP still awaiting discussion); and for others the smallest domain is v*P.

Another line of argument that reaches the same conclusions is based on uninterpretable features: structural Case and redundant agreement. Since the values of these features are determined by context, the simplest assumption is that they are unvalued in the lexicon, thus properly distinguished from interpretable features, and assigned their values in syntactic configurations, hence necessarily by probe-goal relations. Keeping to structural NOM-ACC, NOM (and associated agreement) is assigned at least as high as TP (in fact, CP, we conclude below), and ACC (and associated agreement) is assigned within v*P, independently of the choice of higher Case (e.g., *(for him) to accept the job, accept the job!*, *(his, him) accepting jobs*). On reasonable assumptions that have been familiar since Vergnaud's original ideas on structural Case, valuation always takes place though it is only sometimes manifested. Accordingly, Case-agreement relations are fixed in the configurations v*P and CP, hence by the probes v* and C-T.[22] Object agreement is within this configuration.

Valuation of uninterpretable features clearly feeds A′-movement (e.g., in "whom did you see?"). Hence valuation is "abstract," functioning prior to transfer to the SM interface, as are the uninterpretable features themselves. A-movement is a more subtle case, but the same conclusion appears to hold. Evidence for that conclusion is provided by Eric Reuland's discovery of locality-bound (hence syntactic) reflexivization in which the antecedent does not c-command the reflexive but both are c-commanded by the head that agrees with the antecedent: structures of the form [T...XP...R], where T and XP agree, XP does not c-command R, both XP and R are in the search domain of C-T, and XP binds R – indirectly via the common probe C-T. Again, this must be prior to transfer to the SM interface, hence "abstract."[23]

If transferred to the interface unvalued, uninterpretable features will cause the derivation to crash. Hence both interface conditions require that they cannot be valued after Transfer. Once valued, uninterpretable features may or may not be assigned a phonetic interpretation (and in either case are eliminated before the SM interface), but they still have no semantic interpretation. Therefore they must be removed when transferred to the CI interface. Furthermore, this operation cannot take place after the phase level at which they

[22] On generalization to what he calls "stem features," either agreement or focus, see Miyagawa (2006).

[23] Reuland (2005). See also Legate (2005). Principles of phonetic realization of uninterpretable features are a different matter, though not unrelated.

are valued, because once valued, they are indistinguishable at the next phase level from interpretable features, hence will not be deleted before reaching the CI interface. It follows that they must be valued at the phase level where they are transferred, that is, at the point where all operations within the phase take place and the Transfer operation therefore "knows" that the feature that has just been valued is uninterpretable and has to be erased at (or before) CI. Since all operations take place at the phase level, there is no memory or search problem.[24] It follows again that v*P must be a phase along with CP.

The next question is whether TP is also a phase, as is suggested by surface phenomena of valuation of uninterpretable features and A-movement. From observations of Marc Richards, it follows that the PIC entails that TP cannot be a phase, with operations of valuation and A-movement driven by properties of T.[25] Suppose TP were a phase. Then its interior will be transferred by PIC, but the head T will retain its valued uninterpretable features. The derivation will therefore crash at the next phase, for the reasons just given. Hence the relevant phase for these operations must be CP, not TP. It is, therefore, not only unnecessary but incorrect to add an additional phase TP – the preferred outcome on grounds of computational efficiency, obviously.

For the same reason, Richards points out, the uninterpretable features of C must be "inherited" by T. If they remain at C, the derivation will crash at the next phase.[26] Note that TP cannot be saved as a phase by the same device: if its features are inherited by v*, the derivation will always crash because the external argument is outside the search domain of v*.

From elementary conceptual considerations then, plausibly traceable to SMT, we conclude that v*P and CP are the phases of the clausal skeleton, and that the uninterpretable features of C are assigned to T, which does not head a phase.

---

[24] Alternatives can be devised, but all involve additional search and memory, hence are barred unless empirical evidence is provided for this departure from SMT.

[25] Personal communication, August 2005. See Richards (2006).

[26] A residue may remain at C for phonetic interpretation at the next phase, as sometimes found. That is worth exploring in detail, in a variety of language types, but appears to require only minor technical adjustment. It might be that what appears phonetically at C, in some cases at least, is the result of subsequent concord, not agreement. The principled issues arise at CI. Richards's argument supports the conclusion about inheritance in "On Phases," (Chomsky to appear) there relying partly on empirical observations based on sometimes subtle judgments with interaction of several factors, hence less compelling than Richards's conceptual argument. It is sometimes felt intuitively that "inheritance" is counter-cyclic, but technically that is not the case, any more than the (somewhat similar) probe-goal relation that determines structural Case in situ, for example).

There are further reasons for expecting that TP is not a phase. T has the basic properties of uninterpretable features. It may yield a phonetic reflex, but its ϕ-features are determined by its context, so it should enter the lexicon without values for these features. T bears these features if and only if it is selected by C, hence it should inherit these from C (the precise mechanism does not matter here). The biconditional holds of embedded clauses, but it would make no sense to hold that in root clauses T has different properties. It therefore follows that root clauses must have C, even if it is unpronounced, as is also indicated by other phenomena; e.g., clausal operators in A′-positions outside TP, hence SPEC-C.

What is true of agreement features appears to hold as well for tense: in clear cases, T has this feature if and only if it is selected by C, though C never (to my knowledge) manifests Tense in the manner of ϕ-features in some languages. If that is basically accurate, then there are two possibilities. One is that Tense is a property of C, and is inherited by T. The other is that Tense is a property of T, but receives only some residual interpretation unless selected by C (or in other configurations, e.g., in English-like modal constructions).[27] One advantage of the latter option is that T will then have at least some feature in the lexicon, and it is not clear what would be the status of an LI with no features (one of the problems with postulating AGR or other null elements). Another advantage would be an explanation for why C never manifests Tense in the manner of ϕ-features (if that is correct). Under the former option, with Tense inherited by T, Richards's argument does not independently apply, because tense is interpretable. His argument would also apply, however, if the mechanism of inheritance is generalized (that is, simplified) to all inflectional features of C, not just ϕ-features.

For the same reasons, the inheritance mechanism is simplified if it is generalized to phase heads generally, not restricted to C but extended to v* as well. But as Richards observes, that is necessary anyway, for the same reasons that require that C assign its features to T. Therefore V (or R) must receive ϕ-features from v*. It follows that just as a nominal phrase can raise to SPEC-T within CP, so it should be able to raise to SPEC-V within v*P. There is good evidence for that, going back to work of Paul Postal's on "raising to object" 30 years ago, reformulated and extended by Masatoshi Koizumi, Howard Lasnik, and Mamoru Saito.[28] I personally resisted their evidence and tried to find ways

[27] It would not suffice to have T bear tense with a condition that C must select tense, since that would leave the possibility of tensed T without C, which is impossible in embedded clauses (and root clauses too, for the reasons just given).

[28] For discussion and sources, see Lasnik (2003). These approaches assume raising to

to evade it for some years, because the operation appears to make no sense. It has no visible effect, since V raises to v*, restoring the original order; and there is no semantic motivation at all, though there are semantic consequences. But we now see that there is in fact motivation for this strange and purposeless operation, with its scopal and binding consequences; namely, it follows from SMT. These curious phenomena thus yield further support to the idea that FL may indeed be well-designed to satisfy CI interface conditions (or more radically, that these conditions in part simply reflect SMT).

Note an asymmetry, at least for the languages we are now considering: T may or may not raise to C, but V must raise to v*, which therefore is an affix. There are other asymmetries: what we are calling V has semantic content, and may simply be a root with v* serving as the functional element determining its category, as mentioned earlier. In contrast, C has independent content and is not categorizing a root (and whether T has semantic content remains open). In other languages v* (and v generally) has morphological content, perhaps always affixal. A broader range of languages should, obviously, be considered seriously before conclusions can be drawn about relations among these properties.

Assume the predicate-internal subject hypothesis, which is quite natural on conceptual and semantic grounds: argument structure is determined by EM, in terms of relations within the verbal (or predicate) phrase. For a subject with uninterpretable Case, the value must be assigned by some higher phase head (which will also permit it to undergo A-movement). As we have just seen, that will always cause the derivation to crash unless there is a head selected by C which can inherit its inflectional features, namely T. Again, consequences follow at both interfaces. Displacement to SPEC-T is permitted, with phonetic effects as well as familiar consequences for scope, weak crossover, anaphora, and discourse-related "edge" properties. Problems also remain. It appears that at least some element must raise from v*P, but if so, the reasons remain controversial.

Richards's observation also provides an argument as to why T should exist at all. Uninterpretable features of C must be inherited by an element selected by C, for his reasons, but it cannot be v*, for the reasons mentioned. Therefore T or some counterpart must exist, selected by C and above v*. Why then should T appear in clauses not selected by C: ECM and raising constructions? A possibility is along the lines already suggested for other cases. The UG principle that inserts T before vP is generalized, thus preventing automatic

AGR-O, not V, but it is not clear that the extra assumptions are necessary.

crash at a later stage if C is merged by EM.[29]

The ϕ-features inherited by T probe for agreement and Case-assignment, but a question arises as to whether that happens before or after they are inherited from C, at the phase level. If raising is contingent on the probe-goal relation (as seems plausible from intervention effects), then the inheritance operation must precede probe by the ϕ-features (putting aside here complications about separate probing for person and number), so that T serves as the probe at the C level, not C. Otherwise, there will be no A-movement, contrary to what is empirically observed. There might be a reason for this ordering in terms of computational efficiency: the ordering inheritance-probe yields shorter search by the probe and shorter raising. If reasoning along these lines is tenable, then the A/A′-distinction would follow on computational grounds, yielding the basis for duality of semantics.

Let's look a little more closely at the general character of the mechanisms involved, which can be made precise in various ways – keeping here to NOM-ACC languages, and abstracting from interference effects, quirky Case, double objects, and other complications.

Consider a single phase of the schematic form {P, XP}, where P is the phase head, C or v*. P assigns its inflectional features to the label L of XP, T or V. These labels then probe XP to find the closest matching goal. For P = v*, it is the object of V, subject being outside the search domain even if v* itself is the probe, not V. For P = C, it is either the subject of v*P, or the object of V if v is unaccusative/passive. The uninterpretable features of L receive the values of the goal, which is assigned Case in accord with the properties of P: NOM if P = C, ACC if P = v*. If there are several goals, then all are valued in the same way, as is the probe. Thus in a participial construction of the form [P L (participle)^n object], Case of the object is NOM if P = C (and L = T), ACC if P = v* (and L = V; possibly an ECM construction). P and any participles have the inherent inflectional features of the object and the participles share its Case (presumably an option because of the categorial nature of participles).

Probe-goal agreement may or may not be accompanied by IM. If it is not, then the goal is realized in-situ[30] ; if it is, then the goal moves step-by-step as far as it can, reaching SPEC of the probe that has inherited ϕ-features from the phase head.[31] The intermediate copies reach the CI interface and can have

[29]Presumably control structures are CPs. The status of small clauses raises independent questions.

[30]Unless raised for some other reason. See Chomsky (2001). It is not unlikely, I think, that the observations extend to English constructions involving inversion.

[31]In this case at least, no recourse to the activity condition is needed. See Nevins (2004).

semantic effects: for binding, as in "John seems to her [*John* to appear to X to have left]" (X = himself, her, *herself, *him, *Mary (by Condition (C))]; for scope, "Every child doesn't *every child* seem to his father [to be smart]," with "every child" binding "his" but with scope below negation;[32] in both cases with lower copies italicized. These are basic properties of A-movement.

Note that the notion "label" is playing only an expository role here. In constructions of the form H-XP (H a head), minimal search conditions determine that H is the designated element (label) that enters into further operations. H will be the probe, and wherever selection enters – possibly only at the CI interface – H is the only functioning element, whether selecting or selected. Questions arise about labeling only for XP-YP constructions. For IM, with XP raised from YP with head Y, Y is the probe, and the simplest assumption is that it remains the probe, so that XP need not be searched for a new probe.[33] The most problematic case is XP-YP formed by EM. The primary example is when XP is an external argument EA. But numerous questions arise about the correct analysis (e.g., why does the full v*P including EA never raise, or remain without something extracted from it? Is EA in situ a real phenomenon? Why do idioms typically exclude EA?). It may be that as understanding progresses, the notion "label" will remain only as a convenient notational device, like NP, with no theoretical status.[34]

In any event, reference to labels (as in defining c-command beyond minimal search) is a departure from SMT, hence to be adopted only if forced by empirical evidence, enriching UG.

Consider the Reuland local anaphora cases of the form [C [T...XP...R]], where T and XP agree, XP does not c-command R, and both XP and R are in the local search domain of T. T inherits features from C, and the features of T, XP are valued by the probe-goal relation. TP is transferred to CI, erasing the uninterpretable features that have been valued within the CP phase. But these features cannot be erased before they reach CI, because the T-R relation establishes anaphora – that is, an interpretation by CI making use of the structure presented to it at the interface (much as articulators follow "instructions" given in the phonetic form). Thus there is a strong sense in

[32] Sauerland (2003). The consequences Sauerland draws for phase theory do not apply under the assumptions of "On Phases" (Chomsky to appear). For many problems, see Howard Lasnik, "On a Scope Reconstruction Paradox." reprinted in Lasnik (2003).

[33] Unless XP itself is a head, in which case there is an ambiguity. See "On Phases" (Chomsky to appear) and sources cited.

[34] For a different approach to similar questions, see Collins (2002), with the notion locus replacing label. But neither may be needed.

which transfer to CI at the phase level is "instantaneous," with the structure mapped to CI preserved for CI interpretation. Mapping to the SM interface is sharply different, as well known, another CI-SM asymmetry, consistent with earlier conclusions.

It seems that basic properties of A-movement fall out fairly naturally, though many questions remain unanswered. Among them are the perennial problems of EPP and of why language should have uninterpretable features at all, both involving serious residual problems, which are left as UG stipulation if they receive no principled answer.[35] One approach to the existence of uninterpretable features might be to consider more carefully their consequences. For reasons just discussed, they compel phases to be as small as possible consistent with IM and (possibly) assignment of argument structure, CP and v*P, and they impose cyclicity of Transfer (strict cyclicity, given PIC), thus reducing memory load in computation. Hence they contribute to SMT. They also signal anomaly of derivation quickly, without recourse to selectional features of lexical items that are more naturally understood as part of the interpretive processes at CI. Thus the eventual anomaly of, say, "how many trains did you say that John arrived t" is detected by CI at the earliest possible stage, the lowest v*P. That could be a factor expediting efficient performance.

Many other questions arise when we look more closely at how the principles function; for example, what are the intermediate positions for A-movement?[36] The product of A-movement is an expression with one or more copies of an element that initially entered the computation by EM, receiving its role in argument structure. The collection of these copies is an A-chain (more precisely, a set of occurrences, but we can put the refinement aside). But more is needed. We have assumed that CI permits interpretation of quantification, so that language must generate expressions that yield such interpretation. As discussed earlier, one way to yield such interpretations "comes free," without stipulation: an operator in SPEC-C taking scope over an A-chain regarded as a variable, with restrictions given by the content of the copies. It must be, then, that the edge feature EF of a phase head P can seek a goal in the complement of P, which it can raise to SPEC-P (perhaps covertly). A-movement is IM contingent on probe by uninterpretable inflectional features, while A′-movement is IM driven by EF of P. Like A-movement, A′-movement proceeds step by step, leaving reconstruction sites. Unlike A-movement, it

[35]For a suggestive approach to partial reduction of EPP to general cognitive principles and "canonical surface forms," see Bever (to appear). His proposal bears on expletive-insertion, but not on the more general question of raising to SPEC-T.

[36]For inconclusive discussion of some of these, see "On Phases." (Chomsky to appear)

proceeds successive-cyclically phase by phase, an option because SPEC-P is not transferred at the P-level. That makes broad scope interpretations possible, and opens many questions that are discussed elsewhere.[37]

Consider the interaction of A and A′-movement with regard to improper movement:[38]

(IM) * who [t_1 seems [t_2 C [t_3 T-is t_4 smart]]] (t a copy of *who*).

At the lowest CP, t_4 is Case-marked by C-T and raises by A-movement to t_3. It is also the goal of EF(C), hence raises directly from t_4 to SPEC-C (that is, t_2). There is no defined relation between t_2 and t_3. But t_2 is invisible to C-T of the next higher phase, because it has been inactivated within the lower CP. Therefore A-movement to t_1 is impossible.

While much remains open, at least the general properties of A and A′-movement appear to be within the range of principled explanation.

One might speculate that nominal phrases have structures similar to verbal phrases, and might sometimes also constitute phases. What are intuitively nominal phrases come in two basic varieties, ±definite (maybe specific – put that aside), differentiated by presence or absence of an element with some such property as "referentiality" (meaning "used to refer," not "referential expression," a crucial difference). We may take this element to be D – assuming that D has some real meaning.

Consider first indefinite nominals, lacking D, like "author" or "many authors." The label of the latter cannot be "many," which is not an LI but an XP, so in both cases the label of the phrase must be the label of "author" (which could have a complement, as in "author of the book"; note that its structure differs from one of the options for "picture of the book," with the counterpart "the/a picture, which is of the book"). The best theory, if feasible, would not add any additional distinguishing elements. Assuming that the basic structure corresponds to verbal phrases, the head will be n with the complement [X (YP)] (X perhaps an undifferentiated root, gaining its nominal character from n). X raises to n, just as its counterpart raises to v in the verbal phrase, and the result is a nominal phrase.

Assuming the same correspondence for definite nominal phrases, the head is now n* (analogous to v*) with the complement [X (YP)]. In this case X = D. D inherits the features of n*, so YP raises to its SPEC, and D raises to n*, exactly parallel to v*P. Therefore, the structure is a nominal phrase headed by n*, not a determiner phrase headed by D, which is what we intuitively always

[37] "On Phases", and sources cited.

[38] I adopt here observations of Samuel Epstein (personal communication), adapted to the version of phase theory here.

wanted to say; and D is the "visible" head, just as V is the "visible" head of verbal phrases. The structure is similar to a causative verb phrase analyzed with head v* and complement {cause, {V, complement}}. The complement of *cause* raises to specifier of the element *cause*, which inherits the features of v* and then raises to v*. Further outcomes depend on morphology.

That looks as though it might be roughly on the right track. Both DP and NP are nominal phrases, the natural result. It could be that only the constructions with D are phases, perhaps the source of extraction differences of definite vs indefinite NPs and other properties.[39]

So far, I have kept fairly close to what seems to be a reasonable interpretation of SMT, with assumptions about third factor properties of efficient computation and CI conditions that seem plausible, and can be investigated in independent ways. Just how far this line of inquiry can reach, of course one cannot know. As it proceeds, it approaches more closely the goal of identifying the principles of UG, the residue that constitutes FL once third factor properties of growth and development are extricated, along with others not considered here. The approach proceeds in parallel with a different line of inquiry into UG, the standard one for the past half-century, based on search for explanatory adequacy in terms of conditions for language acquisition: the "bottom-up" and "top-down" approaches to UG discussed earlier. Insofar as the two lines of inquiry converge, they clarify the questions that have been at the heart of theoretical study of language since its origins: to determine the basic properties of FL – a certain biological system, adopting the biolinguistic perspective, apparently a distinctive and crucial component of human nature.

[39]It would follow that structural Case is on n*, not D or N (hence presumably also on n), or the derivation will crash at the phase level.

References

Belletti, Adriana (ed.)
2004 *Structures and Beyond. The Cartography of Syntactic Structure* Vol.3. OUP, Oxford.

Bever, Thomas
to appear Minimalist behaviorism: The role of the the individual in explaining language universals. In Morten Christiansen, Chris Collins, and Shimon Edelman. (eds.), *Language Universals*, Oxford University Press, Oxford, UK.

Carroll, Sean
2005 *Endless Forms Most Beautiful.* Norton, New York.

Chierchia, Gennaro
2004 Scalar Implicatures, Polarity Phenomena, and the Syntax/Pragmatics Interface. In Belletti (2004), pp. 39–103.

Chomsky, Noam
1965 *Aspects of the Theory of Syntax.* MIT Press, Cambridge, Mass.
1980 *Rules and Representations.* Columbia University Press, New York.
1995 *The Minimalist Program.* MIT Press, Cambridge, Mass.
2001 Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A Life in Language*, pp. 1–52. MIT Press, Cambridge, Mass.
2004 Beyond Explanatory Adequacy. In Belletti (2004), pp. 104–131.
2005 Three Factors in Language Design. *Linguistic Inquiry*, 36: 1–22.
to appear On Phases. In Robert Freidin, Carlos Otero, and Maria-Luisa Zubizarreta (eds.), *Foundational Issues in Linguistic Theory.* MIT Press, Cambridge, Mass.

Citko, Barbara
2005 On the Nature of Merge: External Merge, Internal Merge, and Parallel Merge. *Linguistic Inquiry*, 36: 475–496.

Collins, Chris
2002 Eliminating Labels. In Samuel Epstein and Daniel Seely (eds.), *Derivation and Explanation in the Minimalist Program*, pp. 42–64. Blackwell, Oxford.

Fukui, Naoki and Yuji Takano
1998 Symmetry in Syntax: Merge and Demerge. *Journal of East Asian Linguistics*, 7: 27–86.

Gould, Stephen
2002 *The Structure of Evolutionary Theory.* Harvard.

Hasegawa, Hiroshi
2005 Reflexive Binding as Agreement and its Interaction with the Phase System. In Noriko Imanashi (ed.), *The World of Linguistic Research. A Festschrift for Kinsuke Hasegawa on Occasion of his Seventieth Birthday*, pp. 53–69. Kaitakusha, Tokyo.

Hauser, Marc, Noam Chomsky, and Tecumseh Fitch
2002 The faculty of language. What is it, who has it, and how did it evolve? *Science*, 298: 1569–1579.

Hinzen, Wolfram
2006 *Mind Design and Minimal Syntax.* OUP, Oxford.

Kayne, Richard
2004 Antisymmetry and Japanese. In Lyle Jenkins (ed.), *Variation and Universals of Biolinguistics*, pp. 3–35. Elsevier, Oxford.

Lasnik, Howard
2003 *Minimalist Investigations in Linguistic Theory.* Routledge, London.

Legate, Julie Anne
2005 *Morphological and Abstract Case.* Manuscript, University of Delaware, Newark, Delaw.

Marantz, Alec
1997 No escape from syntax. In: Alexis Dimitriadis, Laura Siegel, Clarissa Surek-Clark, and Alexander Williams (eds.), *Proceedings from the 21st Annual Penn Linguistics Colloquium*, pp. 201–225. Penn Working Papers in Linguistics, University of Pennsylvania, Philadelphia, Penn.

Marantz, Alec
2005 *Rederived Generalizations.* Manuscript, MIT, Cambridge, Mass.

Miyagawa, Shigeru
2006 *Strong Uniformity: A Non-Parametric Theory of the Faculty of Language.* Manuscript, MIT, Cambridge, Mass.

Nevins, Andrew
2004 *Derivations without the Activity Condition.* Manuscript, MIT, Cambridge, Mass.

Nissenbaum, Jon
2000 *Investigations of Covert Phrase Movement.* Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, Mass.

Petitto, Laura-Ann
2005 How the Brain Begets Language. In James McGilvray (ed.), *The Cambridge Companion to Chomsky*, pp. 84–101. CUP, Cambridge.

Reinhart, Tanya
2006 *Interface Strategies.* MIT Press, Cambridge, Mass.

Reuland, Eric
2005 Agreeing to Bind. In Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinhenz, and Jan Koster (eds.), *Organizing Grammar. Studies in Honor of Henk van Riemsdijk*, pp. 505–513. Mouton de Gruyter, Berlin.

Richards, Marc
2006 *On Feature-Inheritance: An Argument from PIC.* Manuscript, University of Cambridge

Sauerland, Uli
2003 Intermediate Adjunction with A-Movement. *Linguistic Inquiry*, 34: 308–314.
Svenonius, Peter
2005 *Extending the Extension Condition to Discontinuous Idioms*. Manuscript, CASTL, Tromsø.
Yang, Charles
2002 *Knowledge and Learning in Natural Language*. OUP, Oxford.

The Subject-In-Situ Generalization Revisited*

Artemis Alexiadou and Elena Anagnostopoulou

1 Goal

Alexiadou and Anagnostopoulou (2001; henceforth A&A 2001) proposed that the condition in (1) regulates the availability of vP-internal subjects and objects across languages. The aim of this paper is to re-examine the status of (1) in view of recent developments in syntactic theory.

(1) *The subject-in-situ generalization (SSG)*
By Spell–Out, vP can contain only one argument with a structural Case feature[1].

We argue that (1) is a more general condition than previously recognized. Based on a comparison between Indo–European (IE) and Khoisan languages, we argue that (1) is a universal principle that regulates argument externalization. Our analysis of the phenomena involved will suggest that the condition in (1) forces dislocation of arguments as a consequence of a constraint on Case checking. We will relate this type of dislocation to the EPP, thus viewing the EPP as a principle, regulating argument movement in general. This means that the EPP should no longer be viewed as the guiding force of the movement of the external argument to Spec,TP, but rather EPP features provide landing sites for arguments escaping the condition in (1). On this conception, the EPP forces movement of either the subject or the object, which is independent of the requirement that T bears an EPP feature.

*A preliminary version of this paper was presented at the workshop on "Interfaces + Recursion = Language? The View from Syntax and Semantics" held in Berlin, March 24, 2005. We would like to thank the audience in Berlin, Chris Collins, and Winfried Lechner for comments and discussion. Thanks are also due to one anonymous reviewer and the editors of the volume for their suggestions.

[1]The original formulation is as follows:
(i) By Spell–Out VP can contain no more than one argument with an unchecked Case feature.
See the formulation (1') in section 2.3.

2 Background on the SSG

We will begin the discussion by reviewing the evidence that led A&A (2001) to formulate (1). More specifically, we will present a number of constructions where subjects and objects with structural Case are not allowed to both remain in their base position: one of them must leave the vP.

2.1 Motivating the SSG: Stylistic Inversion and Quotative Inversion

As is well known, in French and English there is a transitivity restriction on subject inversion in constructions containing an expletive subject. While expletive constructions are well–formed with intransitive verbs (2a), (3a), transitive expletive constructions are ungrammatical (2b), (3b):

(2) a. il est arrivé un homme √expl–VS
EXPL is arrived a man
'There has arrived a man'
b. *il a lu un élève le livre *expl–VSO
EXPL has read a student.NOM the book.ACC

(3) a. there arrived a man √expl–VS
b. *there finished somebody the assignment *expl–VSO

It is generally agreed upon that the inverted subjects remain in vP-internal positions (see Bobaljik & Jonas 1996; Déprez 1991 and references therein). In these languages, there are constructions where the subject can remain vP-internal with transitive predicates. These constructions involve movement of the object to a position outside the vP. These are stylistic inversion in French and quotative inversion in English.

2.1.1 Stylistic Inversion

Stylistic Inversion (SI; see Kayne & Pollock 1978; Déprez 1991; Collins & Branigan 1997; Watanabe 1996, among many others) involves postposing of the subject in wh–questions, relative clauses and subjunctive sentential complements (see Déprez 1991: pp. 48–49). We mainly discuss wh–environments here:

(4) a. Je me demande quand partira Marie
I wonder when will.leave Mary
'I wonder when Mary will leave'

b. Les resultats que nous donnent ces experiences
the results that us give these experiments

SI is inapplicable when no wh-movement takes place:

(5) *partira ton ami
will.leave your friend
'Your friend will leave'

Déprez (1991) argues that in SI the subject remains vP-internal. Her arguments include *combien*-extraction, stage vs. individual level subjects and floating quantifiers. We briefly review the latter argument here.

In French there are (roughly) two positions in which floating quantifiers surface: following a tensed verb or following an auxiliary. Sportiche (1988) has argued that floating quantifiers are stranded quantifiers which mark the original DP position of the subject.

(6) *Qu'ont tous fait les enfants?
'what did the children all do'

If postverbal subjects in SI do not undergo movement, then the distribution of stranded quantifiers is correctly predicted. The ungrammaticality of (6) is not expected under an analysis of SI in terms of rightward movement of the subject.

SI is disallowed when the vP contains a direct object, as shown in (7) (see Kayne & Pollock 1978: 604; Kayne & Pollock 1998; Valois & Dupuis 1992; Collins & Branigan 1997; Español–Echevarría, Pinto & de Wind 1998, among others):

(7) *Je me demande quand acheteront les consommateurs les
I wonder when will.buy the consumers.NOM the
pommes
apples.ACC

If, however, the direct object itself is either wh–extracted or cliticized SI becomes possible again:

(8) a. Que crois–tu que manquent un grand nombre
what believe–you that be.absent.from a great number
d'etudiants?
of students

b. Tes cours, a quelle occasion les ont manqués un
your courses at which occasion them have been.absent.from a
grand nombre d'etudiants?
great number of students

The object must either be moved out of the vP, as in (8), or surface as a PP, as in (9) (Collins & Branigan 1997 citing Kayne 1972):

(9) ?Quand écrira ton frère à sa petite amie?
when will.write your brother to his little friend
'When will your brother write to his girlfriend?'

We thus conclude that in order for a subject to remain vP-internal in French, either the DP object must be moved outside the vP or the object must be realized as PP.

2.1.2 Quotative Inversion

Sentences with direct speech complements allow a kind of inversion called *Quotative Inversion (QI)* by Collins & Branigan (1997) and Collins (1997):

(10) "I am happy", said Mary

Collins & Branigan point out that QI is in many respects similar to SI. As with SI, the subject remains in vP-internal position, and the evidence again comes from floated quantifiers. The ungrammaticality of floating quantifiers following the inverted subject in examples like (11c) provides evidence that the subject has not moved outside the vP (assuming again Sportiche 1988):

(11) a. "We must do this again", the guests all declared to Tony
b. "We must do this again", declared all the guests to Tony
c. *"We must do this again", declared the guests *all* to Tony

Similarly to SI, when the quote triggers inversion in sentences with transitive verbs containing an indirect object, the result is ungrammatical (12a) unless the indirect object is expressed as a PP (12b). No such conflict arises in the absence of inversion, as shown in (12c):

(12) a. *"What is the exchange rate?", asked Mary John
b. "What is the exchange rate?", asked Mary of John
c. "What is the exchange rate?", Mary asked John

Thus, QI displays similar characteristics as SI: The transitive subject can remain vP-internally only if vP does not contain another nominal argument. A vP-internal subject is compatible with a vP-internal, co–argumental PP, though (see (12b)).

2.2 The Universality of the SSG

The above facts motivate the generalization in (13):

(13) Subject–inversion with vP-internal subjects is prohibited in the presence of vP-internal DP objects.

(13) follows from the SSG, repeated below:

(1) By Spell–Out vP can contain only one argument with a structural Case feature

In A&A (2001) we claim that the SSG applies universally. There is no language in which both the subject and the object with a structural Case feature can remain vP-internal[2]. We substantiate this claim through a discussion of a number of word order patterns across languages. More specifically, we argue on the basis of *Transitive Expletive Constructions (TECs)* in Icelandic, VSO orders in Celtic and Arabic, and VOS orders in Italian and Catalan, that either the subject or the object or both are cross–linguistically parsed into a vP–external position. Sequences in which both arguments can be shown to remain vP-internal seem to be absent. We refer the reader to A&A (2001) for a detailed discussion of these patterns.

2.3 The SSG in the model of Chomsky (1995): v–to–T raising and Case checking

In A&A (2001) we offer an analysis of the SSG which is formulated in a framework that assumes that (i) overt and covert operations are empirically distinguishable, (ii) Case checking is the result of movement to a Case checking position (Spec,TP or Spec,vP) and (iii) Case checking takes place overtly or covertly (Chomsky 1995). More specifically, we point out that the generalization captured by the SSG can be further decomposed into two parts:

[2]See section 3 below for discussion of some apparent counterexamples to the SSG which support A&A's (2001) analysis.

(14) i. If two DP arguments are merged in the vP domain, at least one of them must externalise.
ii. If two arguments remain vP-internal, one of them must surface as a PP.

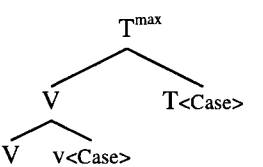
We suggest that the two clauses of (14) can be understood if the SSG derives from the Case constraint in (1'). According to (1'), the presence in the overt syntax of two arguments with an unchecked structural Case feature in the vP domain is prohibited.

(1') By Spell–Out, vP can contain only one argument with an unchecked Case feature.

The two clauses in (14) describe two alternative strategies that can be employed to circumvent (1'): One of the two arguments must leave the vP moving to (or through) its Case checking position (T or v, and from there it can move further to C; *clause i* of 14), or, alternatively, one of the two arguments is a PP lacking a structural Case feature (*clause ii* of 14). In both situations there is only one argument with an unchecked Case feature in the vP domain, conforming with (1').

The next question we address in A&A (2001) is what explains (1'). Our answer is to suggest that there is a link between v–to–T raising and the SSG. In configurations violating the SSG, v and T fall together either overtly (in French / Icelandic and for Fox & Pesetsky 2004, Johnson 1991 also in English) or covertly (in a traditional Emonds 1976 / Pollock 1989–style analysis of English). The Case–features of the arguments must be checked after v–to–T raising takes place creating a complex head with two Case features as in (15):

(15)



We propose that the complex head in (15), with two active Case features, is an illicit item[3]. In this analysis, the SSG (1') results from the improper

[3]There are several reasons why this might be so which are discussed in detail in A&A (2001). They all crucially rely on the assumption that T and v cannot directly enter into Case checking after head adjunction because they fail to c–command outside the non–terminal node

amalgamation of two Case–bearing heads v and T, as stated in (16):

(16) *v and T cannot both bear active Case features when they form a complex head.*

As a consequence of (16), it is necessary that at least one Case feature be checked before the complex head is formed:

(17) T<Case> or v<Case> must be eliminated before the complex head is formed.

Intuitively, a local relationship between an argument and its Case–checking head must be established, which is destroyed by the formation of a complex head with two active Case features.

The clearest example of the effects of (15)/(16) is instantiated by the transitivity restriction in English/French. In these cases, the numeration contains a v and a T which both bear *weak* Case features that can be eliminated without phrasal pied–piping. The derivation proceeds as follows:

(18) (i) First, v is merged, and the object does not raise overtly.
(ii) Then, T is merged.
(iii) The expletive is merged eliminating the EPP feature of T.
(iv) v raises to T overtly or covertly, resulting in the formation of a complex head T^{max} with two unchecked Case features.

As extensively argued for in A&A (2001, pp. 219–224), (15)/(16) does not arise in Icelandic TECs, Celtic and Arabic VSO, and Romance VOS orders, i.e. whenever one of the two arguments (or both) undergo movement to a vP external position.

The analysis of the SSG in terms of (15)/(16) predicts that subjects and objects may remain vP-internal in languages lacking v–to–T raising, in apparent violation of the SSG. As will be seen in the next section, this prediction is indeed borne out.

3 Challenges for SSG: Apparent exceptions

A number of exceptions to the SSG have been noted in the literature. In this section, we will review these cases showing how they can be accommodated under A&A's (2001) analysis. Crucially, whenever subjects and objects remain vP-internal in apparent violation of the SSG, either (i) or (ii) holds: (i) The

dominating them.

subject (or the object) is either Case–less or has its Case checked, and hence (1') is not violated. (ii) Raising of v–to–T fails to take place, and hence (16) is not violated.

3.1 A&A (2001): Greek, Spanish, Romanian

To begin with, A&A (2001) discuss VSO orders in Greek, Spanish, and Romanian, which appear to present counterexamples to the SSG. In the Greek example (19) both arguments remain vP-internal, as is evidenced by the fact that they follow manner adverbs, which have been argued to mark the left edge of the vP. Note that participles precede manner adverbs in Greek, a fact that has been analysed in terms of the proposal that they raise overtly to an aspectual head Asp above Voice where manner adverbs occur (see Alexiadou 1997):

(19) an ehi idi diavasi [$_{vP}$ prosektika [o Janis to vivlio]]
if has already read carefully the John.NOM the book.ACC
'If John has already read the book carefully'

A&A (2001) argue that such orders do not challenge the SSG, understood in terms of the Case checking constraint (1'), because the Case of the in situ subject is realized on the pronominal verbal agreement which has the status of a clitic and checks overtly its (phi and Case) features on T as a result of verb–raising (cf. A&A 1998). In this analysis, the inverted in situ subject does not have an unchecked structural Case feature, despite appearances to the contrary. We link the above mentioned property of Greek, Spanish, and Romanian verbal subject agreement to the clitic doubling parameter which permits the formation of such feature–chains[4] between clitics and in situ DP arguments in clitic doubling languages like Greek, Spanish, and Romanian and prohibits them in non–clitic doubling languages like French, Italian, and Catalan.

3.2 Further exceptions to the SSG

Some further counterexamples to the SSG are reported in the more recent work of Wurmbrand (2004), Carstens (2005) and Baker & Collins (2006). In

[4]See Anagnostopoulou (2003) for an analysis of such feature–chains. Based on locality considerations, Anagnostopoulou argues that clitic doubling languages permit overt feature movement without phrasal pied piping.

what follows, we argue that these cases do not contradict the SSG, understood as either (1') or (16).

3.2.1 German

It has been argued that subjects and objects may both remain vP-internal in German (see e.g. Haider 1993, 2005; Fanselow 2001; Wurmbrand 2004 and others). Evidence for this comes from two sources. First, adverbial placement demonstrates that both arguments remain inside the vP:

(20) weil schon oft ein junger Hund einen Briefträger gebissen hat
since already often a young dog a mailman bitten has
'Since a young dog has already often bitten a mailman'

Second, in contexts of vP–fronting, both arguments can be topicalized[5]:

(21) [Ein junger Hund einen Briefträger gebissen] hat hier schon
a young dog a mailman bitten has here already
oft
often
'It has happened often here already that a young dog has bitten a mailman'

Wurmbrand (2004) notes that these examples present a potential problem for the SSG. We believe, however, that the problem is only apparent. There are two potential explanations for why German permits vP-internal subjects and objects, both of which are compatible with A&A's (2001) analysis: (a) One possibility is that German permits feature–chains between null clitics and *in situ* DP arguments qualifying essentially as a clitic doubling language (following Haider 1985; Fanselow 2001). Hence, there is no violation of the SSG understood as in (1'). (b) Alternatively, German lacks head–movement being a head–final language (Haider 1993, 2005). Hence, German lacks the formation of complex heads like (15) that would lead to a violation of (16).

3.2.2 Kilega

As discussed in Kinyalolo (1991) and Carstens (2005, pp. 238–239), Kilega has transitive inversion constructions in which subjects and objects arguably

[5]Examples (20) and (21) are taken from Wurmbrand (2004). Note that not all speakers judge (21) to be well–formed (Gereon Müller, p.c.).

remain *in situ* in violation of the SSG. Consider (22) (Carstens 2005, ex. (34)), which displays locative inversion in a sentence containing an auxiliary and a main verb:

(22) Ku–Lúgushwá kú–kili ku–á–twag–a
17–Lúgushwá 17.SA–be.still 17.SA–A–stampede–FV
nzogu maswá
10.elephant 6.farm
'At Lugushwa are elephants still stampeding over (the) farms'

Carstens argues that (22) should be analysed as in (23) (raising of the locative omitted):

(23) $[_{TP}$ ku–T $[_{AspP1}$ kili_{Asp1} $[_{AspP2}$ ku–á_{Asp2} $[_{MP}$ twag–a_{Mood}
17.agr–(PRES) be.still 17.agr–A stampeding
$[_{vP}$ SU t_v $[_{VP}$ t_V OB ku Lúgushwá]]]]]
elephants farm LOC Lúgushwá

On the assumption that there is no V–raising beyond the suffixes in Kilega, V–movement always terminates to a position below T in this language (as extensively argued for in Kinyalolo 2003 and Carstens 2005). In compound tenses like (22)/(23) (see Carstens 2005, pp. 226–230) the main verb raises to a low MoodP, which contains the final vowel of Bantu verbs, and is preceded by an aspectual prefix projected under a low aspectual head Asp2. The aspectual auxiliary is in Asp1 while T hosts abstract Tense features (PRES in this example). The fact that both the subject and the object follow the main verb in (22) provides evidence that both arguments remain vP-internal[6].

As pointed out by Carstens (2005, p. 239), lack of V–to–T movement in Kilega provides the key to an understanding of why inversion constructions do not violate the SSG. Under the assumption that the SSG results from the improper amalgamation of v + T with active Case features, as proposed by A&A (2001), the fact that V–to–T movement does not take place in Kilega can explain why this language has transitive inversion constructions with vP-internal arguments. The head (15) is never formed, and the constraint (16) does not arise.

[6]Note that the verb agrees with the moved locative phrase rather than the *in situ* subject in (22), as is always the case in Bantu constructions displaying inversion (see Baker 2003; Carstens 2005 for recent discussion).

3.2.3 Kinande

Baker & Collins (2006, fn. 13) point out that Kinande constructions as in (24) could be taken to violate the SSG:

(24) Mo–ha–teta–sat–a (*a–)mu–kali omo–soko
AFF–there–NEG/PAST–dance–FV AUG–1–woman LOC.18–market
'There danced no woman in the market'

Example (24) features a vP-internal subject and a locative argument which is nominal and bears structural Case, as extensively discussed in Baker & Collins (2006). It therefore looks as if (24) violates the SSG. This is only apparent, however, because the vP-internal subject in (24) lacks a structural Case feature, as Baker & Collins (2006) argue. The crucial evidence for this comes from the observation that the subject is not allowed to bear the initial augment vowel in (24). The augment vowel occurs with all nominals that are interpreted as DPs in Kinande while in its absence, nominals are interpreted as NPs. Under the assumption that structural Case is a feature of D, the obligatory absence of the augment in (24) signifies that the subject lacks Case. In turn, this explains why both nominal arguments remain vP-internally. Thus, Kinande presents one more case where the SSG, understood as the Case constraint (1'), is not violated[7].

4 Expanding the SSG: object movement in linker constructions

Collins (2003) and Baker & Collins (2006) discuss a constraint which forces movement of one object out of the VP–domain in constructions involving two objects in the Khoisan languages Ju'hoansi and ≠Hoan (Collins 2003; Baker & Collins 2006), and in Kinande (Niger–Congo; Baker & Collins 2006). The existence of this constraint, which is strongly reminiscent of the SSG, leads to the view that the SSG is more general than previously recognized. The constraint in question is attested in constructions where a particle, called "linker" by Collins (2003) and Baker & Collins (2006), appears between the direct object and a secondary object or nominal adpositional phrase. Some examples illustrating the linker–construction in Ju'hoansi are provided in (25). In (25a) the linker *ko* appears between the theme and a locative phrase, in (25b) between the theme and an instrument and in (25c) it occurs between

[7] According to Baker & Collins (2006, fn. 5), the verb probably moves to Infl in Kinande, since subject–adverb–verb orders are not found in this language. Hence, the grammaticality of Kinande (24) cannot be attributed to the lack of v–to–T raising, unlike the Kilega example (22).

the two objects of a double object construction, the beneficiary and the theme (from Collins 2003, pp. 1–2):

(25) a. Uto dchuun–a |Kaece ko n!ana n!ang
car hit–TRANS |Kaece KO road in
'A car hit |Kaece in the road'

b. Mi ba ||ohm–a !aihn ko |'ai
My father chop–TRANS tree KO axe
'My father chopped the tree with an axe'

c. Besa komm ||'ama–|'an Oba ko tcisi
Besa EMPH buy–give Oba KO things
'Besa bought Oba some things'

In what follows, we will mainly concentrate on Ju'hoansi locative constructions, as discussed in Collins (2003),[8] because they present particularly clear evidence that the SSG not only regulates the placement of subjects and objects in the vP domain but also regulates the placement of two objects in the VP domain.

To begin with, observe that in Ju'hoansi locative (and instrumental) constructions the particle *ko* co-occurs with the transitivity suffix *–a* (glossed TRANS in (25a) and (25b)). As will be seen immediately, the conditions under which *–a* and *ko* surface are closely related, though not identical. While the transitivity suffix *–a* appears when a locative phrase follows a transitive and an intransitive verb, the particle *ko* appears when a locative phrase follows a transitive verb but not when a locative phrase follows an intransitive verb. More specifically, the transitivity suffix *–a* and the particle *ko* are both disallowed with transitive verbs, as shown in (26), while they are both required when a locative phrase is added to transitive verbs, as shown in (25a) above.

(26) a. Uto dchuun–(*a) |Kaece
Car hit–TRANS |Kaece
'The car hit |Kaece'

b. *Uto dchuun–(a) |Kaece ko
Car hit–TRANS |Kaece KO

[8]Baker & Collins (2006) re-interpret some of the facts discussed in Collins (2003; most notably, the A' extraction facts in Ju'hoansi illustrated in (34) and (35) below) in a way that obscures the similarity between the phenomena studied in A&A (2001) and their linker counterparts. For this reason, we mainly focus on Collins's (2003) interpretation of the facts. Even though a full comparison between A&A (2001), Collins (2003) and Baker & Collins (2006) is beyond the scope of the present paper, we highlight in fn. 10 below one crucial difference between Baker & Collins's (2006) approach and ours.

'The car hit |Kaece'

c. *Uto dchuun–(a) ko |Kaece
Car hit–TRANS KO |Kaece
'The car hit |Kaece'

On the other hand, when a locative phrase is added to an intransitive verb, *–a* is required (see (27)) but *ko* is ungrammatical (compare (28a) to (28b)):

(27) a. Ha ku u
3.SG ASP go
'He was going'

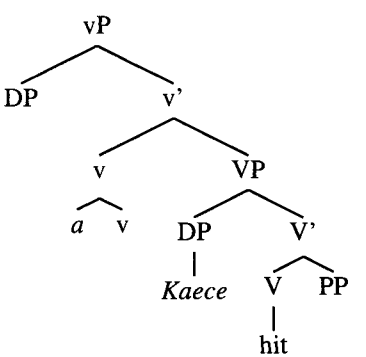
b. *Ha* ku u–a Tjum!kui
3.SG ASP go–TRANS Tjum!kui
'He was going to Tjum!kui'

(28) a. Lena koh djxani–a tju n!ang
Lena PAST dance–TRANS house in
'Lena danced in the house'

b. *Lena koh djxani–a ko tju n!ang
Lena PAST dance–TRANS KO house in

In order to account for the distribution of *–a* in Ju'hoansi, Collins (2003) argues that locative phrases are nominal and have a Case feature to check. The transitivity suffix *–a* is inserted to check the Case of locative phrases. This explains why *–a* is added in transitive (25a) and intransitive (27b), (28a). In (25a) transitive v checks the Case of either DP |*Kaece* or PP *n!ana n!ang*, and the transitivity suffix *–a* checks the Case of the other argument, as shown in (29):

(29)



Similarly, in (27b) and (28a) *–a* is added on an intransitive v (which lacks a Case feature) to check the Case feature of the nominal PPs *Tjum!kui* and *tju n!ang*, respectively.

The question that arises next concerns the distribution of *ko*, which is present in transitive constructions like (25a) and absent in intransitives like (27), (28). Note that the presence of *ko* is obligatory in transitives. Example (30), which lacks *ko*, is ungrammatical:

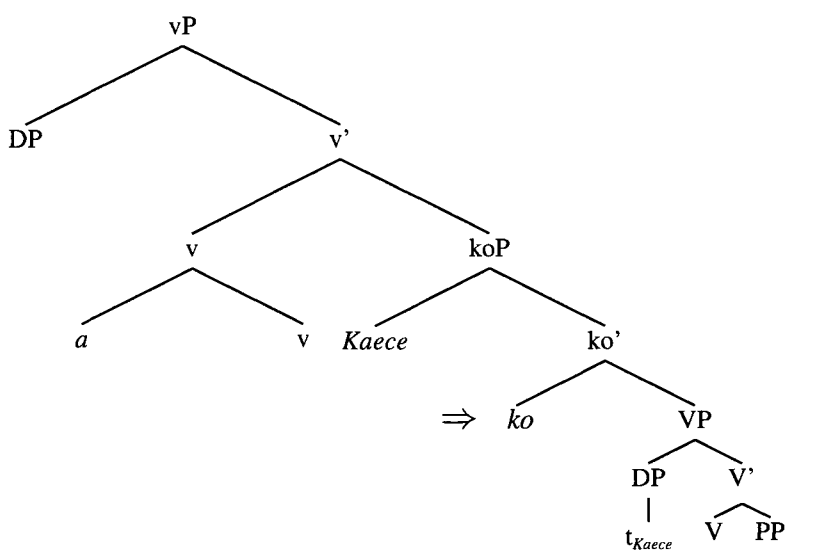
(30) *Uto dchuun–a |Kaece n!ana n!ang
car hit–TRANS |Kaece road in
'A car hit |Kaece in the road'

In order to account for the obligatory presence of *ko* in transitive constructions with a locative (such as (30)) and its obligatory absence in intransitive constructions with a locative (such as (28b)), Collins (2003, pp. 15–16) argues that *ko* is a Last Resort mechanism. It is inserted to provide a landing site for movement in constructions that would otherwise violate a condition which he labels *Multiple Case Condition (MCC)*:

(31) *Multiple Case Condition*
By Spell–Out, VP can contain no more than one argument with a (valued) undeleted Case feature.

In (29) above, the complex functional head [$_v$ *a* v] has two sets of uninterpretable phi-features, one for *a* and one for v. Even though two Agree relations can be established—Agree (v, DP) and Agree (*a*, PP)—there are two Case features internal to the VP that need to be deleted at Spell–Out. In order to avoid a violation of the MCC, *ko* is merged providing a landing site for one of the two arguments, as schematized in (32):

(32)



Being a Last Resort operation, *ko*–insertion is triggered only if a violation of the MCC would ensue, which explains why *ko* is obligatorily absent in intransitives.

In (25a), the DP *Kaece* leaves the VP, as depicted in (32). Movement of locative and instrumental PPs is also possible in Ju'hoansi, as shown by the examples (33) (see Collins 2003: p. 9):

(33) a. Uto dchuun–a n!ana n!ang ko |Kaece
car hit–TRANS road in KO |Kaece
'A car hit Kaece in the road'

b. Ha gu–a ||'aisi ko tju
3SG build–TRANS grass KO house
'He built the house with grass'

For the purposes of the MCC, it doesn't matter which argument moves out of the VP as long as one of them does[9].

The MCC is a version or a close relative of the SSG. Just as the SSG forces movement of either the subject or the object out of the vP when both have structural Case, the MCC forces movement of either the direct object or the adpositional phrase out of the VP when both have structural Case. And, importantly, if one of the VP constituents is extracted by A'–movement, the result is acceptable without *ko*, as shown in (34) and (35):

(34) |Kaece komm uto dchuun–a (*ko) n!ama n!ang
|Kaece EMPH car hit–TRANS KO road in
'|Kaece, the car hit in the road'

(35) N!ama n!ang komm uto dchuun–a (*ko) |Kaece
road in EMPH car hit–TRANS KO |Kaece
'In the road the car hit |Kaece'

This pattern is strongly reminiscent of the conditions licensing *Stylistic Inversion* in French. (Recall that SI is possible when the object undergoes A' movement or cliticization as shown in (8), section 2.1.1 above.)

On the basis of the above discussion, we conclude that the SSG is a broader condition than initially assumed. Descriptively, we distinguish between two types of SSG effects:

[9] See Collins (2003) and Baker & Collins (2006) for further discussion of the constraints of movement of the lower argument in Ju'hoansi double object constructions and ≠Hoan locative, instrumental and double object constructions, which follow from locality considerations.

(i) *The vP–type* is attested in the languages discussed in A&A (2001). Here the competing arguments are the subject and the object, which are not allowed to remain both vP-internal, as illustrated by the Icelandic example (36):

(36) *það klaruðu [$_{vP}$ alveg margar mys ostinn]
there finished completely many mice the.cheese

(ii) *The VP–Type* is attested in the languages discussed in Collins (2003) and Baker & Collins (2006)[10]. Here the competing arguments are the theme-object and the PP-nominal object, which are not allowed to remain both VP–internal, as shown in (37).

(37) *Uto dchuun–a ko |Kaece n!ana n!ang
car hit–TRANS KO |Kaece road in
'A car hit |Kaece in the road'

Note, finally, that the configurations discussed by A&A (2001) and Collins (2003) are formally very similar. In both cases we arrive at the formation of a complex head, which is argued to be an illicit object in A&A (2001). The complex head (15), repeated below as (38), is created in vP-type SSG effects. The complex head (29), repeated below as (39), is created in VP–type SSG effects.

[10]Baker & Collins (2006) redefine the MCC in a way that is less compatible with the SSG:

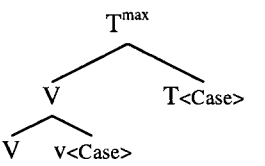
(i) *The Multiple Case Constraint (revised)* (their condition (35)).
By Spell–Out, if the VP complement of v contains a DP with a Case feature, then VP cannot contain any other nominal expression (includes augmentless NPs, semi–nominal locatives, and DPs with Case features)

The basic intuition behind the formulation in (i) is that if v checks the Case of a DP (with an unvalued Case feature) inside its VP complement, then it requires that no other potential Case bearing elements (such as NPs or locatives) are around internal to that VP to compete with it. The reason why Baker & Collins (2006) define the MCC in terms of potential (and not actual) Case bearing elements is because they argue that the MCC applies even when the VP contains locative adjuncts and augmentless NPs in Kinande which do not have a Case feature. This version of the MCC does not seem to us to be unifiable with the SSG for two reasons at least:

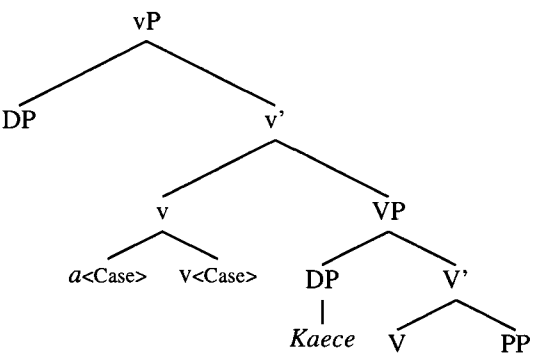
(i) A&A's (2001) account of the SSG crucially relies on the assumption that both vP-internal arguments require Case, and one of them cannot receive it after the formation of the complex head (15), which leads to a crashing derivation.

(ii) As we saw in section 3.2.3 above, constructions with a vP-internal subject and a locative nominal argument such as (24) are well–formed in Kinande, in apparent violation of the SSG, because, crucially, the vP-internal subject in (24) is augmentless and hence lacks a structural Case feature. Augmentless NPs in Kinande, however, still trigger the insertion of a linker to avoid a violation of the MCC, as stated in the revised MCC (i)/(35) above.

(38)



(39)



The two configurations differ only in that the complex head in (38) is formed by v–raising, while the formation of a complex head is derived by suffixation of the transitive marker *a* to v in (39).

5 The SSG revisited

5.1 Generalized EPP and Case

In the preceding sections we saw that a constraint prohibiting more than one arguments with an unchecked structural Case feature applies to distinct domains (vP or VP) forcing argument externalisation. Specifically:

(i) When the constraint under discussion targets the vP it derives the effects of the SSG discussed in A&A (2001):

(1') *The Subject–In–Situ Generalization (SSG)*
By Spell–Out, vP can contain only one argument with an unchecked Case feature.

The SSG applies to subjects and objects and triggers externalisation in the domain above v (spec,TP or spec,vP).

(ii) When the constraint targets the VP it derives the effects of the MCC (a version of the SSG) discussed in Collins (2003):

(31) *Multiple Case Condition (MCC)*
By Spell–Out, VP can contain no more than one argument with a (valued) undeleted Case feature.

The MCC applies to two objects and triggers externalisation in the domain above *ko*.

In view of the tight connection between (1')/(31) and argument externalisation, the SSG / MCC can, in turn, be seen as a constraint on multiple Case checking implementing EPP effects.

In recent literature, there are two different conceptions of the EPP. On one view, EPP is a condition on T or a feature on T. On another view, EPP is a general condition regulating argument movement to v or T (*Generalized EPP*; see A&A 1996, 1997 for an early proposal along these lines). As discussed in A&A (2001), both versions of the EPP are empirically correct, capturing different sets of facts. In the SSG configurations identified and analysed by A&A (2001), the EPP as a property of T is satisfied in four different ways summarized in (40):

(40) *Devices satisfying the EPP property of T*
(i) an expletive in (T)ECs (Icelandic, English).
(ii) a locative PP in Locative Inversion (Branigan 1993, Collins 1997).
(iii) V–raising in contexts with VSO orders (Celtic, Greek; see A&A 1998 for arguments that in these languages, V–to–I checks the EPP feature).
(iv) a (wh)–phrase in environments of Quotative and Stylistic Inversion.

In exactly these contexts, the SSG applies forcing movement of either the object or the subject to a (layered) specifier of v and/or T, respectively. One of the two DP arguments must vacate the vP in order for the derivation to converge, and the derivation converges only if an additional landing site can be provided to host one of the DPs. Crucially, creation of this additional site is independent of the EPP on T. Configurations which escape the effects of the SSG combine three independent parameters that have been proposed in the literature:

(i) the Spec,TP parameter, Bobaljik & Jonas (1996)

(ii) the doubling parameter, A&A (1998)
(iii) the object shift parameter, (see Holmberg & Platzack (1995) among others)

The derivations in the languages discussed in A&A (2001) converge as follows:

(41) *Devices satisfying the SSG*
  (i) *Icelandic TECs:* presence of (an additional) Spec,TP, or an outer Spec,vP.
  (ii) *Celtic VSO:* presence of an additional vP–external specifier (spec, TP) hosting the subject.
  (iii) *Greek:* raising of X° suffices to satisfy the EPP and Case requirements of the subjects (due to the doubling parameter).
  (iv) *Italian, Catalan:* presence of an outer Spec,vP to host the object in VOS.

Interestingly, EPP–related object movement may also lead to improvement in English and in French. Cases in point are cliticization in French (following Kayne 1991; A&A 1997), and marginal TECs in English. ((43) can be taken to involve object movement and v–raising to T; on the latter see Fox & Pesetsky 2004)[11]:

(42) Tes cours, a quelle occasion **les** ont manqué un
your courses at which occasion them have been.absent.from a
grand nombre d'etudiants?
great number of students

(43) ?There entered the room a strange man.

The SSG can thus be formulated as in (44):

(44) At least one argument must vacate the vP.

[11]The limited availability of transitive TECs may be due to the special nature of the expletive construction in English, which requires verbs of existence and appearance (Levin & Rappaport Hovav 1995). Note also that (43) becomes worse if the participial construction is used and is ungrammatical in negated contexts, a fact strongly reminiscent of conditions in QI (judgments courtesy of Jonny Butler and Thomas McFadden).

The requirement in (44) is, in turn, linked to the generalized EPP:

(45) *Generalized EPP and the SSG*
EPP features provide landing sites for arguments escaping the SSG[12].

The conditions under which such EPP features are assigned on T and/or v are subject to parametric variation (see e.g. English vs. Icelandic vs. Italian).

Note, finally, that (45) arises only when the vP domain contains DPs that need to have their Case features valued. Being linked to Case the generalized EPP cannot—unlike the EPP property of T—be satisfied by expletives and PPs. The VP data discussed in Collins suggest that a notion of generalized EPP linked to Case is on the right track. The main function of *ko* consists in introducing an EPP feature into the derivation. *ko* provides a landing site for one of the nominal objects internal to VP. The structure is only projected to allow the DP object / nominal PP to raise out of the VP, escaping the verdict of the SSG / MCC.

5.2 The SSG in a system with cyclic Agree

In this final section, we would like to discuss the status of the SSG in a cyclic Agree system, which creates two problems:

The first problem is that A&A's (2001) account of the SSG (summarized in section 2.3 above) is incompatible with cyclic Agree where the derivation would proceed as follows:

(46) (i) V is Merged with OBJ yielding the low VP shell
(ii) v is merged with VP yielding v'.
(iii) SUBJ is merged with v' yielding vP.
(iv) Agree is established between v and OBJ, and the structural Case of OBJ is assigned the value ACC, while the phi-features of v are valued by OBJ[13].
(v) T is merged with vP
(vi) Agree is established between T and SUBJ, and the structural Case of SUBJ is assigned the value NOM, while the phi-features of T are valued by SUBJ.
(vii) v raises to T

[12](45) is related to the notion of the EPP used in Miyagawa (2005), Chomsky (2005).

[13]It could be that step (iv) precedes step (iii) or, as argued for by Müller 2004, that the relative order between these two steps is a parameter.

In the derivation (46), both arguments enter Case checking relations in a strictly cyclic fashion, as outlined in steps (iv) and (vi). Hence, there is no problem with the Case checking relations established—the vP domain never contains two unvalued Case features since v and OBJ enter Agree before T is merged. As a consequence, there is no reason why both the subject and the object cannot remain *in-situ*. The derivation appears to converge, incorrectly. We conclude that the SSG cannot be (directly) expressed in a system based on cyclic Agree[14].

The second problem is that the A&A (2001) cases now differ qualitatively from the Collins (2003) cases in a crucial way. In Collins's cases, the complete head [v +a] is inserted as one unit (see (29) above), with two unvalued sets of phi/Case features. By contrast, in A&A's cases, the complex head is formed syntactically by v–raising to T (step (vii) in (46)), long after valuation of v's phi-features by OBJ and OBJ's Case feature by v (step (iv) in (46)).

In order to (i) re–state the account of A&A in an Agree based system and (ii) maintain the similarity between the configurations discussed by Collins and by A&A, there are in principle two options. On the first option, it is possible to adopt counter-cyclic Agree, as formulated by the 'T–v–Agree Hypothesis' below:

(47) *T–v–Agree Hypothesis*[15]:
v enters Agree with T and *then* Case valuation takes place, creating a configuration of Case checking ambiguity (v and T could value the Case of SUBJ or OBJ).

Under this hypothesis, the Agree relation between the v–T heads emulates the effects of a complex head in the older system.

The question that arises now is the following: Why should Agree between T and v precede Case valuation? In order to answer this question we would like to tentatively suggest that an Agree relation between T and v can be motivated

[14]Certain complications concerning cyclicity also arise in the system adopted in A&A (2001; see footnotes 34, 36). In particular, it must be assumed that XP movement precedes head movement in the overt syntax, while head movement precedes XP movement at LF. The problem is caused by the assumption that English lacks overt verb raising (the traditional Emonds 1976 / Pollock 1989–style analysis) and does not seem to arise if verb movement in English is always overt (as argued for in Fox & Pesetsky 2004).

[15]Chris Collins (personal communication), suggests that the relation between v and T might be more like Match than Agree since it is not entirely clear which is the goal and which the probe, i.e. whether v is valuing T or the other way round. On the Agree version adopted in this paper, T must be the probe and v the goal, i.e. v must be taken to value an unvalued feature on T. This is more straightforward under option (ii) discussed right below.

in at least two ways: (i) Adopting Distributed Morphology, it can be suggested that the vP is categorially underspecified: it can project a clause or be part of a nominalization. In a clause, T provides the categorial specification of vP[16]. 'Verbal' amounts to 'combining with Tense' (either via head–movement or Agree). (ii) Alternatively, it can be proposed that T and v participate in an Agreement relationship which leads to feature valuation for the purposes of the semantic interpretation of Tense (along the lines suggested in Pesetsky & Torrego 2004). In the new analysis, the relevant steps in the derivation must always be ordered as follows:

(48) (i) First, v Agrees with T, resulting in a Case checking ambiguity / indeterminacy (what will Agree with what first?).
(ii) EPP provides a guideline for ordering the Agree relations:
– If T bears an EPP feature, the first Agree relation involves T and SUBJ.
– If v bears an EPP feature, the first Agree relation involves v and OBJ.
(iii) If both v and T bear EPP features, Agree proceeds strictly cyclically.

On this view, EPP features cancel the ambiguity / indeterminacy configuration[17]. If our proposal is correct, then SSG effects provide an empirical argument that Agree is—unlike Move—locally counter-cyclic.

On the second option, it can be assumed that while Agree is established between v and the OBJ, no valuation takes place. In other words, the relationship is established but nothing happens and Case valuation is delayed till the creation of the complex head. Again here we would need to determine what causes delay of feature valuation, a concept that remains unclear.

6 Conclusion

In this paper we attempted to investigate a question raised in section 1: What is the status of the SSG in the grammar? Based on evidence from Indo–European and Khoisan languages we arrived at the following answers:

(i) The SSG is a constraint on multiple case checking which applies in the vP and the VP domain.

[16]For implementations of this idea, see Alexiadou (2001), Bhatt & Embick (in prog.).

[17]Thanks to Winfried Lechner for discussing this issue and suggesting this option to us.

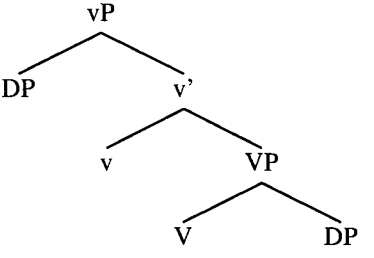
(ii) The EPP functions as a guideline so that a given domain X can escape the SSG.

A further question ensues once the SSG is embedded in a cyclic Agree system: Does the SSG follow from an independent restriction, such as the ban on complex head formation with two unvalued Case features proposed by A&A (2001)? If the answer to this question is positive, then SSG effects provide evidence that the Agree relations in the TP domain are locally counter-cyclic, for the reasons discussed in section 5.2.

Appendix: Richards (2004)

In the recent literature there is a proposal that tries to capture some of the facts discussed in A&A (2001) within a phase based system, namely Richards (2004). According to Richards (2004), the effects of the SSG can be explained as linearization failures inside the vP–phase. Linearization fails whenever the objects to be linearized in a strong phase are insufficiently distinct.

(49) *



Richards proposes that linearization makes reference only to node labels. Moreover, all and only those nodes within a phase must be linearized (Kayne 1994). It follows that syntactic nodes with the same label must not be located too close together in the tree—they need to be separated by a phase boundary, otherwise they cannot be ordered w.r.t. each other. (49) violates this ordering restriction.[18] Richards observes that this is subcase of a more general ban on multiple adjacent objects similar to the Obligatory Contour Principle (OCP) in phonology (see also *Unlike Category Condition* of Hoekstra 1984: chapter 2). Richards's account captures the data discussed in A&A (2001) as well as cases

[18]It is not clear what the status of v–raising is in Richards's accounts. Note that in his discussion of SI and QI the verb vacates the vP.

as in (50), initially discussed in Ross (1972):

(50) *The police are stopping drinking on the campus *double–ing*

Richards's proposal, however, faces a number of problems:

(i) He cannot account for systematic exceptions to the SSG such as the VSO in Greek/Spanish/Romanian. These languages allow VSO orders with two vP-internal DP arguments as discussed in A&A (2001).

(ii) Richards cannot explain the data from German and Kilega discussed in section 3 above: in these languages both DP arguments can remain vP-internal, for principled reasons, as we saw.[19]

(iii) Khoisan languages pose a further challenge for Richards' analysis, for the reasons that (a) the arguments subject to the SSG bear distinct category labels (DP, PP), and (b) the domain in which the condition is computed (VP) is not a phase.

(iv) Finally, Richards would predict that multiple series of PP or NP adjuncts are impossible, contrary to fact.[20]

[19]He can account for Kinande if the subject is an NP and the object a DP (see section 3).

[20]There is a further exception involving DP–internal arguments: multiple "of" phrases in nominalizations seem to fall under the same pattern (Alexiadou 2001, Richards 2004):

(i) a. the enemy's destruction of the city
b. the destruction of the city by the enemy
c. *the destruction of the city of the enemy

The contrast between (c) and (b) is crucial. Richards would have to say that one is Case and the other one is P. For A&A, these examples have a different treatment. Assuming that DPs are not transitive (Alexiadou 2001), the situation we face here is somehow different. What we have is one head that must value two structural Cases, hence no matter what happens one Case feature remains unvalued.

References

Alexiadou, Artemis

1997 *Adverb Placement: a Case Study in Antisymmetric Syntax.* John Benjamins, Amsterdam.

2001 *Functional structure in nominals: nominalization, and ergativity.* John Benjamins, Amsterdam.

Alexiadou, Artemis and Elena Anagnostopoulou

1996 *Symmetries, asymmetries and the role of agreement.* Paper presented at the *19th* GLOW *Colloquium*, 17.4.96, Athens, Greece.

1997 Toward a uniform account of scrambling and clitic doubling. In Werner Abraham and Elly van Gelderen, (eds.), *German: Syntactic Problems—Problematic Syntax*, pp. 142–161. Max Niemeyer Verlag, Tübingen.

1998 Parametrizing AGR: Word Order, Verb–movement and EPP checking. *Natural Language & Linguistic Theory*, 16.3: 491–539.

2001 The subject in situ generalization, and the role of Case in driving computations. *Linguistic Inquiry*, 32: 193–231.

Anagnostopoulou, Elena

2003 *The syntax of ditransitives.* Mouton de Gruyter, Berlin and New York.

Baker, Mark C.

2003 Agreement, Dislocation and Partial Configurationality. In Andrew Carnie, Heidi Harley, and MaryAnn Willie, (eds.), *Formal Approaches to Function in Grammar: In honor of Eloise Jelinek*, pp. 107–132. John Benjamins, Amsterdam.

Baker, Mark C. and Chris Collins

2006 Linkers and the Internal Structure of vP. *Natural Language & Linguistic Theory*, 24: 307–354.

Benmamoun, Elabbas

1998 Spec–Head Agreement and Overt Case in Arabic. In David Adger, Susan Pintzuk, Bernadette Plunkett, and George Tsoulas, (eds.), *Specifiers: Minimalist Approaches*, pp. 110–125. Oxford University Press, New York, NY.

Bhatt, Rajesh and David Embick

In progress *Causative Derivations in Hindi.* Manuscript, UMass and UPenn.

Bobaljik, Jonathan and Andrew Carnie

1992 *A Minimalist Approach to some problems of Irish Word Order.* Unpublished Manuscript, Massachusetts Institute of Technology, Cambridge, MA.

Bobaljik, Jonathan and Diane Jonas

1996 Subject Positions and the Role of TP. *Linguistic Inquiry*, 27: 195–236.

Branigan, Phil
1993 *Locative Inversion and the Extended Projection Principle*. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Carstens, Vicki
2005 Agree and EPP in Bantu. *Natural Language & Linguistic Theory*, 23: 219–279.

Chomsky, Noam
1995 *The Minimalist Program*. MIT Press, Cambridge, MA.
2000 Minimalist inquiries: The framework. In Roger Martin, David Michaels, and Juan Uriagereka, (eds.), *Step by Step. Essays on Minimalist Syntax in Honor of Howard Lasnik*, pp. 89–151. MIT Press, Cambridge, MA.
2001 Derivation by phase. In Michael Kenstowicz, (ed.), *Ken Hale. A Life in Language*, pp. 1–52. MIT Press, Cambridge, MA.
2005 *On phases*. Manuscript, Massachusetts Institute of Technology, Cambridge, MA.

Collins, Chris
1997 *Local Economy*. MIT Press, Cambridge, MA.
2003 The internal structure of vP in Ju'hoansi and Hoan. *Studia Linguistica*, 57: 1–25.

Collins, Chris and Phil Branigan
1997 Quotative Inversion. *Natural Language & Linguistic Theory*, 15: 1–41.

Collins, Chris and Höskuldur Thráinsson
1993 Object shift in double object constructions and the Theory of Case. In Colin Phillips, (ed.), *Papers on Case & Agreement II. MIT Working Papers in Linguistics*, 19: 131–174.
1996 VP–internal structure and Object shift in Icelandic. *Linguistic Inquiry*, 27: 391–444.

Demirdache, Hamida
1988 *Nominative NPs in Modern Standard Arabic*. Unpublished Manuscript, Massachusetts Institute of Technology, Cambridge, MA.

Déprez, Viviane
1991 Two Ways of Moving the Verb in French. In Lisa Lai Shen Cheng and Hamida Demirdache, (eds.), *Papers on Wh-Movement. MIT Working Papers in Linguistics*, 13: 47–85.

Dobrovie–Sorin, Carmen
1987 *Syntaxe du Roumain*. These de Doctorat d'Etat, Université de Paris 7.
1993 *The Syntax of Romanian: comparative studies in Romance*. Mouton de Gruyter, Berlin and New York.

Emonds, Joseph E.
1976 *A transformational approach to English syntax*. Academic Press, New York, NY.

Español–Echevarría, Manuel, Manuela Pinto and Maarten de Wind
1998 *A uniform analysis of case adjacency in Romance.* Paper presented at the Workshop on *Inversion in Romance*, 1998, Amsterdam.

Fanselow, Gisbert
1996 Minimal Link, phi–features, Case and theta–linking. GLOW *Newsletter*, 36: 30–31.
2001 Features, θ–Roles, and Free Constituent Order. *Linguistic Inquiry*, 32: 405–437.

Fassi–Fehri, Abdelkader
1993 *Issues in the structure of Arabic clauses and words.* Kluwer Academic Publishers, Dordrecht.

Fox, Danny and David Pesetsky
2004 *Cyclic Linearization and the typology of movement.* Manuscript, Massachusetts Institute of Technology, Cambridge, MA.

Haider, Hubert
1985 The Case of German. In Jindřich Toman, (ed.), *Studies in German Grammar*, pp. 65–101. Foris Publications, Dordrecht.
1993 *Deutsche Syntax – generativ.* Narr, Tübingen.
2005 How to turn German into Icelandic, and derive OV–VO contrasts. *Journal of Comparative Germanic Linguistics*, 8: 1–53.

Hoekstra, Teun
1984 *Transitivity: Grammatical Relations in Government–Binding Theory.* Foris Publications, Dordrecht.

Hoekstra, Teun and René Mulder
1990 Unergatives as Copular Verbs; Locational and Existential Predication. *The Linguistic Review*, 7: 1–79.

Holmberg, Anders and Thorbjörg Hróarsdóttir
2003 Agreement and movement in Icelandic raising constructions. *Lingua*, 114: 651–673.

Holmberg, Anders and Christer Platzack
1995 *The role of inflection in Scandinavian syntax.* Oxford University Press, New York, NY.

Iatridou, Sabine and Anthony Kroch
1992 The Licensing of CP–recursion and its Relevance to the Germanic Verb–Second Phenomenon. *Working Papers in Scandinavian Syntax*, 50: 1–24.

Johnson, Kyle
1991 Object Positions. *Natural Language & Linguistic Theory*, 9: 577–636.

Jonas, Diane and Jonathan Bobaljik
1993 Specs for Subjects: The Role of TP in Icelandic. In Jonathan D. Bobaljik and Colin Phillips, (eds.), *Papers on Case & Agreement I. MIT Working Papers in Linguistics*, 18: 59–98.

Kayne, Richard
1972 Subject Inversion in French Interrogatives. In Jean Casagrande and Bohdan Saciuk, (eds.), *Generative Studies in Romance Languages*, pp. 70–126. Newbury House, Rowley, MA.
1991 Romance Clitics, Verb Movement, and PRO. *Linguistic Inquiry*, 22: 647–686.
1994 *The Antisymmetry of Syntax*. MIT Press, Cambridge, MA.
Kayne, Richard and Jean–Yves Pollock
1978 Stylistic Inversion, Successive Cyclicity and move NP in French. *Linguistic Inquiry*, 9: 595–621.
1998 *New Thoughts on Stylistic Inversion*. Paper presented at the workshop on *Inversion in Romance*, 1998, Amsterdam.
Kinyalolo, Kasangati K. W.
1991 *Syntactic Dependencies and the Spec–Head Agreement Hypothesis in KiLega*. Ph.D. thesis, UCLA, Los Angeles, CA.
2003 *Limiting the scope of V-movement in Kilega*. Handout of a talk given at Stony Brook, New York, NY.
Levin, Beth and Malka Rappaport Hovav
1995 *Unaccusativity: At the Syntax–Lexical Semantics Interface*. MIT Press, Cambridge, MA.
McCloskey, James
1991 Clause structure, ellipsis and Proper government in Irish. *Lingua*, 85: 259–302.
1996a Subjects and Subject Positions in Irish. In Robert D. Borsley and Ian Roberts, (eds.), *The Syntax of the Celtic Languages: A Comparative Perspective*, pp. 241–283. Cambridge University Press, Cambridge.
1996b On the Scope of Verb-Movement in Irish. *Natural Language & Linguistic Theory*, 14: 47–104.
Miyagawa, Shigeru
2005 On the EPP. In Martha McGinnis and Norvin Richards, (eds.), *Perspectives on Phases: Proceedings of the workshop on EPP and Phases. MIT Working Papers in Linguistics*, 49: 201–235.
Müller, Gereon
2004 Verb-second as vP-first. *Journal of Comparative Germanic Linguistics*, 7: 179–234.
Ordoñez, Francisco
1994 Post–verbal Asymmetries in Spanish. GLOW *Newsletter*, 32: 40–41.
1997 *Word Order and Clause Structure in Spanish and other Romance languages*. Ph.D. thesis, CUNY, New York, NY.
Pesetsky, David and Ester Torrego
2004 Tense, Case and the Nature of Syntactic Categories. In Jacqueline Guéron and Jacqueline Lecarme, (eds.), *The Syntax of Time*, pp. 495–537. MIT Press, Cambridge, MA.

Pollock, Jean–Yves
1989 Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry*, 20: 365–424.
Richards, Norvin
2004 *A distinctness condition on linearization*. Unpublished Manuscript, Massachusetts Institute of Technology, Cambridge, MA.
Rivero, Maria-Luisa
1994 Verb Movement and the Structure of IP in the languages of the Balkans. *Natural Language & Linguistic Theory*, 12: 63–120.
Ross, John R.
1972 Double -ing. *Linguistic Inquiry*, 3: 6–86.
Sigurðsson, Halldor
1989 *Verbal Syntax and Case in Icelandic*. Ph.D. thesis, University of Lund.
Sportiche, Dominique
1988 A Theory of Floating Quantifiers and Its Corollaries for Constituent Structure. *Linguistic Inquiry*, 19: 425–449.
Valois, Daniel and Fernande Dupuis
1992 On the Status of (Verbal) Traces in French: The Case of Stylistic Inversion. In Paul Hirschbühler and E.F.K. Koerner, (eds.), *Romance Languages and Modern Linguistic Theory*, pp. 325–338. John Benjamins, Amsterdam.
Vikner, Sten
1994 Scandinavian Object Shift and West Germanic Scrambling. In Norbert Corver and Henk van Riemsdijk, (eds.), *Studies on Scrambling*, pp. 487–517. Mouton de Gruyter, Berlin.
1995 *Verb Movement and Expletive Subjects in the Germanic Languages*. Oxford University Press, New York, NY.
Watanabe, Akira
1996 *Case Absorption and Wh-Agreement*. Kluwer Academic Publishers, Dordrecht.
Wurmbrand, Susi
2004 *Licensing Case*. Unpublished Manuscript, University of Connecticut, Storrs, CT.

Towards a Relativized Concept of Cyclic Linearization*

Gereon Müller

1 Introduction

Suppose that a grammar is an optimal satisfaction of requirements imposed by the interfaces LF and PF (see Chomsky (2000, 2001b, 2005a)). Against this background, Fox and Pesetsky (2003, 2005a,b)) discuss a number of phenomena involving shape conservation effects with movement operations. They suggest that these effects should not be accounted for by invoking specific constraints demanding just that (as, e.g., in Müller (2001), Williams (2003)), or by syntax-internal constraints (like the Minimal Link Condition, as in Collins and Thráinsson (1996)). Rather, they should be taken to follow from an independently motivated system of cyclic linearization applying to local spell-out domains (phases). The present paper is an attempt to revise Fox and Pesetsky's system. More specifically, I will try to substantiate the following three claims. First, a cyclic linearization approach to shape conservation effects is in principle worth pursuing. Second, the specific cyclic linearization approach developed by Fox and Pesetsky faces certain empirical and conceptual problems. And third, these problems can be avoided if cyclic linearization is assumed to be relativized rather than rigid, and if more emphasis is placed on the derivational nature of the system.[1]

*For helpful comments, I would like to thank Petr Biskup, Gisbert Fanselow, Kleanthes Grohmann, Fabian Heck, Gunnar Hrafn Hrafnbjargarson, Jóhannes Gísli Jónsson, Denisa Lenertová, David Pesetsky, Wolfgang Sternefeld, and particularly Hans-Martin Gärtner, as well as the audiences of the ZAS workshop *Interfaces + Recursion = Language?* in March 2005 in Berlin, the *InterPhases* Conference in May 2006 at the University of Cyprus, and the Norms workshop on *Verb Placement* in January 2007 at the University of Iceland. I alone am responsible for any misunderstandings and errors.

[1]The concept of relativized cyclic linearization as an alternative to rigid cyclic linearization is inspired by a predecessor in the domain of locality theory, viz., relativized minimality (see Rizzi (1990), Fanselow (1991)) as an alternative to rigid minimality (cf. Chomsky (1986), Baker (1988)).

I will proceed as follows. In section 2, I outline the main features of Fox and Pesetsky's cyclic linearization approach. In section 3, I present a number of empirical and conceptual problems raised by this approach. The conclusion is that a theory of cyclic linearization should rely on (i) a strictly derivational organization of grammar, (ii) a relativization of ordering statements that are sensitive to two fundamentally different kinds of Merge operation (feature-driven vs. non-feature-driven), and (iii) a principled theory of successive-cyclic movement that incorporates the distinction among Merge operations. Section 4 then sketches such an approach to successive-cyclic movement along these lines (based on Heck and Müller (2000, 2006a)), as a prerequisite for a new approach to cyclic linearization. Finally, section 5 develops such an approach: a strictly derivational system of relativized linearization.

2 Cyclic Linearization

2.1 Basic Assumptions

Fox and Pesetsky (2003, 2005a,b) adopt a number of basic minimalist assumptions. For instance, syntactic structure is created incrementally, bottom-up, by alternating applications of external and internal Merge (i.e., Merge and Move, respectively) (see Chomsky (1995, 2001a)). Certain XPs count as special derivational units, i.e., phases (see Chomsky (2001b, 2005a)). In this kind of approach, it is standardly assumed that all syntactic movement operations must be (a) local and (b) triggered by something (typically, certain features). As for (a), Chomsky (2000, 2001b) proposes a Phase Impenetrability Condition (PIC) that restricts search space in derivations; in its strictest form, the PIC states that a syntactic operation (like movement) can only see as far as the edge (i.e., specifier and head) domain of the previous phase, and not below this area. The PIC forces movement to apply successive-cyclically, via edge domains of phases that act as escape hatches. Crucially, Fox and Pesetsky dispense with the PIC; there is no notion of escape hatch here. Rather, the necessity of successive-cyclic movement via phase edges is assumed to be derivable from cyclic linearization. As for (b), it is clear that if movement must take place via phase edges, and if all movement must be triggered, there must be a trigger (independent of locality considerations) that permits movement to phase edges. Here, one common suggestion is that optional EPP features can be inserted at phase edges if this "has an effect on outcome" (see Chomsky (2001b)). Fox and Pesetsky simply presuppose that there is some such condition that permits intermediate movement steps to phase edges, i.e., successive-cyclic movement. Thus, whereas they are concerned with deriving

the *necessity* of successive-cyclic movement (from cyclic linearization), they are not concerned with deriving the *possibility* of successive-cyclic movement.

The heart of their proposal can then be summarized as in (1).

(1) *Cyclic linearization*:
   a. Linearization of syntactic structure applies cyclically, to spell-out domains (phases).
   b. Spell-out domains are CP, VP/vP, and DP.
   c. Linearization adds new ordering statements to the set of statements established by the linearization of previous spell-out domains.
   d. A new ordering statement generated in a spell-out domain must not contradict an ordering statement of a previous spell-out domain.

It follows from the system of cyclic linearization in (1) that shape conservation effects emerge as by-products: The linear ordering of items is regulated by external and internal Merge operations within a spell-out domain (and can repeatedly be changed within this domain), but it is fixed for the remainder of the derivation at the end of each spell-out domain. Essentially, spell-out domains (phases) act as stages of the derivation where a photograph is taken, and the information provided by this photograph is filed away and cannot be contradicted later in the derivation. Section 2.2 shows how this system of cyclic linearization derives the necessity of successive-cyclic movement.

2.2 Successive-Cyclic Wh-Movement

(2-a) illustrates a typical long-distance *wh*-dependency that crosses an embedded CP; and (2-b) exemplifies a highly local *wh*-movement operation. Under traditional assumptions (see Chomsky (1973)), (2-a) involves an intermediate movement step to the edge of the embedded clause (the CP phase). Given that vP/VP is also a phase, there must also be intermediate movement steps to the edge of this domain in (2-a) and (2-b).[2]

(2) a. $[_{CP}$ What$_1$ do you $[_{VP}$ t_1''' think $[_{CP}$ t_1'' that she $[_{VP}$ t_1' read t_1]]]] ?
   b. $[_{CP}$ What$_1$ did she $[_{VP}$ t_1' read t_1]] ?

[2]Fox and Pesetsky (2005a) tentatively assume that it is in fact VP that acts as a spell-out domain (phase) in a language like English; see below.

As noted above, it must be ensured that intermediate movement steps are both possible and necessary. One option for deriving the possibility is a condition like (3), here called Optional EPP Feature Condition (OFC), that controls the insertion of optional EPP features (see Chomsky (2000, 2001b)).

(3) *Optional EPP Feature Condition* (OFC):
The head X of phase XP may be assigned an EPP-feature (after the phase XP is otherwise complete), but only if that has an effect on outcome.

It remains to be shown that the necessity of successive-cyclic movement via phase edges follows from cyclic linearization. Consider first a derivation in which *wh*-movement skips a phase edge on the way to its ultimate landing site, as in (4), which yields (2-b) without successive-cyclic movement.

(4) a. $[_{VP}$ read what$_1$ $]$ → *read* < *what*
b. $[_{CP}$ what$_1$ did she $[_{VP}$ read t$_1$ $]]$ → **what* < *read*

There is no *wh*-movement to the edge of the spell-out domain VP in (4-a) (such movement is optional, given the OFC). Consequently, linearization of VP generates the ordering statement *read* < *what*, which must not be contradicted in the remainder of the derivation. However, it *is* contradicted by (regular, feature-driven) *wh*-movement to the edge of CP in (4-b): Linearization of CP generates the ordering statement *what* < *read*, and ungrammaticality arises due to conflicting ordering statements. The situation is different if *wh*-movement applies successive-cyclically, as in (5).

(5) a. $[_{VP}$ what$_1$ read t$_1$ $]$ → *what* < *read*
b. $[_{CP}$ what$_1$ did she $[_{VP}$ t$'_1$ read t$_1$ $]]$ → *what* < *read*

Here, movement to the edge of VP applies first; therefore, VP linearization generates the ordering statement *what* < *read*. Subsequent movement and linearization in the CP domain generates the very same odering statement *what* < *read* – if the *wh*-phrase precedes VP-internal material in the VP domain, it will also precede that material in the CP domain. More generally, it now follows that a *wh*-phrase originating in a non-edge position of VP can only end up in a SpecC position (where it precedes all other items of a clause) without contradicting the ordering statements for the spell-out domain VP if it first

moves to the left-peripheral edge in VP.[3] Thus, successive-cyclic movement is effected without recourse to a specific constraint like the PIC.[4]

2.3 Simple Object Shift and Holmberg's Generalization

A second welcome consequence of Fox and Pesetsky's system of cyclic linearization is that shape conservation effects with object shift can be straightforwardly derived. Relevant generalizations that characterize the operation of (simple) object shift in the Scandinavian languages include the following (see Holmberg (1986, 1998), Vikner (1990, 1994), Collins and Thráinsson (1996)): Object shift moves DPs out of the VP, to a higher position that is TP-internal. The operation can affect (unstressed) pronominal DPs and non-pronominal DPs. In Icelandic, both types of DPs can undergo object shift; in the Mailand Scandinavian languages, only (unstressed) pronominal DPs can undergo object shift. The property of object shift that has arguably received most attention in the literature is its dependence on main verb raising to a position in front of the shifted object ('Holmberg's Generalization'). Pronominal object shift is obligatory (in contexts where it respects Holmberg's Generalization); non-pronominal object shift is optional throughout.

The obligatoriness of pronominal object shift in Danish is illustrated by the contrast between (6-a) and (6-b); the operation's dependence on main verb raising is exemplified by the examples in (6-cd), which involve a finite auxiliary and a non-finite main verb, and hence an absence of main verb raising in verb-second contexts.[5]

(6) a. *Hvorfor købte$_V$ Peter – ikke t$_V$ den$_1$?
why bought Peter not it

b. Hvorfor købte$_V$ Peter den$_1$ ikke t$_V$ t$_1$?
why bought Peter it not

[3]Note that it follows that items which are already left-peripheral in a given spell-out domain as a result of external Merge or regular, feature-driven (i.e., not OFC-triggerd) internal Merge do not have to move to a specific escape hatch. This issue can become relevant with, e.g., *wh*-objects and verbs in SOV languages (depending on whether vP or VP acts as a spell-out domain).

[4]There are, however, certain PIC effects that cannot be derived from cyclic linearization. One example is the *wh*-island sensitivity of empty operator movement in, e.g., English relative clauses; see Heck (2005).

[5]Mainland Scandinavian languages lack V-to-T movement but exhibit V-(to-T-)to-C movement in verb-second contexts; consequently, object shift can only take place in verb-second contexts in these languages. The situation is different in Icelandic, which has standard (i.e., non-intermediate) V-to-T movement.

c. Hvorfor skal Peter ikke købe den$_1$?
why shall Peter not buy it
d. *Hvorfor skal Peter den$_1$ ikke købe t$_1$?
why shall Peter it not buy

Similarly, the examples in (7-ab) illustrate the obligatoriness of pronominal object shift in Swedish.[6] The examples in (7-cd) and (7-ef) show that pronominal object shift in Swedish is impossible without raising of the main verb, in verb-second clauses where an auxiliary is the finite verb and in embedded clauses without verb-second, respectively.

(7) a. (*)Jag kysste inte henne$_1$
I kissed not her
b. Jag kysste henne$_1$ inte t$_1$
I kissed her not
c. Jag har inte kysst henne$_1$
I have not kissed her
d. *Jag har henne$_1$ inte kysst t$_1$
I have her not kissed
e. att jag inte kysste henne$_1$
that I not kissed her
f. *att jag henne$_1$ inte kysste t$_1$
that I her not kissed

The data in (8-ab) show that non-pronominal object shift in Icelandic is optional, and the data in (8-cd) illustrate that this type of object shift requires main verb raising, too.

(8) a. Jón las ekki bækurnar$_1$
Jón read not the.books
b. Jón las bækurnar$_1$ ekki t$_1$
Jón read the.books not
c. Jón hefur ekki lesið bækurnar$_1$
Jón has not read the.books
d. *Jón hefur bækurnar$_1$ ekki lesið t$_1$
Jón has the.books not read

Following earlier work by Williams (2003), Fox and Pesetsky (2005a) assume

[6]Note, however, that there is some variation in judgements concerning the lack of pronominal object shift in main verb raising contexts.

that Holmberg's Generalization should be viewed as a shape conservation effect with object shift. Two assumptions prove crucial: First, the landing site of object shift is outside the spell-out domain (phase) VP (but below TP). Second, object shift cannot target a phase-edge position as an intermediate landing site (in contrast to *wh*-movement). It is this second assumption in particular that is indispensible if the shape conservation effect with object shift is to be derived from cyclic linearization.[7] On this basis, Holmberg's Generalization follows without further ado. Linearization of VP invariably generates an ordering statement V < DP (optional intermediate movement to the phase edge is not available with object shift). This ordering statement remains present througout the rest of the derivation; and it must not be contradicted by another ordering statement that is generated later. However, if object shift out of VP takes place and is not accompanied by further raising of the main verb, subsequent linearization of CP generates a contradictory ordering statement DP < V, and ungrammaticality results. If, on the other hand, object shift is accompanied by further raising of the main verb to a higher position, subsequent linearization of CP generates an ordering statement V < DP that is identical to the statement generated earlier; so the movement operation is licit. Thus, Holmberg's Generalization is derived as a shape conservation effect that follows automatically from general assumptions about cyclic linearization.

2.4 Multiple Object Shift

There is a similar shape conservation effect with multiple object shift of pronouns and non-pronominal DPs in double object constructions (see Vikner (1990), Johnson (1991), and Collins and Thráinsson (1996), among others). In Scandinavian double object constructions, the order is indirect object before direct object. This order must be strictly respected by multiple object shift of two pronouns in Danish; cf. (9).

(9) a. Peter viste hende$_1$ den$_2$ jo t_1 t_2
Peter showed her it indeed
b. *Peter viste den$_2$ hende$_1$ jo t_1 t_2
c. *Peter viste – – jo hende$_1$ den$_2$

[7]There are various ways to motivate this assumption. One possibility would be to invoke the fact that object shift has A-movement properties (see Vikner (1994)); if a phase-edge position qualifies as an A-bar position, an intermediate movement step to the edge of the spell-out domain in the course of object shift might induce an improper movement effect (reducible to Principle C of the binding theory, as in Chomsky (1981), or to the Principle of Unambiguous Binding in Müller and Sternefeld (1993)).

d. *Peter viste – – jo den_2 $hende_1$
e. *Peter viste $hende_1$ – jo t_1 den_2
f. *Peter viste – den_2 jo $hende_1$

If only one of the two objects is a pronoun, it must be the indirect object; object shift of a direct object pronoun across a non-pronominal indirect object fails to preserve the VP-internal order and is blocked; see (10).

(10) a. *Peter viste den_2 jo $Marie_1$ t_2
Peter showed it indeed Marie
b. Peter viste $hende_1$ jo t_1 $bogen_2$
Peter showed her indeed the.book

Essentially the same situation obtains with multiple optional object shift of non-pronominal DPs in Icelandic. The examples in (11) show that multiple object shift must preserve the VP-internal order; and the examples in (12) show that if only one of the two objects undergoes object shift, it must be the indirect object, so that the pre-movement order is maintained after the operation.

(11) a. Ég lána $Maríu_1$ $bækurnar_2$ ekki t_1 t_2
I lend Maria the.books not
b. *Ég lána $bækurnar_2$ $Maríu_1$ ekki t_1 t_2
I lend the.books Maria not

(12) a. *Ég lána $bækurnar_2$ ekki $Maríu_1$ t_2
I lend the.books not Maria
b. Ég lána $Maríu_1$ ekki t_1 $bækurnar_2$
I lend Maria not the.books

The account that can be given in Fox and Pesetsky's framework is exactly as before: Linearization of VP generates the three ordering statements V $<$ DP_{IO}, V $<$ DP_{DO}, and DP_{IO} $<$ DP_{DO}. These ordering statements can only be respected after (multiple) object shift if (a) the main V moves to a higher position in front of both objects, and (b) the two shifted objects reassemble in their pre-movement order: V $<$ DP_{IO}, V $<$ DP_{DO}, and DP_{IO} $<$ DP_{DO}. A derivation in which the main verb fails to move contradicts an ordering statement in the VP domain, as with simple object shift (Holmberg's Generalization): DP_{IO} $<$ V, DP_{DO} $<$ V. However, for the same reasons, a derivation in which the two shifted objects fail to preserve the pre-movement order established in VP is also ruled out: DP_{DO} $<$ DP_{IO} contradicts the earlier linearization statement DP_{IO} $<$ DP_{DO}.

To sum up so far, Fox and Pesetsky's set of simple assumptions about cyclic linearization succeeds in deriving the necessity of intermediate movement steps to phase edges with *wh*-movement (more generally, A-bar movement) without invoking a special constraint like the PIC, and it accounts for the shape-preserving nature of object shift without invoking a special requirement demanding just that (like the Shape Conservation constraints in Williams (2003), Müller (2000)), given that the former movement type can target phase edges as intermediate landing sites, and the latter one cannot. Still, Fox and Pesetsky's approach also raises a number of questions.

3 Problems

3.1 Spell-Out Domains

A first potential problem is related to the notion of spell-out domain. As noted in footnote 2, Fox and Pesetsky (2005a) assume that it is VP rather than vP that acts as a spell-out domain in English and Scandinavian; thus, this spell-out domain does not correspond to the notion of phase in Chomsky (2000, 2001b). As things stand, this assumption is unavoidable. To see this, suppose that vP is the spell-out domain (and VP is not). Then, an ordering statement $DP_S <$ V established in vP (with the external argument DP merged in Specv) would invariably be contradicted by a possible later ordering statement $V < DP_S$ after verb-second movement to C. This problem is illustrated for verb-second movement *wh*-questions in Danish in (13). The external argument precedes the finite verb in vP in (13-a), and follows the finite verb in (13-b). However, the first linearization statement is not generated if VP rather than vP is the relevant spell-out domain (the external argument has not yet been merged when VP linearization takes place).

(13) a. $[_{vP}$ Peter købte den]
Peter bought it

b. $[_{CP}$ Hvorfor købte$_0$ $[_{TP}$ Peter$_2$ t'_0 den$_1$ ikke $[_{vP}$ t_2 t_0 t_1]]] ?
why bought Peter it not

Thus, the assumption that VP rather than vP is the spell-out domain in Danish seems crucial.[8] This may be incompatible with the idea that phases are semantically motivated as propositional units (see Chomsky (2000, p. 106)), which

[8] See, however, Fox and Pesetsky (2003) for a more elaborate approach that relies on a set of further assumptions (concerning overt vs. covert movement) under which this consequence does not hold.

then includes the base position of the external argument. It follows that if VP is a spell-out domain, either spell-out domains do not have to equal phases, or phases are not semantically motivated in the way Chomsky suggests. Furthermore, there is independent phonological motivation for vPs as phases (see Legate (2003), Ishihara (2004), Richards (2004), Kratzer and Selkirk (2007))). Finally, Fox and Pesetsky (2005a) (based on Ko (2004)) suggest that there may be a parametrization with respect to vP or VP as the relevant spell-out domain (with Korean choosing the former option, English and Scandinavian the latter one); again, this may possibly be considered problematic from a conceptual point of view if one assumes a semantic grounding of phases.

More importantly, there are empirical problems with the assumption that VP (and not vP) is the relevant spell-out domain. The evidence comes from shape conservation phenomena with multiple pronominal object shift in German and multiple *wh*-movement in Bulgarian. In both cases, there is good reason to assume (based on Fox and Pesetsky's general assumptions) that an ordering statement must be generated for an external argument in Specv and an object in the complement position of V. I begin with pronoun fronting in German.

3.1.1 Pronoun Fronting in German

The following generalizations describe the basic behaviour of unstressed personal pronouns in German (see, among others, Bierwisch (1963, pp. 99–101), Lenerz (1977), and Heidolph et al. (1981)). Unstressed pronouns are obligatorily moved across adverbs and non-pronominal DPs; see (14-ab) vs. (14-c). There is only one kind of element that may precede unstressed pronouns and follow C (i.e., remain TP-internal), and that is an external argument bearing nominative case (i.e., a subject); cf. (14-de). This follows if scrambling targets a lower position than pronoun fronting, and if only external arguments bearing nominative case may undergo movement to the subject position SpecT; this latter movement is always optional.

(14) a. *dass gestern der Fritz es_1 gelesen hat
that yesterday the Fritz it read has

b. *dass gestern es_1 der Fritz t_1 gelesen hat
that yesterday it the Fritz read has

c. dass es_1 gestern der Fritz t_1 gelesen hat
that it yesterday the Fritz read has

d. dass der Fritz es_1 gestern t_1 gelesen hat
that the Fritz it yesterday read has

e. *dass der Fritz gestern es$_1$ gelesen hat
that the Fritz yesterday it read has

If more than one DP argument in a clause is an unstressed pronoun, pronoun fronting applies multiply. As with Scandinavian object shift, there are shape conservation effects (see Müller (2001)). These effects occur, e.g., when the two unstressed pronouns are subject and object: The pre-movement order $DP_S < DP_O$ must strictly be adhered to after multiple pronoun fronting. This is illustrated by the sentence pair in (15).

(15) a. dass sie$_1$ es$_2$ gestern t_1 t_2 gelesen hat
that she.NOM it.ACC yesterday read has
b. *dass es$_2$ sie$_1$ gestern t_1 t_2 gelesen hat
that it.ACC she.NOM yesterday read has

The strict $DP_S < DP_O$ order after pronoun fronting of subject and object cannot be due to obligatory subject raising (as it may be in comparable cases in the Scandinavian languages) because, as we have just seen (cf. (14-c)), German does not have obligatory raising to subject position. However, the phenomenon can straightforwardly be derived from cyclic linearization – but only if an external argument merged in Specv and an internal argument merged with V have a common spell-out domain. The relevant domain must then be vP. Of course, this reasoning raises immediate questions within Fox and Pesetsky's framework. Most notably, how can an object pronoun move across a subject DP in Specv in the first place? Assuming Specv to be available as an intermediate landing site for German pronoun fronting cannot be the solution because we would then not expect any shape conservation effect with pronoun fronting. So we face a dilemma: The shape conservation effect in (15-b) strongly suggests cyclic linearization at work; but then, it becomes unclear how sentences like (14-c) can be permitted. Furthermore, since German has verb-second clauses in which a finite main verb in C precedes a subject (in SpecT or in Specv), assuming vP to be a spell-out domain in German creates the very same problem that it does in Scandinavian languages like Danish (see (13)).

In what follows, I will assume that the shape conservation effects with Scandinavian object shift and German pronoun fronting can and should be derived in essentially the same way (but note that this does not imply that the two operations are identical).

3.1.2 Wh-Movement in Bulgarian

Wh-movement in Bulgarian exhibits the following properties (cf., among others, Rudin (1988), Billings and Rudin (1996), Grewendorf (2001), Richards (2001), and Bošković (2002b)): All *wh*-phrases are fronted to the C domain in multiple questions. An (agentive) *wh*-subject (which we may take to be the prototypical external argument merged in Specv) and a *wh*-object exhibit strict shape conservation; they always show up in the order $DP_S < DP_O$. The situation is somewhat more involved with two objects (an indirect object DP_{IO} and a direct object DP_{DO}). Here, the order is often $DP_{IO} < DP_{DO}$, but there are intervening factors (related to, e.g., animacy and categorial status (DP vs. PP)), and there is often optionality. Here and henceforth, I will assume that the variation in linear order in the C domain observed with multiple *wh*-movement applying to two objects can be traced back to variation that exists already within VP.[9] The generalization that is more interesting in the present context concerns the fixed linear order of subject and object *wh*-phrase in the C domain; compare, e.g., (16):

(16) a. $[_{CP}$ Koj_1 $kogo_2$ C $[_{vP}$ t_1 vižda t_2]] ?
who.NOM whom.ACC sees

b. *$[_{CP}$ $Kogo_2$ koj_1 C $[_{vP}$ t_1 vižda t_2]] ?
whom.ACC who.NOM sees

As before, on the basis of Fox and Pesetsky's general assumptions, the clear shape conservation effect that can be observed here suggests that subjects and objects are part of one linearization domain, which then implies that vP is indeed a spell-out domain (note that this holds independently of whether or not there is raising to SpecT). And again, this assumption is incompatible with V-to-C movement in front of subject DPs – an operation that is available in Bulgarian as it is in Danish or German.

3.2 A-Movement in Passive Constructions

A second problem arises under the assumption that unaccusative, passive and raising vPs (or VPs – the difference is immaterial in the present context) are phases (spell-out domains); this assumption has recently been argued for from different empirical and theoretical perspectives (see, e.g., Legate (2003) and Richards (2004)). If unaccusative and passive vPs are spell-out domains, Fox and Pesetsky's analysis makes wrong predictions: An ordering

[9] This assumption may not be entirely uncontroversial, but it is not crucial here either.

statement V < DP_O in the vP/VP domain is later followed by a reverse ordering statement DP_O < V in the CP domain, and we should thus wrongly expect subject movement to SpecT in passive or raising constructions to lead to ungrammaticality (assuming that standard A-movement, like object shift, cannot proceed successive-cyclically, via the edge of the spell-out domain). Consider the derivations in (17-a) (passive) and (17-b) (raising) in English (with (i) the lower spell-out domain, and (ii) the higher, final spell-out domain).

(17) a. (i) $[_{vP}$ hit-v $[_{VP}$ t $John_1$]]
(ii) $[_{CP}$ C $[_{TP}$ $John_1$ was $[_{vP}$ hit-v [t t_1]]]]
b. (i) $[_{vP}$ v $[_{VP}$ seems $[_{TP}$ $Mary_1$ to be smart]]]
(ii) $[_{CP}$ C $[_{TP}$ $Mary_1$ T $[_{vP}$ seems-v $[_{TP}$ t_1 to be smart]]]]

As observed by Bobaljik (2005, p. 121), the same phenomenon shows up with A-movement in passivized double object constructions in Icelandic; but here the problem might be even more pressing because in addition to a violation of shape conservation with DP_1 in SpecT and the verb, there is what might be interpreted as a shape conservation effect among the two DPs (only DP_1 can move to SpecT; DP_2 cannot undergo such movement).[10] Consider the derivation in (18), with the vP/VP spell-out domain in (18-a), and the CP spell-out domain in (18-b) (data from Zaenen et al. (1985)).

(18) a. $[_{vP}$ v $[_{VP}$ gefnar konunginum$_1$ ambáttir$_2$]]
given the.king.DAT slaves.NOM
b. $[_{CP}$ Um veturinn voru $[_{TP}$ konunginum$_1$ $[_{vP}$ gefnar-v $[_{VP}$ t
in the.winter were.PL the.king.DAT given
t_1 ambáttir$_2$]]]]
slaves.NOM

3.3 Verb-Second in SOV languages

So far, we have seen that the existence of verb-second constructions in SVO languages necessitates the assumption that VP rather than vP is the spell-out domain (so that an external argument in Specv can end up in a position that follows a verb that has undergone movement to C); and we have also seen that the same consequence holds for SOV languages. Now I will argue that the problem posed by verb-second constructions in SOV languages is in fact

[10]However, this may alternatively be accounted for as a minimality effect, derivable from a constraint like the Minimal Link Condition.

much more general than that, and cannot be solved by making assumptions concerning the parametrization of spell-out domains (vP vs. VP).

Verb-second in an SOV language like German may systematically reverse the ordering statements of the lower spell-out domain. The two examples in (19-ab) show this for subjects merged in vP; V follows DP_1 in the lower spell-out domain vP (embedded clauses have verb-final order), and V precedes DP_1 on the CP cycle. This case can in principle be handled in Fox and Pesetsky's system if we assume that VP is the relevant spell-out domain (but recall that there is evidence from pronoun fronting which indicates that subjects do belong to the lower spell-out domain).

(19) a. Gestern arbeitete$_0$ [$_{vP}$ Maria$_1$ [$_{VP}$ zu Hause t_0]]
yesterday worked Maria.NOM at home
b. dass [$_{vP}$ Maria$_1$ [$_{VP}$ zu Hause arbeitete$_0$]]
that Maria.NOM at home worked

However, the verb-second construction in (20-a) cannot be accounted for by parametrizing spell-out domains: As evidenced by the analogous construction without verb-second in (20-b), V follows DP_2 in the lower spell-out domain (be it vP or VP) in (20-a) but precedes DP_2 in the higher spell-out domain (CP). Thus, the only kind of verb-second clause that could be derived on the basis of the VP in (20-b) in accordance with Fox and Pesetsky's assumptions (given that vP is not a spell-out domain) would be (20-c), with a topicalized object.

(20) a. Maria$_1$ las$_0$ [$_{vP}$ t_1 [$_{VP}$ das Buch$_2$ t_0]]
Maria.NOM read the book.ACC
b. dass [$_{vP}$ Maria$_1$ [$_{VP}$ das Buch$_2$ las$_0$]]
that Maria.NOM the book.ACC read
c. Das Buch$_2$ las$_0$ [$_{vP}$ Maria$_1$ [$_{VP}$ t_2 t_0]]
the book.ACC read Maria.NOM

To conclude, the order reversal involving a finite verb and an object that takes place between the verb-final VP (or vP) domain and the verb-second CP domain poses a problem for Fox and Pesetsky's analysis.[11] Possible solutions

[11] Note that this problem persists under an antisymmetric approach according to which OV predicate phrases are derived from a basic VO structure (see Kayne (1994)), as long as we make the assumption that the derived OV structure then feeds further derivational steps on, e.g., the TP and CP cycles (cf. Zwart (1993)). However, things are different if this latter assumption is not made; see Hallman (2000) and the following remarks in the main text.

would then minimally seem to require non-trivial and somewhat unusual assumptions about verb placement in SOV languages with verb-second.

Here is a first attempt at a solution. Suppose that SOV languages have a uniform base order where the verb precedes the object (V < DP), as contemplated in the last footnote. Suppose next that an object DP moves across the verb before the spell-out domain is reached in verb-final clauses (as in (20-b)); and that the object DP fails to undergo such movement to a linearization-domain internal pre-verbal position in verb second clauses (as in (20-a)) – unless, that is, the object DP eventually ends up in front of the verb in topic position, as in (20-c), in which case it has to undergo local movement across the verb. Such an approach might work, but it would arguably be somewhat ad hoc; and it is far from clear how verb movement to C could block local object movement (as a kind of Anti-Holmberg's Generalization effect) in most contexts and force local object movement in contexts where the object is eventually topicalized.

Let me therefore consider a second solution, one that does not rely on local object movement across the verb, but rather on local verb movement across the object. Suppose, as before, that SOV languages have a uniform base order; however, this time the uniform order is DP < V. In verb-second clauses like (20-c), the verb raises to a position in front of the DP before the spell-out domain is reached, so there is a V > DP linearization on the vP/VP cycle that subsequent verb-second movement adheres to. On the other hand, the verb stays in situ in verb-final clauses like (20-b). For object-initial verb-second clauses like (20-c), it can then be assumed that the object must move to a phase-edge position first if it is to undergo movement to SpecC later, and since it thus precedes a locally moved verb before spell-out of the lower linearization domain takes place, the pattern vivsible in (20) is derived.

Still, this kind of approach is not completely unproblematic either. The main problem I see is this: The apparent instance of order reversal that is visible in (20) can only be accounted for if an extremely abstract base structure of the vPs in question is assumed. However, this assumption threatens to undermine the whole approach: If highly abstract linearization domains (that are never attested on the surface) are available for the SOV language German, one might wonder why they are not available for the Scandinavian SVO languages – in the latter, a surface-oriented approach seems crucial. For these reasons, then, verb-second in SOV languages continues to pose a problem in Fox and Pesetsky's approach.[12]

[12]It should be noted that the issue of verb-second in SOV languages is eventually addressed at the very end of the extended handout that is Fox and Pesetsky (2003) (viz., on p. 44). Fox and Pesetsky suggest that "covert merge of V and object" is involved. Notwithstanding

3.4 Intermediate Landing Sites

The crucial difference between movement types that respect (a pre-theoretic concept of) shape conservation (e.g., Scandinavian object shift) and movement types that do not (e.g., English *wh*-movement) boils down to the ability or inability to move successive-cyclically via SpecC. This fact has sometimes been regarded as a problem because an important building block of the overall approach in Fox and Pesetsky (2005a) has thus been left unspecified (see Nilsen (2005), Williams (2005), Bobaljik (2005), and Müller (2005)). Fox and Pesetsky (2005a) are well aware of this; they note that "our proposals say nothing in themselves [...] about the circumstances under which movement to these left-edge positions is allowed or prohibited." As noted above (see footnote 7), the availability of an intermediate phase-edge position for *wh*-movement, and the unavailability of such a position for Scandinavian object shift, may well follow from the theory of improper movement, given an A/A-bar distinction among movement types. Now, such an approach is not entirely unproblematic. For instance, it fails to account for shape conservation effects with pronoun fronting in German (which shows A-bar properties like parasitic gap licensing; see Vikner (1994) and below). It also does not provide an immediate account of the fact that A-movement in passive contexts can to some extent violate shape conservation (viz., with respect to the verb). However, it seems to me that these problems of classifying the circumstances under which a movement operation may target an intermediate phase-edge position are to some extent orthogonal to Fox and Pesetsky's main claims, and may in principle be solved within their approach.

Still, I would like to contend that there is a much more general conceptual problem lurking here. As has been noted by Sells (2001) and Richards (2004), it does not seem accidental that the shape conservation property of object shift (with respect to V) is correlated with the fact that this operation is extremely local, and that the failure of, say, *wh*-movement to obey shape conservation (with respect to V) is correlated with the fact that this operation is inherently non-local (and can apply long-distance). In fact, even A-movement to SpecT is less local than object shift, and, as noted above, it does not obey shape conservation with respect to V. Thus, the correct generalization underlying shape conservation effects with respect to V does not involve a concept like

---

other problems with the concept of covert operations under Fox and Pesetsky's general set of assumptions (concerning the interaction of quantifier raising and deletion; see Heck (2005)), this can be taken as evidence that verb-second in SOV languages is a priori unexpected, and requires extra assumptions, in their system of cyclic linearization.

the A-bar vs. A distinction; rather, the correct generalization seems to be that only extremely local operations (i.e., operations whose final landing site is very close to the base position) obey shape conservation with respect to V. This generalization cannot be captured in Fox and Pesetsky's approach in a natural way.

3.5 Derivational Syntax

Finally, I would like to point out a potential conceptual peculiarity. Fox and Pesetsky (2005a) assume a derivational organization of syntax. However, closer inspection reveals that there is in fact a large representational residue (also see Sells (2005)). First, note that all ordering statements that have been generated for a given spell-out domain remain active and visible throughout the rest of the derivation. Arguably, in a strictly derivational approach, information that has undergone cyclic spell-out should become inaccessible, and irrelevant, at subsequent stages of the derivation (see, e.g., Epstein and Seely (2002)). Second, recall that the *Phase Impenetrability Condition* (PIC) is abandoned in Fox and Pesetsky's (2005a) approach. The PIC derives locality constraints on movement; however, it is motivated in Chomsky (2000, 2001b) primarily by complexity considerations (viz., reduction of search space) rather than empirically, as a locality constraint. The latter task is accomplished by cyclic linearization in the present approach; but the former task is not: Search space is in principle unlimited here. Thus, it is not so clear that dispensing with the PIC is a virtue, assuming that the PIC's main task is that of reducing search space.

To conclude this section, I have argued that there are a number of empirical and conceptual problems with Fox and Pesetsky's system of cyclic linearization. On the empirical side, the approach turns out to be both too strong (e.g., it rules out non-object initial verb-second in SOV languages, and A-movement in passive contexts on the assumption that a local spell-out domain is involved here), and too weak (e.g., it does not derive shape conservation effects with pronoun fronting in German and *wh*-movement in Bulgarian). On the conceptual side, the approach relies on a parametrization of spell-out domains (VP vs. vP) that may be regarded as conceptually suboptimal (and also empirically problematic, as argued for pronoun fronting in German). Furthermore, it does not capture the generalization that only extremely local movement types exhibit shape conservation effects with respect to V. Finally, by providing syntactic access to the complete representation built up so far at every step of the derivation, there is a theoretical heterogeneity that may be regarded as

conceptually unattractive. In view of this, I will attempt to revise Fox and Pesetsky's (2005a) approach to shape conservation effects in terms of cyclic linearization.[13]

The revised approach to cyclic linearization as a source of shape conservation effects to be developed in what follows relies on three main assumptions. The first assumption is that syntax exhibits a strictly derivational organization. This implies that pieces of information that are accessible at one stage of the derivation (including ordering statements) may be inaccessible at later stages; i.e., information may be lost in the course of the derivation. Second, the approach is based on a relativization of ordering statements: Instead of fixed spell-out domains (like vP/VP and CP), the creation of ordering statements is assumed to be a relativized property of two structure-building operations – feature-driven vs. non-feature-driven Merge.[14] Third and finally, a theory of successive-cyclic movement is presupposed which captures the difference between movement that reaches its target position and movement that does not, and which has the effect of forcing *all* movement to apply successive-cyclically, in steps that are highly local. Since the theory of successive-cyclicity is a prerequisite for the derivational, relativized approach to cyclic linearization, it is addressed first, in the following section.

4 Successive-Cyclic Movement

I assume the (stricter version of the) Phrase Impenetrability Condition introduced in Chomsky (2000, 108; 2001, 13); see (21).

(21) *Phase Impenetrability Condition* (PIC):
The domain of a head X of a phase XP is not accessible to operations outside XP; only X and its edge are accessible to such operations.

The PIC requires successive-cyclic movement to take place via phase edges. Given that unforced movement is blocked (cf. the Last Resort requirement on syntactic operations), there must be another constraint that forces movement to an intermediate position. One candidate for this constraint is what I have called the "Optional EPP Feature Condition" (OFC; see Chomsky (2000,

[13] I hasten to add that the empirical coverage will not be identical. For instance, it follows without further assumptions in Fox and Pesetsky's approach that preposing of negative quantifiers is incompatible with verb movement to a higher position in Icelandic (an Anti-Holmberg's Generalization effect); the present approach will have nothing to say about this restriction.

[14] Also see Fanselow and Lenertová (2007) for a related assumption (viz., that XPs with an operator feature are not serialized immediately).

2001b) – recall that, according to the OFC, the head X of phase XP may be assigned an EPP-feature, subject to the requirement that this has then an effect on outcome). Another possibility is the constraint Phase Balance that has first been introduced in Heck and Müller (2000) (also see Müller (2004), Fischer (2004), Heck (2004), and Heck and Müller (2006a)). It is an approach based on this latter constraint that I will adopt here.[15] Phase Balance can be defined as in (22).

(22) *Phase Balance*:
Every phase has to be balanced: For every feature [•F•] in the numeration there must be a potentially available feature [F] at the phase level.

Phase Balance presupposes a distinction between two kinds of features relevant for movement: [•F•] is a probe feature with an EPP (or, more generally, movement-inducing) property, and [F] is a matching goal feature.[16] To find

[15]One should bear in mind, however, that the alternative approach could in principle also be adopted, as long as it is ensured that the system can distinguish between regular feature-driven movement, and movement triggered by optional EPP-features. The main reason for adopting the Phase Balance-based approach here is that it is fully explicit about the circumstances under which intermediate movement is triggered, and does not resort to a vague requirement like "having an effect on output". That said, the present approach can to some extent be viewed as a formal elaboration of what it means to "have an effect on output." From this perspective, the only fundamental difference between these two approaches concerns the presence or absence of features as triggers for intermediate movement steps. However, this difference is not crucial either. Phase Balance, which will momentarily be shown to act as a trigger for non-feature-driven movement, could in principle also be viewed as a trigger for the insertion of features (which in turn force movement), in which case the OFC-based approach and the Phase Balance-based approach might eventually emerge as closely related variants of the same concept. See Heck and Müller (2006b); and Biskup (2005) for a related proposal based on feature insertion. A feature-based version of the Phase Balance approach to successive-cyclic movement will nevertheless not be pursued here, mainly because it adds unnecessary complexity to the system. (McCloskey (2002) argues that a feature-based approach is better equipped to handle morphological reflexes of successive-cyclicity, but we argue in Heck and Müller (2006a) that this is not the case under current approaches to morphological spell-out (like distributed morphology).)

[16]See Frampton and Gutman (1999), Gärtner (2002), Adger (2003), Roberts and Roussou (2002), Sternefeld (2006), and Heck and Müller (2006b) for feature systems involving this kind of diacritic (with similar, but not necessarily identical interpretations). In earlier work (cf., e.g., Heck and Müller (2000, 2006a)), we rendered probe features that are accompanied by an EPP property as [*F*] rather than [•F•]; however, we will now reserve the [*F*] notation for cases of 'pure' probe features that trigger Agree and lack an EPP property; see Heck and Müller (2006b).

out whether a phase is balanced, the derivation looks into the numeration and scans the [•F•] features still located there. For each of these [•F•] features, there must be a corresponding [F] that is potentially available. There are two ways of how an [F] feature can be potentially available for a matching [•F•] feature. The straightforward one is that [F] is present either in the numeration, or in a tree that has been constructed outside of the present derivation by a separate derivation, and is waiting to be merged with the present tree at some later point; these domains can be referred to as the workspace of the present derivation. However, there is a possibility that no matching [F] feature is present in the workspace because the item bearing [F] has already been merged in the present derivation.[17] Then, it can only be potentially available for [•F•] in the numeration if it is part of the edge domain of the current phase; and (with the exception of underived phase specifiers; but see below) the only way to ensure the presence of [F] in the edge domain of a phase is by movement of the category bearing [F]. This will then derive intermediate steps of successive-cyclic movement, in (minimal) violation of the Last Resort constraint that blocks non-feature-driven movement. The notion of potential availability can thus be defined as in (23).

(23) *Potential availability*:
A feature [F] is potentially available if (i) or (ii) holds:
(i) [F] is on X or edgeX of the present root of the derivation.
(ii) [F] is in the workspace of the derivation.

As an abstract example of how the system works, consider a stage of a derivation in which a phase Σ is created, and there is a [•wh•] feature on a C item that is still part of the numeration. Σ must be balanced, so there must be a [wh] feature that is potentially available for the [•wh•] feature waiting in the numeration. Suppose now first that there is a [wh] feature on some other item in the numeration (say, a bare *wh*-phrase like *who*), or on some other item outside the numeration but within the workspace (say, a complex *wh*-phrase like *which woman*, which may have been formed already). Then, Phase Balance does not trigger any movement to the specifier of Σ, regardless of

[17] In principle, it might also be possible that [F] is neither part of the workspace, nor of the current derivation. In that case, the derivation will crash. Arguably, such an option is unwanted from the perspective of crash-proof syntax. It can be excluded in a simple and principled way by imposing a count invariant requirement (as it is known from categorial grammar), such that the number of [•F•]s must equal the number of [F]s in the numeration; see van Benthem (1988), Stabler (1996) on count invariants.

whether there is a [wh] feature within Σ or not.[18] In contrast, suppose now that there is in fact no [wh] feature for the [•wh•] probe in the workspace; then, a phase-internal *wh*-phrase bearing [wh] must move to the specifier of the phase in order to respect Phase Balance, even if this violates the Last Resort requirement because there is no feature that might trigger the operation. This way, successive-cyclic movement from phase to phase is brought about.

Next, the question arises of what constitutes a phase. I assume that all *saturated* XPs qualify as phases. Thus, all XPs are phases, except for VP, which counts as part of the vP phase.[19] VP is special in being systematically accompanied by an additional projection vP, and it is unique among XPs in realizing certain types of arguments of its head (viz., external arguments) outside the latter's maximal projection. If every XP (except for VP) is a phase, *wh*-movement must proceed via every XP to its ultimate target position (the $C_{[\bullet wh\bullet]}$ node that attracts it).[20]

Based on this notion of phase, Phase Balance and the PIC interact as shown in (24) to generate the derivation of a simple embedded *wh*-question like *what John read*. On the vP cycle, *what* must move to a specifier because this is the only way to satisfy Phase Balance for [•wh•] on C in the numeration; [•D•] on T in the numeration has a similar requirement, which is met trivially by the external argument (but see below); furthermore, V moves to v, an operation that I assume to take place without exception (due to the deficient character of VP). Next, on the TP cycle the same reasoning applies with respect to [•wh•] as before: To ensure Phase Balance, the *wh*-phrase must be moved to T's specifier, even though this movement is not feature-driven and thus violates the Last Resort condition (in addition, the external argument moves to SpecT,

[18] As shown in Müller (2004) and Heck and Müller (2006a), this derives various types of superiority and superiority-like effects, without recourse to a constraint like the MLC – a *wh*-phrase cannot move to the specifier of a phase if there is another [wh] feature on some item outside the present derivation.

[19] Similar considerations may apply with respect to NP if there is an nP/NP distinction; see, e.g., Adger (2003).

[20] There are many predecessors of this approach; see van Riemsdijk (1978), Koster (1978), Sportiche (1989), Takahashi (1994), Agbayani (1998), Bošković (2002a), Boeckx (2003), and Chomsky (2005b). Similar concepts are employed in the SLASH feature percolation analysis of *wh*-movement employed in GPSG (see Gazdar (1981), Gazdar et al. (1985)), and in Koster's (2000) approach in terms of feature percolation in gap phrases. Also compare the remark that "phases should be as small as possible, to minimize memory for [spell-out]" in Chomsky (2001a). – Chomsky (2001b) argues that phases must be somewhat larger objects based on Agree relations holding between T and a VP-internal object (as in nominative object constructions in Icelandic). In the present approach, Agree will have to be successive-cyclic, just like movement. Also see Fischer (2004) for the same consequence with respect to binding.

as an instance of feature-driven movement). Finally, on the CP cycle, the *wh*-phrase moves to SpecC, triggered not by Phase Balance, but by the [•wh•] feature on C directly. Note that given the PIC in (21), tree-internal search space is constantly reduced in this derivation (this is indicated by striking through the domains that have become inaccessible in (24)): On the TP cycle, the complement domain of vP is not accessible anymore; and on the CP cycle, vP is invisible for syntactic operations.

(24) (I wonder) what John read

a. $[_{vP}$ what$_1$ John$_2$ v+read$_3$ $[_{VP}$ t$_3$ t$_1$]]
→ workspace: {$C_{[\bullet wh\bullet]}$, $T_{[\bullet D\bullet]}$}

b. $[_{TP}$ what$_1$ John$_2$ T $[_{vP}$ t$'_1$ t$_2$ v+read$_3$ ~~[$_{vP}$ t$_3$ t$_1$]~~]]
→ workspace: {$C_{[\bullet wh\bullet]}$}

c. $[_{CP}$ what$_1$ C $[_{TP}$ t$''_1$ John$_2$ T ~~[$_{vP}$ t$'_1$ t$_2$ v+read$_3$ [$_{vP}$ t$_3$ t$_1$]]~~]]
→ workspace: {–}

This may suffice as a brief illustration of the Phase Balance approach to successive-cyclic movement. There are two consequences of this analysis that will prove important for the approach in terms of relativized linearization that is developed in the next section: First, there are now two types of movement that can be clearly distinguished: On the one hand, there is feature-driven movement; on the other hand, there is movement that is brought about by Phase Balance. This difference will make it possible to relativize ordering statements. Second, in interaction with the narrow concept of phases employed here, the PIC drastically reduces search space. This opens up the possibility that ordering statements that were generated at some earlier point in the derivation can be forgotten at later stages because they are rendered invisible by the PIC.

With this approach to successive cyclicity as background, let me now turn to the system of relativized linearization itself.

5 Relativized Linearization

5.1 Analysis

Suppose first that syntactic representations do not tolerate contradictory ordering statements, exactly as assumed by Fox and Pesetsky (2003, 2005a,b). However, note that this does not imply that contradictory ordering statements can never be generated in the course of the derivation. Contradictory statements $\alpha < \beta$ and $\beta < \alpha$ can be generated under this assumption as long as

only one of the two statements is accessible in the syntax at any given point of the derivation. Syntactic inaccessibility of an ordering statement follows from cyclic spell-out: As soon as a phase is completed, the domain of its head, including all ordering statements generated for this domain, is spelled out and thus rendered inaccessible for further syntactic operations, in accordance with the PIC. This opens up the possibility of a new, contradictory statement being generated in the syntax. When this statement is sent to phonology, it replaces the older, contradictory one there (i.e., whereas syntax does not permit contradictory ordering statements, phonology can handle this situation by deleting the older statement). Thus, it is the fact that information can be lost in the course of the derivation (which in turn is a characteristic property of derivational systems) that makes it possible to assume that contradictory ordering statements can arise in syntax.

Next, the question arises of when and how ordering statements are generated. I assume that the domain in which ordering statements are generated is extremely local: It is the syntactic Merge operation.[21] The ordering statements themselves are generated according to precedence rules which are partly language-specific; e.g.: A head precedes its complement in English; a head follows its complement in Korean; and a [+V] head (V, v, A, T) follows its complement, whereas a [–V] head (N, D, P, C) precedes its complement in German. Crucially, now, ordering statements are relativized, in the sense that only a subset of the ordering statements that could in principle be generated are in fact generated by syntactic operations. The underlying hypothesis is that the system of cyclic spell-out exhibits optimal design: Given the ubiquity of displacement operations in syntax, a system of cyclic spell-out that does not take this into account by selectively ignoring possible ordering statements that will invariably have to be undone (because an item that participates in the ordering relation will subsequently have to move and thereby likely reverse the original order) would exhibit poor design.[22] Thus, I would like to suggest that a Merge operation applying to two categories α and β generates a linearization statement for α and β, and for items that are dominated by α or β, except for those categories that involve a difference in Merge status. To see how this systems works, consider first two general constraints that are standardly taken to govern syntactic linearization: the Exclusivity Condition in (25-a) and the Nontangling Condition in (25-b) (see Partee et al. (1993, p. 440)).

[21] See Epstein and Seely (2002) for a related but more general assumption.

[22] Also compare the discussion of generalization (I) in Chomsky (2005a, p. 21).

(25) a. *Exclusivity Condition*:
In any well-formed constituent structure tree, for any nodes x and y, x and y stand in the precedence relation P (i.e., either $<x,y> \in P$ or $<y,x> \in P$) iff x and y do not stand in the dominance relation D (i.e., neither $<x,y> \in D$ nor $<y,x> \in D$).

b. *Nontangling Condition*:
In any well-formed constituent structure tree, for any nodes x and y, if x precedes y, then all nodes dominated by x precede all nodes dominated by y.

In the present analysis, the incompatibility of dominance and precedence expressed by the Exclusivity Condition follows from the fact that ordering statements are tied to Merge operations, which do not involve dominance; the total ordering of terminals implied in the Exclusivity Condition follows because, at the end of the derivation, all categories will have participated in ordering statements. As for the Nontangling Condition, I will assume that, while basically valid, it does not hold without exception; the instances where it does not hold in the derivation are accounted for by the concept of relativized linearization, to which I now turn.

(26) *Relativized Linearization*:
For all categories x reflexively dominated by a category α and for all categories y reflexively dominated by a category β, Merge (α,β) generates an ordering statement for $<x,y>$ if x and y have an identical Merge status.

The final notion to be clarified here is the concept of "identical Merge status." At this point, the approach to successive-cyclic movement sketched in section 4 becomes important. In this approach, it is possible to formally distinguish between feature-driven and non-feature-driven Merge and Move operations; and this is what underlies the differences in Merge status. Thus, I assume that there are two types of Merge status: A category α has Merge status $[+\psi]$ at a given stage of the derivation if the reason why it shows up in its present position is related to a feature in the local environment, and α has Merge status $[-\psi]$ at a given stage of the derivation if the reason why it shows up in its present position is related to a feature that is not present in the local environment (but that exists in the numeration). External, basic Merge patterns with $[+\psi]$-marked internal Merge (i.e., Move), and so it would seem natural to assume that external Merge is feature-driven. I will indeed presuppose that this is the case (see Svenonius (1994), Collins (2003), Adger (2003),

Stabler (1996, 1997, 1998), Kobele (2006), Sternefeld (2006), and Heck and Müller (2006b), among others, for arguments to this effect): External Merge is triggered by subcategorization features [•F•] on heads that act as probes, just as with feature-driven internal Merge.

Closer inspection reveals that a bit more must be said about the Merge status of categories that are dominated by a category that has been merged (α): On the one hand, it may be that a category γ included in a category α that undergoes feature-driven Merge has itself undergone non-feature-driven movement in α; in that case, γ clearly has a different Merge status. On the other hand, it may be that γ has undergone feature-driven movement in α (or no movement at all), whereas α is moved solely for reasons of Phase Balance; in this case, however, γ must have the same Merge status as α (its position is not fixed with respect to α-external material). The notion of Merge status can thus be defined as in (27) (where a *non-local feature* is a feature that is not part of the current tree).[23]

(27) *Merge status*:

a. A category γ in a position P has Merge status [–ψ] iff (i) or (ii):
   (i) γ is merged in P, and γ is required in P by a non-local Merge-inducing feature.
   (ii) γ is dominated by (a segment of) a category with Merge status [–ψ].

b. A category γ in a position P has Merge status [+ψ] iff (i) and (ii):
   (i) γ is merged in P, and γ is not required in P by a non-local Merge-inducing feature.
   (ii) γ is not dominated by a (segment of a) category with Merge status [–ψ].

(27) predicts a complementary distribution of Merge status [+ψ] and Merge status [–ψ] among the positions that are generated by Merge (see below on adjunction). However, there are two kinds of positions that (27) systematically classifies as [–ψ] even though they are base positions (and should thus arguably be [+ψ]). First, this holds for a Specv position in which an external argument

[23] A remark is due on the ontological status of [+ψ] and [–ψ]. These symbols do not act as genuine features that encode pieces of information (even though I will sometimes say that a category is 'marked' [+ψ] or [–ψ]); their only purpose is to simplify exposition. This should be particularly evident when categories change their [±ψ] status in the derivation, but it holds throughout: Whether a category has Merge status [+ψ] or [–ψ] can always be inspected by looking at the syntactic context; this is not information that exists independently on the category. Thus, there is no violation of the Inclusiveness Condition here (see Chomsky (2001b)).

is merged that will undergo raising to SpecT; and second, it holds for all heads that need to undergo movement to a higher head. This would imply that subjects (that will move) and heads (that will move) never participate in linearization statements in their base positions. To avoid this consequence, two additional assumptions are necessary. As for subjects, I will deviate from the standard assumption that the external argument's base position is a position that is accessible to the next higher phase; i.e., that this position is part of the edge of vP. More generally, suppose that phase edges must be *derived* positions, in the sense that they can only be reached by movement, as in (28).

(28) *Edge*:
A category is part of the edge of a phase iff it is a specifier of a phase head that is created by Move.

Given (28), the PIC requires movement of an external argument from an inner to an outer specifier of v, and this movement may often be string-vacuous. By adopting (28), Specv ceases to be exceptional with respect to edge properties (other escape hatches of phases – e.g., SpecT, SpecC – are reached by internal rather than external Merge).[24]

Turning next to heads, the situation is slightly different because heads cannot undergo successive-cyclic movement, assuming excorporation (at least of the type that would be required here) to be impossible (see Baker (1988), Roberts (1991) for relevant discussion). So it seems that if heads are to participate in the [+ψ]/[–ψ] system at all, they have to do so in situ. What we want is that a head may be [+ψ] (so that it can generate linearization statements with externally merged items) and also [–ψ] (if it is to undergo movement to a higher head). Assuming that heads can be [+ψ]/[–ψ] at the same time will not help because they would then never have a Merge status that is identical to that of either a [–ψ]-marked or a [+ψ]-marked element. The solution that I would like to propose in view of this is that heads always start out as [+ψ] and may then change their Merge status to [–ψ] under the conditions regulated by (27), *but only if they have discharged all their subcategorization features* (and thereby lost their 'base property'); see (29), which acts as an addition to (27-a-i).

(29) A head can only have Merge status [–ψ] if its subcategorization features have all been discharged; otherwise it has Merge status [+ψ].

It now remains to be shown how this revised approach to cyclic linearization

[24]Independent arguments for (28) are given in Heck and Müller (2006a).

can account for the shape conservation effects discussed in this paper without incurring the problems mentioned in section 3 above. I will do this in the next subsection.

5.2 Sample Derivations

5.2.1 Abstract Scenarios

Consider the abstract representation in (30), which results from a derivational step where an XP α is merged with another XP β, yielding ω (headed by α or β), with $\delta_1,...,\delta_n$ the categories dominated by α and $\gamma_1,...,\gamma_m$ the categories dominated by β.

(30) $[_\omega\ [_\alpha\ \delta_1,...,\delta_n\ ]\ [_\beta\ \gamma_1,...,\gamma_m\ ]]$

In (30), α and β may have an identical Merge status [+ψ]; they may in principle have an identical Merge status [–ψ] (although this state of affairs is independently excluded on general grounds, at least as long as interarboreal operations are ruled out); or they may differ in Merge status ([+ψ]/[–ψ]). In the first two cases, an ordering statement is generated for α and β; in the last case, no ordering statement is generated for α and β. Note, however, that categories δ_i and γ_j do not necessarily participate in the same ordering statements as α and β; δ_i may have a Merge status different from α, and γ_j may have a Merge status different from β (at least as long α, β do not have status [–ψ]). Thus, it may well be the case that, e.g., α and β have Merge status [+ψ], whereas some γ_j in β has Merge status [–ψ], which then prevents γ_j from participating in an ordering statement. (An instantiation of this abstract situation would be Merge (T,vP), with T and vP marked [+ψ], and vP including a *wh*-phrase in its (outer) specifier that has undergone non-feature-driven movement, and therefore has Merge status [–ψ].) Moreover, both δ_i and γ_j may in principle have Merge status [–ψ] even though their dominating categories α and β have Merge status [+ψ].[25] With this background, let me now turn to derivations of actual sentences.

5.2.2 Simple Clauses

Consider first simple English sentences like those in (31).

[25]Relevant constructions may well exist, in violation of Gazdar's (1981) one hole restriction on syntactic dependencies; see Pesetsky (1982) for relevant discussion. Furthermore, parasitic gaps would instantiate such a case, under some analyses.

(31) a. $[_{CP}$ C $[_{TP}$ Mary$_1$ $[_{vP}$ t$'_1$ $[_{v'}$ t$_1$ wrote$_2$-v $[_{VP}$ t$_2$ a book]]]]]
b. $[_{CP}$ C $[_{TP}$ John$_1$ $[_{vP}$ t$'_1$ $[_{v'}$ t$_1$ gave$_2$-v $[_{VP}$ $[_{DP_3}$ Mary] $[_{V'}$ t$_2$ $[_{DP_4}$ a book]]]]]]]

I have argued above that VP is a defective projection (and therefore not a phase). The conclusion that VP is special is reinforced by considerations related to cyclic linearization. If Merge of V and DP (in (31-ab)), or of V′ and DP (in (31-b)), generates ordering statements, subsequent obligatory movement of V to v will invariably generate contradictory ordering statements in the second case, which will then lead to a crash of the derivation. This outcome can be avoided if VP is defective with respect to linearization statements, just as it is defective with respect to phase status. Thus, I would like to conclude that no ordering statements are generated in the VP that involve V; the underlying rationale is that vP and VP act in certain respects like a single projection, and the position occupied by V after head movement to v is the one that determines base linearization. With this in mind, consider (31-b).

On the VP cycle, there are ordering statements $DP_3 < V'$ and, accordingly, $DP_3 < DP_4$ after Merge of DP_3, plus further statements triggered by (26); however, by stipulation, there is no statement $DP_3 < V$. DP_3, V′ and DP_4 have the same Merge status [+ψ] because they have undergone (feature-driven) external Merge. Next, v is merged with VP and attracts V; both operations are driven by probe features on v (for subcategorization and movement, respectively). This generates the ordering statement V+v < VP, and also V+v < DP_3, and V+v < DP_4 because all these categories have the same Merge status [+ψ]. In the next step, the external argument DP_1 is merged as v's specifier, which generates the ordering statement $DP_1 < v'$, and also $DP_1 < V+v$, $DP_1 < DP_3$, and $DP_1 < DP_4$. Thus, all external Merge operations affecting vP are completed, and v does not bear a probe feature anymore. Given (29), this means that v may in principle now acquire Merge status [–ψ]; however, it does not do so in the case at hand because English does not have V+v-to-T movement (i.e., T does not have a probe feature attracting V+v; hence, V+v is not required in its position by a non-local feature).

Next, the external argument DP_1 must be available for checking T's [•D•] feature. Given the assumption that phase edges can only be reached by movement (see (28)), an external argument that is required in the edge domain by Phase Balance (via T's [•D•] feature) must undergo non-feature-driven and (in the case at hand) string-vacuous movement to an outer specifier of vP (this is indicated by the intermediate traces t$'_1$ in (31)). In this case, no new ordering statement is generated because the moved external argument is [–ψ], and the

v′ category it is merged with is [+ψ] (furthermore, in the present case, there is no other category within v′ that is [–ψ], which would generate an ordering statement with the external argument). This completes the vP in (31); the domain of v then undergoes spell-out, and the ordering statements generated in this domain become inaccessible (and irrelevant) for the remainder of the derivation.

After this, T is merged with vP; both categories have Merge status [+ψ], and this generates some ordering statements, all of them straightforward.[26] However, at this point, no ordering statement is generated for T ([+ψ]) and the external argument DP_1 ([–ψ]). The next operation is subject raising to SpecT, driven by [•D•] on T. Now, an ordering statement is generated for DP_1 ([+ψ]) and T′ ([+ψ]), and for DP_1 and all other categories in T′ (because the latter are all marked [+ψ]). In principle, the new ordering statements thus generated could now contradict earlier ones (which are rendered inaccessible by spell-out of the earlier vP phase); but since the external argument preceded all vP-internal items in the first place, contradicting ordering statements do not arise (but see subsection 5.2.8 below on passive and raising constructions). After this, the domain of T is spelled out, and the next phase head (C) is merged with TP. Ordering statements for C ([+ψ]) and TP ([+ψ]) (as well as material within TP) are generated, but there are no interesting further consequences.

5.2.3 Wh-Movement

Consider next a simple case of *wh*-movement, as in (32):

(32) (I wonder) [$_{CP}$ what$_2$ C [$_{TP}$ t''_2 Mary$_1$ T [$_{vP}$ t'_1 [$_{v'}$ t'_2 [$_{v'}$ t_1 saw t_2]]]]]

On the vP level, the main difference to (31) is that both the external argument DP_1 and the internal argument DP_2 are required in the edge domain of vP by Phase Balance (for [•D•] on T and [•wh•] on C, respectively). External Merge of DP_1 ([+ψ]) and v′ ([+ψ]) inter alia generates a statement $DP_1 < DP_2$; subsequent non-feature-driven movement of DP_1 ([–ψ]) and DP_2 ([–ψ]) to outer specifiers of v′ generates a new ordering statement. If the two DPs reassemble in their base order, as in (32), no contradictory ordering statement is generated: First, movement of DP_2 across DP_1 (which is in its base position)

[26]Of course, ordering statements can be generated via (26) only for those categories that are still accessible in the structure (i.e., that have not yet undergone cyclic spell-out). To ensure a total order of terminals in the PF component, missing statements are added there according to the Nontangling Condition, which ensures transitivity; this process overwrites contradictory earlier statements, as stated above.

does not generate a new ordering statement because DP_2 is [–ψ] and DP_1 is [+ψ]; and second, subsequent movement of DP_1 across DP_2 reproduces the earlier ordering statement (both are [–ψ] now). However, if they reassemble in the edge domain in the reversed order, a new, contradictory ordering statement $DP_2 < DP_1$ is generated. This latter ordering statement would lead to a crash of the derivation because the earlier ordering statement $DP_1 < DP_2$ is still accessible at this point of the derivation. Of course, this theoretical prediction does not yet have empirical consequences: It is hard to see how the two derivations of (32) – with DP_1 vs. DP_2 as the highest specifier of vP – could be empirically distinguished. However, I will show below that there are contexts where different empirical predictions do arise. For the time being, we can conclude that the vP phase is thereby finished, and the complement of v is spelled out, together with the linearization statements established in the domain of vP.

On the TP level, DP_1 undergoes raising to SpecT (after Merge (T,vP), which can be neglected here) and, having acquired Merge status [+ψ], induces ordering statements with T′ ([+ψ]) and [+ψ]-marked categories included in T′; crucially, no statement is generated by Merge of DP_1 and T′ for DP_1 in SpecT ([+ψ]) and DP_2 in Specv ([–ψ]). Next, DP_2 moves to an outer specifier of T; this movement is not feature-driven but forced by Phase Balance. No new ordering statement is generated by this Merge operation because DP_2 in SpecT is [–ψ] and its sister T′ and all categories included in T′ have Merge status [+ψ]. After this, the domain of T is spelled out, and the only ordering statements that are still accessible in the derivation are $DP_1 < T'$ and $DP_1 < T$ (since T is phonologically empty, this can never be relevant).

On the CP cycle, DP_2 undergoes feature-driven movement to SpecC. This generates the ordering statement $DP_2 < DP_1$ (among others). Note that this contradicts the earlier ordering statement(s) $DP_1 < DP_2$ generated on the vP cycle. However, this is unproblematic because cyclic spell-out has long removed the conflicting linearization information from the derivation; and phonology, by assumption, resolves the conflict by simply replacing the earlier statement with the new one.

After these general illustrations, I will now return to the original idea that Holmberg's Generalization can be derived from cyclic linearization, and show how this follows in the present approach.

5.2.4 Simple Object Shift

Recall Sells' (2001) and Richards' (2004) insight that the shape conservation property of object shift is due to the strict locality of the operation. In the present approach, this translates into a generation of an ordering statement that contradicts an earlier ordering statement which is still present in the derivation. For concreteness, I assume that object shift is feature-driven movement to Specv.[27] This implies that negation and those adverbials which are crossed by object shift show up in specifier positions of v, and do not head functional projections that intervene between T and v in the main clausal projection line (see Bobaljik (2002), Thráinsson (2007) vs. Alexiadou (1997), Cinque (1999)).[28]

Consider now first licit cases of object shift accompanied by main verb raising, as in the Danish example (6-b), which is repeated here as (33) (with structural information added).

(33) [$_{CP}$ Hvorfor købte [$_{TP}$ Peter$_2$ t$'_{V+v+T}$ [$_{vP}$ t$'_2$ [$_{v'}$ [$_{DP_1}$ den]
why bought Peter it
[$_{v'}$ ikke [$_{v'}$ t$_2$ [$_{v'}$ t$_{V+v}$ [$_{VP}$ t$_V$ t$_1$]]]]]]] ?
not

On the vP cycle, an ordering statement V+v < DP_1 is generated after v has been merged with VP and V has undergone head movement to v. The external argument is merged, negation is added, and both operations trigger ordering statements with v′ and the categories included therein (because so far, all categories are marked [+ψ]). Next, the following two movement operations apply: There is non-feature-driven movement of the external argument DP_2 to an

[27] For present purposes, this feature can be referred to as [π]; and it can be further assumed that a [•π•] feature is inserted on v in the numeration (see Chomsky (2000, 2001b)) – obligatorily so for every [D] that is an unstressed pronoun argument of V in Danish and Icelandic, and optionally so for other [D] arguments in Icelandic. Under these assumptions, more must eventually be said to capture the fact that unstressed pronouns are possible in situ when verb movement is not possible, as in (6-c). However, I will not explore these matters here any further since they are independent of the main issues currently under consideration: The only important assumption is that an object-shifted DP has Merge status [+ψ] rather than Merge status [–ψ], and this should be uncontroversial.

[28] There are two implicit assumptions here: First, adverbials can only enter the structure if all subcategorization features of a head have been discharged. And second, internal Merge operations follow (i) external, feature-driven Merge operations and (ii) adverb insertion. If there is no movement to a position below a base position (but see the main text on movement to a position below a derived position), a moved object will invariably end up in a higher specifier than an adverbial.

outer specifier, forced by Phase Balance (for the [•D•] feature of T in the numeration), and there is feature-driven object shift of the internal argument DP_1 to an outer specifier as well. The Strict Cycle Condition essentially demands that every XP movement operation extends the tree created so far; however, head movement (see Chomsky (1993)) does not fall under this constraint. In addition, it has often been argued that movement to a non-unique (multiple) specifier may minimally violate the Strict Cycle Condition by ending up in lower specifier position ('tucking in'; see Richards (2001)). Assuming this latter option, the present system is compatible with two derivations of (33). In both of them, the external argument DP_2 must move first, acquiring status [$-\psi$] in the process, which precludes a generation of ordering statements with any category in its sister v'. If, alternatively, the internal argument DP_1 moves first to Specv, both DP_1 (in the object shift position, where it is required by a local probe feature on v, viz., the object shift-triggering feature [•π•]) and DP_2 (in situ) are marked [$+\psi$], and this movement operation then leads to a crash of the derivation because the new ordering statement $DP_1 < DP_2$ contradicts the earlier (and still accessible) statement $DP_2 < DP_1$. Thus, Phase Balance-driven movement of DP_2 to Specv has to apply first. The question then is which Specv position is targetted by subsequent object shift of the internal argument DP_1. Given present assumptions, DP_1 may either move to a Specv position above DP_2, or to a Specv position below DP_2, as an instance of tucking in (but above the adverb) – since DP_2 is now [$-\psi$], and DP_1 is [$+\psi$], no new statement is generated in either case. The structure in (33) represents the latter option, with tucking in of the shifted object, but I see no reason to rule out the former.[29]

When DP_1 undergoes object shift to Specv, it (potentially) crosses (or affects, see the last footnote) three categories containing lexical material – DP_2 (in one derivation), adverb, and V+v –, and for each of these categories, it must be shown that no contradictory ordering statement is generated. This is evident for DP_2: DP_1 ([$+\psi$]) and DP_2 ([$-\psi$]) differ in Merge status. What about the adverb *ikke*? It is often assumed that adverbial categories do not enter syntactic derivations as a result of feature-driven Merge operations since it does not seem plausible to assume that the discharge of a subcategorization probe feature (on either the modified category or the adverbial category itself) can be involved here; in fact, it is sometimes postulated that adverbial categories do

[29] Given the option of tucking in, relativized linearization (see (26)) must be modified in such a way that movement of a category α to a lower derived specifier may still trigger an ordering statement with a category β in a higher specifier of the same head. This is completely parallel to the modification of the Strict Cycle Conditon that tucking in necessitates.

not enter phrase markers by Merge in the first place, but by some alternative structure-building operation (Adjoin; cf. Chomsky (2000, 2001b), Adger (2003), among others). I will here follow this latter proposal and assume that adverbs are integrated into syntactic representations by a separate operation Adjoin. Adjoin is not feature-driven and follows all external Merge operations that are triggered by subcategorization features of a head X in an XP; but, crucially, Adjoin precedes all internal Merge operations (feature-driven or not). Since adverbial categories do not enter the structure by Merge, they have neither Merge status [+ψ] nor Merge status [–ψ] (as long as they stay in situ). This means that they cannot participate in ordering statements generated via (26) at all unless they undergo movement (feature-driven or non-feature-driven). Consequently, when the adverb *ikke* is merged, no linearization statement with DP_1 (or any other category) is generated; and the same goes for movement of DP_1 across the adverb.[30]

With DP_2 (*Peter*) and the intervening adverb (*ikke*) now accounted for, the final and most important question is why movement of DP_1 across V+v (*købte*) does not trigger a new ordering statement $DP_1 < V+v$, which would contradict the one established earlier, viz., $V+v < DP_1$. The answer is given by (29): When V+v has discharged all its subcategorization probe features (i.e., after external, feature-driven Merge of DP_2), it can in principle acquire Merge status [–ψ], thereby losing Merge status [+ψ]. In the case at hand, V+v does in fact now obtain Merge status [–ψ] because it is required in its position by a probe feature of T (which is still in the numeration at this point); this probe feature ([•v•]) will trigger head movement on the next cycle. Since V+v is marked [–ψ] immediately after the external argument is merged, object shift of DP_1 to an outer specifier of v gives rise to the following situation: DP_1 is marked [+ψ], V+v is marked [–ψ]. The Merge status of DP_1 and the Merge status of V+v are therefore not identical, and no new ordering statement is generated by the operation. More generally, then, object shift as in (33) never leads to contradictory ordering statements when the derivation "knows" that the finite verb will have to move later in the derivation (and will thereby invariably have to end up in front of the shifted object, given that T precedes its complement).[31]

[30]The question arises of how adverbial categories can then ultimately be linearized with respect to other categories. For present purposes, it may suffice to assume that ordering statements for an adverbial category (that is not part of the edge domain of a phase) are generated when a phase is complete, and the adverbial then undergoes spell-out together with the rest of the domain of the head of the phase.

[31]The same account suggests itself for cases where the verb ends up in front of an object-

Turning now to illicit cases of object shift that are not accompanied by verb movement to a higher position, the crucial difference between the vP of a sentence like (6-d), repeated here as (34), and the vP of a sentence like (6-b) (= (33)) is that v is non-finite in (34), and does not undergo movement to T on the next cycle. More technically, v is not required in its position by T because T does not have a [•v•] probe feature in this derivation. Consequently, v never acquires Merge status [–ψ] in (34). Therefore, with both items marked [+ψ], object shift of DP_1 creates an ordering statement $DP_1 < V+v$, which contradicts the earlier (and still syntactically accessible) statement $V+v < DP_1$. This leads to a crash of the derivation.

(34) *Hvorfor skal Peter$_2$ [$_{vP}$ t'_2 [$_{v'}$ [$_{DP_1}$ den] [$_{v'}$ ikke [$_{v'}$ t_2 [$_{v'}$
why shall Peter it not
[$_V$ købe]+v [$_{vP}$ t_V t_1]]]]]] ?
buy

In this analysis, the prohibition against object shift without verb movement can essentially be viewed as a kind of anti-locality effect (see also Richards (2004)). More generally, the present approach in terms of relativized cyclic linearization systematically derives a certain class of anti-locality effects (see Grohmann (2003a,b) and Abels (2003)): Whenever a complement β of a head α (or some γ that is more deeply embedded in β but has participated in an ordering statement with α which is still accessible, as in the case just discussed) undergoes feature-driven movement to the specifier of α, ungrammaticality arises because of two conflicting linearization statements. Thus, e.g., feature-driven local movement of TP to SpecC across a verb-second head in German (as discussed by Geilfuß (1988), Abels (2003)) will invariably result in ungrammaticality because of two conflicting ordering statements: C < TP before movement vs. TP < C after movement; see (35).[32]

(35) *[$_{CP}$ [$_{TP_1}$ Fritz gestern hier geschlafen t_2] [$_{C'}$ hat$_2$-C t_1]]
Fritz yesterday here slept has

shifted pronoun by topicalization rather than head movement; see Holmberg (1998) and Vikner and Engels (2006).

[32]Needless to say, though, both Grohmann's and Abels' approaches to anti-locality differ significantly from the present approach (and from each other) in their empirical consequences.

5.2.5 *Multiple Object Shift*

Multiple object shift, as in the Danish examples involving pronouns in (9), and the Icelandic examples involving non-pronominal DPs in (11), requires verb movement in exactly the same way, and for the same reasons, as simple object shift. In addition, multiple object shift is only possible if the base order among the two objects is maintained (indirect objects uniformly precede direct objects in the base in the Scandinavian languages); compare, e.g., the Icelandic examples in (11-a) and in (11-b), which are repeated here as (36-ab), with relevant structure added.

(36) a. Ég$_3$ lána [$_{vP}$ t$'_3$ [$_{v'}$ Maríu$_1$ [$_{v'}$ bækurnar$_2$
I lend Maria the.books
[$_{v'}$ ekki [$_{v'}$ t$_3$ [$_{v'}$ t$_{V+v}$ [$_{VP}$ t$_1$ t$_V$ t$_2$]]]]]]]
not

b. *Ég$_3$ lána [$_{vP}$ t$'_3$ [$_{v'}$ bækurnar$_2$ [$_{v'}$ Maríu$_1$
I lend the.books Maria
[$_{v'}$ ekki [$_{v'}$ t$_3$ [$_{v'}$ t$_{V+v}$ [$_{VP}$ t$_1$ t$_V$ t$_2$]]]]]]]
not

The shape conservation effect with multiple feature-driven object movement (for *bækurnar*$_2$ and *Maríu*$_1$ in (36-ab)) follows without further ado: When DP$_1$ is merged with V′ (which includes DP$_2$), an ordering statement DP$_1$ < DP$_2$ is generated (because both categories have Merge status [+ψ]). When DP$_1$ and DP$_2$ undergo object shift to outer specifiers of v, both categories have Merge status [+ψ] again. Since V+v acquires Merge status [–ψ] before object movement (after the external argument DP$_3$ has been merged), there is no linearization problem with respect to verb raising in (36-ab) because DP$_1$/DP$_2$ and V+v do not have the same Merge status (i.e., Holmberg's Generalization is respected). However, there is a linearization issue with DP$_1$ and DP$_2$: If the two objects reassemble in specifiers of v in an order that reverses the base order (as in (36-b)), an ordering statement DP$_2$ < DP$_1$ is generated that contradicts the earlier ordering statement DP$_1$ < DP$_2$, which leads to a crash of the derivation. Note that the indirect object DP$_1$ must move first in (36-a), followed by movement of the direct object DP$_2$ to a lower Specv position (tucking in; movement of the subject DP$_3$ may alternatively end up in a higher or lower Specv position, as before, since DP$_3$ has a different Merge status in a derived Specv position: [–ψ]). In contrast, (36-b) is ungrammatical independently of whether the DP$_1$ or DP$_2$ moves first.

This analysis also explains why a direct object pronoun cannot shift across an indirect object non-pronominal DP; see (10-a) vs. (10-b) in Danish, and (12-a) vs. (12-b) in Icelandic. The latter set of examples is repeated in (37).

(37) a. *Ég$_3$ lána [$_{vP}$ t$'_3$ [$_{v'}$ bækurnar$_2$ [$_{v'}$ ekki [$_{v'}$ t$_3$ [$_{v'}$ V+v
I lend the.books not
[$_{VP}$ Maríu$_1$ [$_{V'}$ tV t$_2$]]]]]]]
Maria

b. Ég$_3$ lána [$_{vP}$ t$'_3$ [$_{v'}$ Maríu$_1$ [$_{v'}$ ekki [$_{v'}$ t$_2$ [$_{v'}$ V+v [$_{VP}$ t$_1$ [$_{V'}$ tV
I lend Maria not
bækurnar$_2$]]]]]]]
the.books

In both (37-a) and (37-b), an ordering statement $DP_1 < DP_2$ is generated when DP_1 is first merged. This ordering statement is still visible when object shift applies. This is unproblematic with object shift of DP_1 in (37-b) because it generates a new statement $DP_1 < DP_2$; but ungrammaticality arises if DP_2 undergoes object shift alone, as in (37-a): A new statement $DP_2 < DP_1$ is generated, which contradicts the earlier statement $DP_1 < DP_2$ that is still accessible. From a more general point of view, movement of DP_2 in (37-a) is again too local. If DP_2 undergoes topicalization or *wh*-movement rather than object shift in the same context, no problem arises: DP_2 then has Merge status [–ψ] in Specv, and consequently does not generate a new ordering statement with respect to DP_1. Hence, the wellformedness of examples like (38) in Danish, where a direct object undergoes A-bar movement across an indirect object, is expected. The important fact here is that DP_2 does not generate an ordering statement with DP_1 when it is merged in the position of t'_2 ([–ψ] vs. [+ψ]). Note also that the order of subject DP_3 ([–ψ]) and direct object DP_2 ([–ψ]) is fixed in outer Specv positions here; however, it may be derived either by first moving the subject and then moving the direct object via tucking in, or by first moving the object, and then raising the subject to a higher specifier.[33]

[33] Subject DPs can stay in situ, within the vP, in transitive expletive constructions in Icelandic. Unless further assumptions are made, the present analysis predicts that object shift across the in situ-subject should be impossible. This conforms to observations made in Vikner (1995, pp. 198–200) and Bobaljik and Jonas (1996, pp. 212–214) for non-pronominal DPs (see, e.g., (i)), but is incompatible with the conflicting evidence put forward in Collins (1997, p. 18) and Alexiadou and Anagnostopoulou (2001, 2006). It is also incompatible with the evidence from *pronominal* object shift, which may regularly cross a subject; see Jónsson (1996).

(38) [$_{DP_2}$ Denne bog] viste [$_{TP}$ t''_2 [$_{T'}$ Peter$_3$ t_{V+v+T} [$_{vP}$ t'_3 [$_{v'}$ t'_2 [$_{v'}$ ikke
this book showed Peter not
[$_{v'}$ t_3 [$_{v'}$ t_{V+v} [$_{vP}$ Marie$_1$ [$_{V'}$ t_V t_2]]]]]]]]]
Marie

5.2.6 Pronoun Fronting in German

One might take the null hypothesis to be that German pronoun fronting is the same kind of operation as Scandinavian object shift. However, there are a number of differences between the two operations, and at least some of these differences shed doubt on such an assumption. A difference that can be accounted for independently is that an unstressed object pronoun may precede a subject DP in German (cf. (14-c) vs. (14-d)) but not in, say, Danish; this follows straightforwardly from the fact that external arguments are moved to the subject position SpecT optionally in German, and obligatorily in Danish.[34] Another difference that is not really surprising from the present perspective is that Scandinavian object shift depends on verb movement whereas movement of pronouns in German does not seem to do so in an obvious way (particularly if one assumes that there is no obligatory V-to-T movement in German; see Haider (1993) vs. Sabel (1996)); this simply follows from the assumption that Holmberg's Generalization instantiates a shape conservation effect which cannot show up in the same way in OV languages like German.

However, other asymmetries between Scandinavian object shift and German pronoun fronting are not as readily explained by invoking independent factors. In particular, whereas Scandinavian object shift shows A-movement properties, German pronoun fronting exhibits A-bar movement properties (see Vikner (1994)). Thus, consider the different behaviour with respect to parasitic gap licensing (a typical A-bar movement property) between Danish object shift in (39-a) and German pronoun fronting in (39-b).[35]

(i) *θaδ lauk verkefninu$_2$ einhver$_1$ t_2
there finished the.assignment someone

For present purposes, I will leave it at that, noting that if object shift across an in situ-subject in transitive expletive constructions turns out to be possible, this could be taken to support the hypothesis that there is overt subject movement in these constructions after all, with the expletive emerging as a partial realization of the moved subject DP, and the putatitive in situ-subject DP emerging as a fully spelled out trace. Needless to say, spell-out of traces would require a further modification of the system developed here, such that linearization is established for three rather than two items as a result of Merge operations.

[34]However, also cf. Josefsson (1992) on Swedish.

[35]Fanselow (2001, p. 412) argues that the construction in (39-b) does not actually involve a

(39) a. *Peter inviterede dem$_1$ ikke t$_1$ [$_{CP}$ uden at kende e$_1$ på forhånd]
Peter invited them not without to know beforehand

b. dass Peter sie$_1$ [$_{CP}$ ohne e$_1$ zu kennen] t$_1$ ins Regal
that Peter them without to know into.the bookshelf
gestellt hat
put has

Thus, I would like to conclude (deviating from the set of assumptions in Müller (2001)) that the differences between German pronoun fronting and Scandinavian object shift result from the fact that the two operations have different landing sites: Whereas object shift is feature-driven movement to the specifier of v, pronoun fronting targets a higher functional projection μP that intervenes between TP and vP; A-movement properties are associated with the former position, A-bar movement properties are associated with the latter. With this in mind, consider now first derivations of the sentences in (14-c) and (14-d), which are repeated here in (40-ab) (again, with additional structural information).[36]

(40) a. dass [$_{TP}$ [$_{\mu P}$ es$_1$ [$_{vP}$ t$'_1$ [$_{v'}$ gestern [$_{v'}$ [$_{DP_2}$ der Fritz] [$_{VP}$ t$_1$ t$_V$ [$_v$
that it yesterday the Fritz
[$_{V+v}$ gelesen]]]]]]]] hat
read has

b. dass [$_{TP}$ [$_{DP_2}$ der Fritz] [$_{\mu P}$ t$''_2$ [$_{\mu'}$ es$_1$ [$_{vP}$ t$'_2$ [$_{v'}$ t$'_1$ [$_{v'}$ gestern [$_{v'}$
that the Fritz it yesterday
t$_2$ [$_{VP}$ t$_1$ t$_V$ [$_v$ [$_{V+v}$ gelesen]]]]]]]]] hat
read has

The only difference between (40-a) and (40-b) is that T has an EPP feature [•D•] in the latter case, but not in the former (recall that this feature is optional in German). The relevant parts of the derivation of (40-a) proceed as

parasitic gap (because it may affect nonreferential and *wh*-pronouns, and may involve multiple gaps), and is hence irrelevant for the question of whether pronoun fronting is an A- or A-bar movement operation. Notwithstanding potential empirical problems with some of the relevant data, Fanselow's observations strike me as interesting and may well have interesting consequences for the analysis of, e.g., inherently reflexive verbs in German, but, in my view, they do not call into question a standard parasitic gap analysis.

[36]I assume that the auxiliary in (40) is the head of a vP-external verbal functional projection; but I know of no evidence to decide whether this projection is higher or lower than μP – or, for that matter, identical to it. (For instance, both unstressed pronouns and finite auxiliaries stay behind in predicate ('VP') topicalization contexts.)

follows: V obligatorily moves to v (which is right-peripheral in German), which produces an ordering statement $DP_1 < V+v$; DP_2 is merged in Specv, which generates the ordering statements $DP_2 < V+v$ and, more relevantly in the present context, $DP_2 < DP_1$. After Merge of DP_2, V+v can acquire Merge status [$-\psi$]. However, following essentially Haider (1993), I assume that there is no V-to-T movement (without subsequent T-to-C movement) in German, and so V+v continues to have only Merge status [$+\psi$] here. The next step is movement of DP_1 to a specifier of v. Given that pronoun fronting targets Specμ, it is clear that DP_1 does not undergo feature-driven movement to Specv in (40-a); rather, this movement step is driven by Phase Balance, via the non-local movement-inducing probe feature on μ in the numeration. Importantly, no new ordering statement $DP_1 < DP_2$ is generated because DP_1 has Merge status [$-\psi$], and DP_2 has Merge status [$+\psi$] (there is, by assumption, no [•D•] feature on T in the numeration). After the completion of vP, the domain of V+v (together with the adverb) is spelled out, and so are the ordering statements generated for this domain. Next, μ is merged with vP, and DP_1 undergoes feature-driven movement to Specμ. At this point, a new ordering statement $DP_1 < \mu'$ is generated (consequently also $DP_1 <$ vP, and therefore ultimately $DP_1 < DP_2$), but this is unproblematic: The earlier linearization statements, including $DP_2 < DP_1$, are not syntactically accessible anymore on the μP cycle.[37]

As noted, (40-b) differs from (40-a) in having subject raising to SpecT. Thus, both DP_1 and DP_2 undergo Phase Balance-driven movement to an outer specifier of v in (40-b); since they both have Merge status [$-\psi$] in the landing site, the two categories have to reassemble in the order dictated by the earlier, still visible, ordering statement, viz., $DP_2 < DP_1$. On the μP cycle, both DPs move again to specifier positions; however, no new ordering statement is generated because DP_1 has Merge status [$+\psi$], and DP_2 has Merge status [$-\psi$] (it is required in this position by a non-local feature, viz., [•D•] on T in the numeration). Finally, on the TP cycle, DP_2 undergoes feature-driven movement to SpecT, and a new ordering statement $DP_2 < DP_1$ is generated.

Let us now turn to the case of multiple pronoun fronting in German, as in (15); the examples are repeated here as (41-ab). In this case, there is a shape conservation effect: The subject pronoun must precede the object pronoun in situ and in the ultimate landing site; and, as noted, this cannot be due

[37]The assumption that pronoun fronting in German is not feature-driven movement to Specv proves crucial here: Otherwise, Merge (DP_1,v′) would produce a contradictory ordering statement while the original ordering statement is still accessible, and the derivation would be expected to crash.

to obligatory raising to subject position because such raising is optional in German.

(41) a. dass [$_{\mu P}$ sie$_1$ [$_{\mu'}$ es$_2$ [$_{vP}$ t'_1 [$_{v'}$ t'_2 [$_{V'}$ gestern [$_{v'}$ t_1 [$_{v'}$ [$_{VP}$
that she.NOM it.ACC yesterday
t_2 t_V] gelesen]]]]]]] hat
read has

b. *dass [$_{\mu P}$ es$_2$ [$_{\mu'}$ sie$_1$ [$_{vP}$ t'_1 [$_{v'}$ t'_2 [$_{V'}$ gestern [$_{v'}$ t_1 [$_{v'}$ [$_{VP}$
that it.ACC she.NOM yesterday
t_2 t_V] gelesen]]]]]]] hat
read has

On the vP cycle, when the external argument DP_1 is merged, it generates an ordering statement with the internal argument DP_2 (among others) because both categories have Merge status [+ψ]: $DP_1 < DP_2$. Next, DP_1 and DP_2 undergo non-feature-driven movement to outer specifiers of v, forced by Phase Balance via the features on μ that will later trigger pronoun fronting to this position. Since DP_1 and DP_2 are both marked [–ψ] after the first movement step, a new ordering statement is generated at a point of the derivation where the original ordering statement is still visible. If DP_1 and DP_2 reassemble in the pre-movement order, as they do in (41-ab), the new ordering statement is $DP_1 < DP_2$, and no problem arises; if, however, DP_1 is moved first, and DP_2 then ends up in front of it (or DP_2 is moved first, and DP_1 is then moved to a lower position, via tucking in), an ordering statement $DP_2 < DP_1$ is generated that leads to a crash of the derivation. Exactly the same reasoning applies on the μP cycle, the only difference being that DP_1 and DP_2 are now both marked [+ψ] again. The (second) ordering statement $DP_1 < DP_2$ generated on the vP cycle is still accessible; and therefore this ordering statement demands that the ultimate order of DP_1 and DP_2 is as in (41-a), and not as in (41-b) (the last step may involve tucking in, but does not have to).

There is an independent piece of evidence for the difference between object shift and pronoun fronting with respect to the landing site. Recall that object shift cannot move an object across another, c-commanding object – if it does, a new ordering statement will be generated that contradicts the original one, which is still accessible. If German pronoun fronting does not target Specv (but a higher specifier), the prediction is that an intervening VP-internal object can be crossed in the process; the reason is that an object pronoun that undergoes non-feature-driven movement to Specv ([–ψ]) does not enter into a new ordering relation with another object that remains in the VP ([+ψ]). This

prediction is borne out: Both a direct object pronoun and an indirect object pronoun can undergo pronoun fronting, with the remaining object staying in situ; see (42-ab). In my view, there is good reason to assume that the base order of arguments in German is neither dative before accusative (as it is standardly assumed, and as it seems correct for the Scandinavian languages), nor variable and dependent on the argument structure of individual verbs (as argued by Haider (2000)), but uniformly (i.e., with all types of verbs) accusative before dative; see Müller (1993, 2001) for arguments to this effect. If so, the argument that is crossed by pronoun fronting in (42) is the direct object DP_1 (*das Buch* ('the book')) in (42-b).

(42) a. dass es_1 $[_{vP}$ t'_1 $[_{v'}$ gestern $[_{v'}$ die Maria $[_{VP}$ t_1 $[_{V'}$ $[_{DP_2}$
that it.ACC yesterday the Maria.NOM
dem Karl] t_V]] gegeben]]] hat
the Karl.DAT given has

b. dass ihm_2 $[_{vP}$ t'_2 $[_{v'}$ gestern $[_{v'}$ die Maria $[_{VP}$ $[_{DP_1}$ das
that him.DAT yesterday the Maria.NOM the
Buch] $[_{V'}$ t_2 gegeben]]]]] hat]
book.ACC given has

Finally, given a uniform base order accusative before dative, we expect that if both objects are unstressed pronouns, they have to show up in exactly this order after pronoun fronting. Again, the prediction is confirmed: The order of direct and indirect pronouns after multiple fronting must be accusative before dative; see (43-a) vs. (43-b).[38]

(43) a. dass $[_{\mu P}$ es_1 $[_{\mu'}$ ihm_2 $[_{vP}$ t'_1 $[_{v'}$ t'_2 $[_{v'}$ der Fritz $[_{VP}$ t_1 $[_{V'}$ t_V t_2]]
that it him ART Fritz
gegeben]]]]] hat
given has

b. *dass $[_{\mu P}$ ihm_2 $[_{\mu'}$ es_1 $[_{vP}$ t'_1 $[_{v'}$ t'_2 $[_{v'}$ der Fritz $[_{VP}$ t_1 $[_{V'}$ t_V t_2]]
that him it ART Fritz
gegeben]]]]] hat
given has

[38]There are a few complications. One concerns the order of pronouns in coherent infinitive constructions, where the expected linearization shows up but might initially be unexpected in some cases (assuming that two clauses are involved); see Müller (2001). Other complications involve deviations from the expected order in certain contexts, which may show the influence of other factors on the order of unstressed pronouns. See Zifonun et al. (1997) (and literature cited there) and Anagnostopoulou (2005).

The analysis is almost exactly as in (41), where one pronoun is a subject and one an object: After Merge of the accusative object DP_1, a linearization statement is generated that includes the dative object DP_2: $DP_1 < DP_2$. Since both pronouns first undergo non-feature-driven movement to an outer specifier of v, and consequently they both acquire Merge status [–ψ] as a result of the movement operations, a new ordering statement for DP_1 and DP_2 is generated. If the order in the edge domain of v is $DP_2 < DP_1$, the derivation crashes at this point. In (43-ab), however, the original order is maintained, and a new ordering statement $DP_1 < DP_2$ is generated that is compatible with the older one. The same reasoning applies on the next, and final, cycle, viz., μP. Now both pronouns have Merge status [+ψ], and a new ordering statement is generated while the ordering statement from the edge domain of vP is still accessible. In (43-b), a contradictory statement is generated, with ungrammaticality resulting; in (43-a), the previous order is maintained, and a non-contradictory ordering statement is generated.

Taking together the analyis of examples with two object pronouns, and of examples with a subject and a (direct) object, the prediction is that there is an obligatory order subject before direct object before indirect object in examples with three unstressed pronouns. This is borne out.

5.2.7 Multiple Wh-Movement in Bulgarian

The account of the shape conservation effect with multiple *wh*-movement in Bulgarian that can be given in the present approach is completely parallel to the account of multiple pronoun fronting in German.[39] The relevant examples are repeated here from (16).

(44) a. $[_{CP}$ Koj$_1$ $[_{C'}$ kogo$_2$ $[_{C'}$ C $[_{TP}$ t''_1 $[_{T'}$ t''_2 $[_{vP}$ t'_1 $[_{v'}$ t'_2
who.NOM whom.ACC
$[_{v'}$ t_1 vižda $[_{VP}$ t_V t_2]]]]]]]]] ?
sees

b. *$[_{CP}$ Kogo$_2$ $[_{C'}$ koj$_1$ $[_{C'}$ C $[_{TP}$ t''_1 $[_{T'}$ t''_2 $[_{vP}$ t'_1 $[_{v'}$ t'_2
whom.NOM who.ACC
$[_{v'}$ t_1 vižda $[_{VP}$ t_V t_2]]]]]]]]] ?
sees

As before, when the external argument DP_1 is merged with v′, an ordering statement $DP_1 < DP_2$ is generated (because at this point, both categories

[39] Recall, however, that things are less clear with two *wh*-objects.

have Merge status [+ψ]). Next, Phase Balance forces both DPs to undergo non-feature-driven movement to an outer specifier of v, and since the DPs both have Merge status [–ψ] in this position (and therefore give rise to a new ordering statement), they must show up in the pre-movement order in the edge domain of vP. On the TP cycle, the same reasoning applies, and DP_1 and DP_2 must again show up in the original order (DP_1 is [–ψ] in SpecT even if T has a [•D•] feature, given that C – which is still in the numeration – has two [•wh•] features, and DP_1 is thus required in SpecT by a non-local Merge-inducing feature). Finally, on the CP cycle, both categories have reached their ultimate landing site; they have Merge status [+ψ], which again generates an ordering statement, which must then not contradict the earlier ordering statement on the TP cycle which is still accessible. In principle, this mechanism is unbounded: Whenever two categories end up in the same edge domain as a result of non-feature-driven movement, they must respect an original order determined by external Merge, and they will have to preserve that order in all subsequent cycles (with the derivation proceeding via tucking in) until either both reach their target position on the same cycle (then the original order must be maintained here as well), or one reaches its ultimate landing site and the other one moves on (then the original order can be reversed because no new statement is generated, and the earlier statements are now inaccessible in the derivation).

5.2.8 A-Movement in Passive and Raising Constructions

Recall that A-movement constructions like those in (17) pose a problem for Fox and Pesetsky's approach if unaccusative and passive vPs have phase (spell-out domain) status; see (45-ab).

(45) a. $[_{CP}$ C $[_{TP}$ John$_1$ was $[_{vP}$ t'_1 $[_{v'}$ hit-v $[$ t t_1 $]]]]]$
b. $[_{CP}$ C $[_{TP}$ Mary$_1$ T $[_{vP}$ seems-v $[_{TP}$ t_1 to be smart $]]]]$

There is no problem in deriving these examples in the present approach. When V is moved to v, an ordering statement V+v < DP_1 is generated. Next, DP_1 moves to Specv because of Phase Balance. Since this movement is not feature-driven, DP_1 now has Merge status [–ψ]. V+v does not have to move on the next cycle, so it continues to have only Merge status [+ψ]; hence, no new ordering statement is generated. After this, T ([+ψ]) (= *was*) is merged with vP. This generates ordering statements (with vP and V+v), but not with DP_1 ([–ψ]). Finally, on the TP cycle DP_1 undergoes feature-driven movement to SpecT. This generates an ordering statement DP_1 < T′, which does not

contradict any earlier ordering statement that would still be visible.

5.2.9 Verb-Second

The final two constructions to be addressed here involve verb-second in SVO languages and verb-second in SOV languages. Both constructions raise problems for the approach in Fox and Pesetsky (2005a) but turn out to be fully compatible with the present, relativized approach to cyclic linearization.

Consider first verb-second in an SVO language like Danish, as exemplified by the sentences in (13-b) (= (33)) and (38); the former is repeated here as (46). Recall that the main problem with verb-second in SVO languages in Fox and Pesetksky's approach is that an order of subject before verb can be changed into an order of verb before subject.[40]

(46) [$_{CP}$ Hvorfor [$_{C'}$ [$_{V+v+T+C}$ købte] [$_{TP}$ Peter$_2$ [$_{T'}$ t$_{V+v+T}$ [$_{vP}$ t$'_2$ [$_{v'}$ den$_1$ [$_{v'}$
why bought Peter it
ikke [$_{v'}$ t$_2$ [$_{v'}$ t$_{V+v}$ [$_{VP}$ t$_V$ t$_1$]]]]]]]]] ?
not

We can abstract away from all ordering statements involving the object pronoun because this issue has already been discussed in section 5.2.4. Focussing on the verb and the external argument DP$_2$, an ordering statement DP$_1$ < V+v is generated when the external argument is merged with v′. Next (after its subcategorization probe features have been discharged), V+v acquires Merge status [–ψ], and DP$_2$ undergoes non-feature-driven movement to an outer specifier because of Phase Balance, thereby acquiring Merge status [–ψ]. A second, identical ordering statement for DP$_2$ and V+v is thus generated on the vP cycle. On the TP cycle, note first that no ordering statement is generated for T ([+ψ]) and DP$_2$ ([–ψ]). Next, V+v moves to T; V+v+T acquires Merge status [–ψ] since T has used up its subcategorization probe features but is needed by a probe feature triggering verb-second movement that is located on C in the numeration. After this, DP$_2$ moves to SpecT, obtaining Merge status [+ψ]. Consequently, no new ordering statement is generated for subject and verb. Next, the domain of T is spelled out, which removes all existing ordering statements for subject and verb from the syntactic derivation. Finally, T moves to C. This generates a new ordering statement V+v+T+C < DP$_1$ (since both categories are now marked [+ψ]), which contradicts the earlier statement DP$_1$

[40](33) has been analyzed in the present approach already, but only with respect to the interaction of the two DPs; now the focus is on the relation of the verb and the subject DP.

< V+v. However, by the time the new statement is generated by the derivation, the earlier, contradictory statement has been removed from it.

Consider finally verb-second in SOV languages. This construction has resisted a simple analysis in Fox and Pesetsky's approach because the order of an argument (in particular, an object) and the verb may be systematically reversed from the lower spell-out domain to the higher spell-out domain; recall the examples in (20), the first of which is repeated here with additional structural information.

(47) $[_{CP}\ [_{DP_1}$ Maria] $[_{C'}\ [_{V+v+T+C}$ las $[_{TP}\ t''_1\ [_{T'}\ [_{vP}\ t'_1\ [_{v'}\ t_1\ [_{v'}$
Maria.NOM read
$[_{VP}\ [_{DP_2}$ das Buch] $t_V\ ]\ t_{V+v}\ ]]]\ t_{V+v+T}\ ]]]]]$
the book.ACC

When V is moved to v, an ordering statement $DP_2 < V+v$ is generated. Next, DP_1 is merged with v', which results in (among others) the ordering statements $DP_1 < V+v$ and $DP_1 < DP_2$. Then DP_1 undergoes string-vacuous, non-feature-driven movement to an outer specifier of v because Phase Balance requires this for the [•top•] feature on C in the numeration; this feature acts as the trigger for topicalization.[41] Since V+v is [–ψ] by now (v has discharged its arguments and is required in its position by a non-local feature on C in the numeration), a new statement $DP_1 < V+v$ is generated that corresponds to the one established earlier. After this, the verb and the subject argument move up hand in hand in the structure, generating identical ordering statements with respect to each other, until T ends up in C, and DP_1 undergoes feature-driven movement to SpecC. At this point, both V+v+T+C and DP_1 are [+ψ], and a final compatible ordering statement is generated ($DP_1 < T$). As for the relation of the verb and the object DP_2, the initial linearization statement $DP_2 < V+v$ is not followed by any other statement involving these two categories after the vP cycle is completed. When V+v+T moves to C, it generates a new ordering statement with TP – hence ultimately, at PF, with the object DP_2, via the Nontangling Condition. However, DP_2 has already undergone spell-out as part of the domain of v after vP is completed. Thus, whereas the verb follows the object at the beginning of the derivation, it precedes the object at the end of the derivation. This is entirely unproblematic because thanks to the fact that phases are small units in the present approach, the window in which

[41] We can assume that DP_1 has a matching [top] feature; alternatively, it might be that C's relevant feature is a bare EPP feature in the case at hand, and minimality requirements then force movement of the highest argument. See Fanselow and Mahajan (2000), Fanselow (2003), and Fanselow and Lenertová (2007) for relevant discussion.

ordering statements are accessible is fairly small throughout the derivation; and by the time the CP cycle is reached, all ordering statements of the vP cycle which have not been constantly renewed throughout the derivation by multiple non-feature-driven movement to edge domains are irrelevant for syntactic operations.[42]

5.2.10 Scrambling

The question arises of how order-changing scrambling fits into the approach to relativized cyclic linearization outlined here. Given that some notion of shape conservation plays a role for syntactic movement operations, order-changing scrambling can be viewed as initially unexpected; and it does indeed turn out that the present approach is incompatible with many theories of scrambling.[43] It is incompatible with the assumption that scrambling is local, feature-driven movement to Specv; as has been amply illustrated above, local movement to Specv obeys shape conservation (an object can never be crossed by another object, and subject DPs can stay in situ, within vP, in German, in which case they block feature-driven movement of an object to Specv). The present approach is also incompatible with the hypothesis that scrambling in German and other languages is base-generated, i.e., that there is no fixed base order (see, e.g., Haider (1988), Fanselow (2001)); the reason is that the shape conservation effects with pronoun fronting can only be derived by cyclic linearization if there is a fixed base order.

What, then, could be possible approaches to scrambling that are compatible with relativized cyclic linearization? Perhaps the most straightforward possibility would be to assume that scrambling targets unique specifiers of functional categories (and not multiple specifiers); see, e.g., Meinunger (1995). These functional categories must be located above vP, but below μP (because all scrambled material follows all unstressed pronouns). Another approach might try to assign scrambled items a special Merge status that would systematically prevent the generation of ordering statements by Merge (and that might treat scrambled items on a par with adverbial categories). Third, it might be a constitutive property of scrambling that this movement operation deletes existing ordering statements. And fourth, a more radical approach

[42] As far as object-initial verb-second clauses are concerned (see (20-c)), the derivation is straightforward in SVO and SOV languages: When an object DP moves to Specv because of Phase Balance, an ordering statement DP > V+v is established because both items are [$-\psi$] at this point, and this order is then maintained throughout the derivation, including the final landing sites, where both items are [$+\psi$].

[43] Incidentally, this does not hold for Fox and Pesetsky's (2005a) approach.

might directly take into account that it has sometimes been argued that a reversal of the pre-movement order is in fact the very reason for the existence of scrambling (see, e.g., Ross (1967), Haider (1993)). On this view, a scrambling operation might apply *in order to yield contradictory ordering statements*; a technical implementation of this idea might then involve concepts like constraint ranking and constraint violability. I will have to leave these questions open.

6 Conclusion

To sum up, the present system of relativized cyclic linearization predicts two kinds of shape conservation effects. First, there are *anti-locality effects*: Two conflicting ordering statements arise for a head and an XP within the same phase; this effect is instantiated by simple object shift that is not accompanied by verb movement out of the vP. Second, there are *multiple movement effects*: These involve two items that originate in the same domain, undergo successive-cylic, non-feature-driven movement together, and finally end up in the same domain again, as is the case with multiple object shift, multiple pronoun fronting, and multiple *wh*-movement. The analysis relies on a derivational organization of grammar where (i) information is lost as the derivation proceeds; (ii) the role of representations is minimized (ordering statements are generated for structure-building operations, not for representations); and (iii) intermediate movement steps can be formally distinguished from final movement steps.

References

Abels, Klaus
2003 *Successive Cyclicity, Anti-Locality, and Adposition Stranding*. Ph.D. thesis, University of Connecticut, Storrs, Connecticut.

Adger, David
2003 *Core Syntax*. Oxford University Press, Oxford, New York.

Agbayani, Brian
1998 *Feature Attraction and Category Movement*. Ph.D. thesis, UC Irvine.

Alexiadou, Artemis
1997 *Adverb Placement*. Benjamins, Amsterdam.

Alexiadou, Artemis and Elena Anagnostopoulou
2001 The subject in situ generalization, and the role of Case in driving computations. *Linguistic Inquiry*, 32: 193–231.
2006 The subject in situ generalization revisited. Ms., Universität Stuttgart and University of Crete. This volume.

Anagnostopoulou, Elena
2005 Judgements on pronoun restrictions in German. Ms., University of Crete.

Baker, Mark
1988 *Incorporation. A Theory of Grammatical Function Changing*. University of Chicago Press, Chicago.

Bierwisch, Manfred
1963 *Grammatik des deutschen Verbs*. Studia Grammatica II. Akademie Verlag, Berlin.

Billings, Loren and Catherine Rudin
1996 Optimality and superiority: A new approach to overt multiple wh ordering. In Jindřich Toman, (ed.), *Proceedings of Formal Approaches to Slavic Linguistics. The College Park Meeting 1994*, pp. 35–60. Ann Arbor.

Biskup, Petr
2005 Phase featuring. Ms., Universität Leipzig.

Bobaljik, Jonathan
2002 A-chains at the PF interface. *Natural Language and Linguistic Theory*, 20: 197–267.
2005 Re: CycLin and the role of PF in object shift. *Theoretical Linguistics*, 31: 111–125.

Bobaljik, Jonathan and Dianne Jonas
1996 Subject positions and the roles of tp. *Linguistic Inquiry*, 27: 195–236.

Boeckx, Cedric
2003 *Islands and Chains. Resumption as Stranding*. Benjamins, Amsterdam.

Bošković, Željko
2002a A-movement and the EPP. *Syntax*, 5: 167–218.

2002b On multiple wh-fronting. *Linguistic Inquiry*, 33: 351–383.

Chomsky, Noam

1973 Conditions on transformations. In Stephen Anderson and Paul Kiparsky, (eds.), *A Festschrift for Morris Halle*, pp. 232–286. Academic Press, New York.

1981 *Lectures on Government and Binding*. Foris, Dordrecht.

1986 *Barriers*. MIT Press, Cambridge, Mass.

1993 A minimalist program for syntactic theory. In Ken Hale and Samuel Jay Keyser, (eds.), *The View from Building 20*, pp. 1–52. MIT Press, Cambridge, Mass.

1995 *The Minimalist Program*. MIT Press, Cambridge, Mass.

2000 Minimalist inquiries: The framework. In Roger Martin, David Michaels, and Juan Uriagereka, (eds.), *Step by Step*, pp. 89–155. MIT Press, Cambridge, Mass.

2001a Beyond explanatory adequacy. Ms., MIT, Cambridge, Mass.

2001b Derivation by phase. In Michael Kenstowicz, (ed.), *Ken Hale. A Life in Language*, pp. 1–52. MIT Press, Cambridge, Mass.

2005a On phases. Ms., MIT, Cambridge, Mass.

2005b Three factors in language design. *Linguistic Inquiry*, 36: 1–22.

Cinque, Guglielmo

1999 *Adverbs and Functional Heads*. Oxford University Press, Oxford.

Collins, Chris

1997 *Local Economy*. MIT Press, Cambridge, Mass.

2003 Eliminating labels. In Samuel David Epstein and T. Daniel Seely, (eds.), *Derivation and Explanation in the Minimalist Program*, pp. 42–64. Blackwell, Oxford.

Collins, Chris and Höskuldur Thráinsson

1996 VP-internal structure and object shift in Icelandic. *Linguistic Inquiry*, 27: 391–447.

Epstein, Samuel David and T. Daniel Seely

2002 Rule applications as cycles in a level-free syntax. In Samuel David Epstein and T. Daniel Seely, (eds.), *Derivation and Explanation in the Minimalist Program*, pp. 65–89. Blackwell, Oxford.

Fanselow, Gisbert

1991 Minimale syntax. Habilitation thesis, Universität Passau.

2001 Features, theta-roles, and free constituent order. *Linguistic Inquiry*, 32: 405–436.

2003 Münchhausen-style head movement and the analysis of verb-second. In Anoop Mahajan, (ed.), *Head Movement and Syntactic Theory*, volume 3 of *Syntax at Sunset*, pp. 40–76. UCLA & Universität Potsdam Working Papers in Linguistics, Los Angeles & Potsdam.

Fanselow, Gisbert and Denisa Lenertová
2007 Movement to Spec,CP and linearization. Ms., Universität Potsdam & Universität Leipzig.
Fanselow, Gisbert and Anoop Mahajan
2000 Towards a minimalist theory of wh-expletives, wh-copying, and successive cyclicity. In Uli Lutz, Gereon Müller, and Arnim von Stechow, (eds.), *Wh-Scope Marking*, pp. 195–230. Benjamins, Amsterdam.
Fischer, Silke
2004 *Towards an Optimal Theory of Reflexivization.* Ph.D. thesis, Universität Tübingen.
Fox, Danny and David Pesetsky
2003 Cyclic linearization and the typology of movement. Ms., MIT (long handout).
2005a Cyclic linearization of syntactic structure. *Theoretical Linguistics*, 31: 1–45.
2005b Cyclic linearization and its interaction with other aspects of grammar: A reply. *Theoretical Linguistics*, 31: 235–262.
Frampton, John and Sam Gutman
1999 Cyclic computation. *Syntax*, 2: 1–27.
Gärtner, Hans-Martin
2002 *Generalized Transformations and Beyond.* Akademie-Verlag, Berlin.
Gazdar, Gerald
1981 Unbounded dependencies and coordinate structure. *Linguistic Inquiry*, 12: 155–184.
Gazdar, Gerald, Ewan Klein, Geoffrey Pullum, and Ivan Sag
1985 *Generalized Phrase Structure Grammar.* Blackwell, Oxford.
Geilfuß, Jochen
1988 *Reanalyse?* Master's thesis, Universität Tübingen.
Grewendorf, Günther
2001 Multiple wh-fronting. *Linguistic Inquiry*, 32: 87–122.
Grohmann, Kleanthes
2003a *Prolific Domains: On the Anti-Locality of Movement Dependencies.* John Benjamins, Amsterdam & Philadelphia.
2003b Successive cyclicity under (anti-)local considerations. *Syntax*, 6: 260–312.
Haider, Hubert
1988 Theta-tracking systems – evidence from German. In Laszlo Marácz and Pieter Muysken, (eds.), *Configurationality*, pp. 185–206. Foris, Dordrecht.
1993 *Deutsche Syntax – generativ.* Narr, Tübingen.
2000 Branching and discharge. In Peter Coopmans, Martin Everaert, and Jane Grimshaw, (eds.), *Lexical Specification and Insertion*, pp. 135–164. Benjamins, Amsterdam.

Hallman, Peter
2000 Verb-final as a subcase of verb-second. In Masako Hirotani, Andries Coetzee, Nancy Hall, and Ji-Yung Kim, (eds.), *Proceedings of NELS 30*, pp. 287–298. GLSA, Amherst, Mass.

Heck, Fabian
2004 *A Theory of Pied Piping*. Ph.D. thesis, Universität Tübingen.
2005 Remarks on Fox & Pesetsky (2003). Lecture Notes, Universität Leipzig.

Heck, Fabian and Gereon Müller
2000 Successive cyclicity, long-distance superiority, and local optimization. In Roger Billerey and Brook D. Lillehaugen, (eds.), *Proceedings of WCCFL*, volume 19, pp. 218–231. Cascadilla Press, Somerville, MA.
2006a Derivational optimization of wh-movement. Ms., Universität Leipzig. To appear in special issue of *Linguistic Analysis*: Dynamic Interfaces; ed. Kleanthes Grohmann.
2006b Extremely local optimization. Ms., Universität Leipzig.

Heidolph, Karl Erich, Walter Flämig, and Wolfgang Motsch
1981 *Grundzüge einer deutschen Grammatik*. Akademie-Verlag, Berlin.

Holmberg, Anders
1986 *Word Order and Syntactic Features in the Scandinavian Languages and English*. Ph.D. thesis, University of Stockholm.
1998 Remarks on Holmberg's Generalization. *Studia Linguistica*, 53: 1–39.

Ishihara, Shinichiro
2004 Prosody by phase. Ms., Universität Potsdam.

Johnson, Kyle
1991 Object positions. *Natural Language and Linguistic Theory*, 9: 577–636.

Jónsson, Jóhannes Gísli
1996 *Clausal Architecture and Case in Icelandic*. Ph.D. thesis, University of Massachusetts, Amherst.

Josefsson, Gunlög
1992 Object shift and weak pronominals in Swedish. *Working Papers in Scandinavian Syntax*, 49: 59–94.

Kayne, Richard
1994 *The Antisymmetry of Syntax*. MIT Press, Cambridge, Mass.

Ko, Heejeong
2004 Asymmetries in scrambling and cyclic linearization. Ms., MIT, Cambridge, Mass.

Kobele, Greg
2006 *Generating Copies*. Ph.D. thesis, UCLA, Los Angeles.

Koster, Jan
1978 *Locality Principles in Syntax*. Foris, Dordrecht.
2000 Variable-free grammar. Ms., University of Groningen.

Kratzer, Angelika and Lisa Selkirk
2007 Phase theory and prosodic spellout: The case of verbs. Ms., University of Massachusetts, Amherst.
Legate, Julie Anne
2003 Some interface properties of the phase. *Linguistic Inquiry*, 34: 506–516.
Lenerz, Jürgen
1977 *Zur Abfolge nominaler Satzglieder im Deutschen*. Narr, Tübingen.
McCloskey, James
2002 Resumptives, successive cyclicity, and the locality of operations. In Samuel David Epstein and T. Daniel Seely, (eds.), *Derivation and Explanation in the Minimalist Program*, pp. 184–226. Blackwell, Oxford.
Meinunger, André
1995 *Discourse Dependent DP (De-)Placement*. Ph.D. thesis, Universität Potsdam.
Müller, Gereon
1993 *On Deriving Movement Type Asymmetries*. Ph.D. thesis, Universität Tübingen.
2000 Shape conservation and remnant movement. In Masako Hirotani, Andries Coetzee, Nancy Hall, and Ji-Yung Kim, (eds.), *Proceedings of NELS 30*, pp. 525–539. GLSA, Amherst, Mass. Revised and extended version appeared 2002 in Artemis Alexiadou, Elena Anagnostopoulou, Sjef Barbiers, & Hans-Martin Gärtner (eds.), *Dimensions of Movement*, pp. 209–241. Amsterdam: Benjamins.
2001 Order preservation, parallel movement, and the emergence of the unmarked. In Géraldine Legendre, Jane Grimshaw, and Sten Vikner, (eds.), *Optimality-Theoretic Syntax*, pp. 279–313. MIT Press, Cambridge, Mass.
2004 Phrase impenetrability and wh-intervention. In Arthur Stepanov, Gisbert Fanselow, and Ralf Vogel, (eds.), *Minimality Effects in Syntax*, pp. 289–325. Mouton de Gruyter, Berlin.
2005 Remarks on cyclic linearization and order preservation. *Theoretical Linguistics*, 31: 159–171.
Müller, Gereon and Wolfgang Sternefeld
1993 Improper movement and unambiguous binding. *Linguistic Inquiry*, 24: 461–507.
Nilsen, Øystein
2005 Some notes on cyclic linearization. *Theoretical Linguistics*, 31: 173–183.
Partee, Barbara, Alice ter Meulen, and Robert Wall
1993 *Mathematical Methods in Linguistics*. Kluwer, Dordrecht.
Pesetsky, David
1982 *Paths and Categories*. Ph.D. thesis, MIT, Cambridge, Mass.

Richards, Marc
2004 *Object Shift and Scrambling in North and West Germanic: A Case Study in Symmetrical Syntax.* Ph.D. thesis, University of Cambridge, Cambridge, UK.

Richards, Norvin
2001 *Movement in Language.* Oxford University Press, Oxford.

Rizzi, Luigi
1990 *Relativized Minimality.* MIT Press, Cambridge, Mass.

Roberts, Ian
1991 Excorporation and minimality. *Linguistic Inquiry*, 22: 209–218.

Roberts, Ian and Anna Roussou
2002 The extended projection principle as a condition for the tense-dependency. In Peter Svenonius, (ed.), *Subjects, Expletives, and the EPP.* Benjamins, Amsterdam.

Ross, John
1967 *Constraints on Variables in Syntax.* Ph.D. thesis, MIT, Cambridge, Mass.

Rudin, Catherine
1988 On multiple questions and multiple wh fronting. *Natural Language and Linguistic Theory*, 6: 445–501.

Sabel, Joachim
1996 *Restrukturierung und Lokalität.* Akademie-Verlag, Berlin.

Sells, Peter
2001 *Structure, Alignment and Optimality in Swedish.* Stanford Monographs in Linguistics. CSLI Publications, Palo Alto.
2005 Comments on object shift and cyclic linearization. *Theoretical Linguistics*, 31: 185–198.

Sportiche, Dominique
1989 Le mouvement syntaxique: Contraintes et paramètres. *Langages*, pp. 35–80.

Stabler, Edward
1996 Acquiring and parsing languages with movement. Book ms., UCLA.
1997 Derivational minimalism. In Christian Retoré, (ed.), *Logical Aspects of Computational Linguistics*, pp. 68–95. Springer, Heidelberg.
1998 Acquiring languages with movement. *Syntax*, 1: 72–97.

Sternefeld, Wolfgang
2006 *Syntax.* Stauffenburg, Tübingen. Two volumes.

Svenonius, Peter
1994 C-selection as feature-checking. *Studia Linguistica*, 48: 133–155.

Takahashi, Daiko
1994 *Minimality of Movement.* Ph.D. thesis, University of Connecticut.

Thráinsson, Höskuldur
2007 Verb placement in embedded clauses: Some issues. Ms., University of Iceland, Reykjavík.
van Benthem, Johan
1988 The Lambek calculus. In Richard Oehrle, Emmon Bach, and Deirdre Wheeler, (eds.), *Categorial Grammars and Natural Language Structures*, pp. 35–68. Reidel, Dordrecht.
van Riemsdijk, Henk
1978 *A Case Study in Syntactic Markedness: The Binding Nature of Prepositional Phrases*. Foris, Dordrecht.
Vikner, Sten
1990 *Verb Movement and the Licensing of NP-positions in the Germanic Languages*. Ph.D. thesis, University of Geneva.
1994 Scandinavian object shift and West Germanic scrambling. In Norbert Corver and Henk van Riemsdijk, (eds.), *Studies on Scrambling*, pp. 487–517. Mouton de Gruyter, Berlin.
1995 *Verb Movement and Expletive Subjects in the Germanic Languages*. Oxford University Press, New York and Oxford.
Vikner, Sten and Eva Engels
2006 Object shift, remnant VP-topicalization and optimality theory. Ms., University of Aarhus.
Williams, Edwin
2003 *Representation Theory*. MIT Press, Cambridge, Mass.
2005 A note on "cyclic linearization". *Theoretical Linguistics*, 31: 229–234.
Zaenen, Annie, Joan Maling, and Höskuldur Thráinsson
1985 Case and grammatical functions: The Icelandic passive. *Natural Language and Linguistic Theory*, 3: 441–483.
Zifonun, Gisela, Ludger Hoffmann, Bruno Strecker, et al.
1997 *Grammatik der deutschen Sprache*. de Gruyter, Berlin.
Zwart, Jan-Wouter
1993 *Dutch Syntax. A Minimalist Approach*. Ph.D. thesis, Rijkuniversiteit Groningen.

Strategies of Subject Extraction*

Luigi Rizzi and Ur Shlonsky

1 Introduction

A major motivation for the classical ECP is the explanation of subject-object asymmetries, illustrated by the French examples in (1). Traces must be properly governed, the core case of proper government being government by a lexical head (Chomsky 1981). Subject traces fail to be properly governed, so that subject extraction is often problematic cross-linguistically. No comparable problem arises for object traces, which are properly governed by the verb:

(1) a. * Qui crois-tu que t_{qui} va gagner?
'Who do you believe that will win?'

b. Qui crois-tu que Paul va aider t_{qui}?
'Who do you believe that Paul will help?'

Contrary to the principles licensing other types of null or overt elements (overt DP's, pro or PRO), the ECP is hard to state in terms fully compatible with minimalist guidelines.

Minimalism envisages two fundamental types of principles, interface-driven and economy principles. The ECP is not naturally amenable to either (reliance on "government" is another frequently mentioned problem, but, apart from terminology, the issue doesn't really seem to arise in variants of minimalism allowing direct relations to be established between a head and a local DP in its domain, as in Chomsky 2000 and subsequent work). A separate and more theory-neutral problem is that it is hard to find a natural characterization of the class of proper governors going beyond a simple enumeration of cases: Why should local c-command by certain functional heads license a trace in some cases and not in others? For instance, why should Agr, T and other inflectional

*The authors are grateful to the editors of this volume and to a reviewer. Portions of this paper have been presented at numerous venues over the last three years and the comments and questions of many listeners have served to sharpen many points. An extension of the present paper appears in the bibliography as Rizzi & Shlonsky 2006.

elements suffice to license subject movement from the Theta position to the EPP position, possibly through intermediate positions, while normal C (*que*, *that*, *for*, etc.) would not? Why would only special complementizers such as French *qui* and English ∅ count as proper governors?

In this article, we would like to explore a different analytic path to subject object asymmetries which eschews the pitfalls of the ECP approach.[1] We adopt two ideas presented in Rizzi 2003:

(2) a. An element moved to a position dedicated to some scope-discourse interpretive property, a criterial position, is frozen in place (Criterial Freezing).

b. Classical EPP, the requirement that clauses have subjects, can be restated as a criterial requirement, the Subject Criterion, formally akin to the Topic Criterion, the Focus Criterion, the Q or Wh Criterion, etc., Rizzi 1996, 1997.

In accordance with (2b), thematic subjects move to the criterial subject position. By (2a), they are frozen there by Criterial Freezing. Thus, the non-extractability of subjects in cases like (1a) is explained. Movement of objects and other complements is not similarly constrained since there is no Object Criterion, parallel to the Subject Criterion.

From this viewpoint, subject extraction is only possible when the thematic subject is allowed to skip the criterial Subject position. Strategies of subject extraction that different languages use amount to ways of skipping the freezing position. Alternatively, languages may develop strategies for forming A'-chains on embedded subjects without moving them. In section 2, we introduce the basic theoretical background of this analysis. Sections 3-9 describe some of these ploys. Section 10 concludes the article.

2 Criterial Freezing and the Subject Criterion

Rizzi 2003 observes that a phrase meeting a Criterion (= reaching a position dedicated to a particular scope-discourse interpretive property in the terms of Chomsky 2001), is frozen in place and resists further movement to a distinct and higher criterial position (Criterial Freezing). For instance, a wh-phrase satisfying the Q-Criterion in an embedded question cannot undergo further

[1] For other, recent attempts to explain subject-object extraction asymmetries in non-ECP terms, see, in particular, Boeckx 2003, Landau 2007 and Roussou 2002. No comparison with these or with other recent approaches will be undertaken in this paper.

focus movement to the main clause as in (4b), an operation which is normally available in Italian, i.e., to the direct object in (3a).

(3) a. Pensavo che avessero scelto la RAGAZZA, non il ragazzo
'I thought they had chosen the GIRL, not the boy'

b. La RAGAZZA pensavo che avessero scelto ___, non il ragazzo
'The GIRL I thought they had chosen ___, not the boy'

(4) a. Mi domandavo quale RAGAZZA avessero scelto, non quale ragazzo
'I wondered which GIRL they had chosen, not which boy'

b. * Quale RAGAZZA mi domandavo ___ avessero scelto,
non quale ragazzo
'Which GIRL I wondered they had chosen, not which boy'

So, criteria cannot be satisfied "in passing". For instance, a complex phrase like *[quanti libri del quale]* 'how many books by whom$_{REL}$' can't be pied-piped from the complementizer-system of an indirect question to a higher relative complementizer.

(5) Gianni, [___ C_{REL} [non è ancora stato chiarito [[quanti libri del quale] C_Q [siano stati censurati t_{DP}]]]]
'Gianni, it has not been clarified yet how many books by whom have been censored'

Given an intermediate representation like (5), it is completely impossible to pied-pipe to the relative complementizer-system the complex phrase satisfying the Q-criterion in the indirect question, as in (6b). Two other derivational options are available from (5): either the PP *del quale* is subextracted, yielding (6a), or the whole indirect question is pied piped to the relative C-system, as in (6c) (a marginal structure, as clausal pied-piping generally is in Italian).

(6) a. Gianni, [del quale C_{REL} [non ancora stato chiarito [[quanti libri t_{PP}] C_Q [siano stati censurati t_{DP}]]]]
'Gianni, by whom it has not been clarified yet how many books have been censored'

b. *Gianni, [[quanti libri del quale] C_{REL} [non è ancora stato chiarito [t_{DP} C_Q [siano stati censurati t_{DP}]]]]
'Gianni, how many books by whom it has not been clarified yet have been censored'

c. ?(?) Gianni [[[quanti libri del quale] C_Q [siano stati censurati t_{DP}]] C_{REL} [non è ancora stato chiarito t_{CP}]]
'Gianni, how many books by whom have been censored, it has not been clarified yet'

So, the following principle seems to hold:

(7) Criterial Freezing: A phrase meeting a criterion is frozen in place.[2]

Criterial Freezing may be thought of as a condition clearing an element from narrow syntax as soon as it has reached a position dedicated to scope-discourse semantics; as such, it has an economy flavor which makes it similar to other devices intended to minimize memory resources in syntactic computations.

The second ingredient we need to deal with ECP effects is a characterization of EPP in criterial terms. Rizzi 2003 proposes that classical EPP ("Clauses must have subjects") can be advantageously reanalyzed as a Subject Criterion: the functional head Subj, distinct from and higher than T and other heads in the functional structure of the clause (Cinque 1999), attracts a nominal to its Spec and determines the subject-predicate articulation. Subj gives rise to the following configuration:

(8) [DP [Subj XP]]

Configuration (8) receives an interpretation paraphrasable as "About DP, I'm reporting event XP". Subjects thus share an interpretive property of topics, the "aboutness" relation linking subjects and predicates as well as topics and comments. In other respects, subjects are distinct from topics: contrary to topics, subjects do not require D-linking, so that a Subject-Predicate

[2]This statement should be sharpened to the effect that only the feature-bearing element is subject to criterial freezing. This formulation correctly rules in (6a). Moreover, subextraction in (6a) should be contrasted with preposition stranding by wh-movement to a matrix Comp from e.g., an embedded topic position: (ib) cannot be derived from (ia) by wh-movement from the topic position.

(i) a. I think that with this guy it would be interesting to exchange ideas.

b. * Which guy do you think that with ___ it would be interesting to exchange ideas

(ib) is reminiscent of cases discussed in Postal 1972 and may be treated as a violation of criterial freezing, on the assumption that the DP is the source of both relevant criterial features in (i), Top and Wh. The DP satisfies the Topic Criterion in the embedded sentence and is blocked from moving to the matrix Q position.

structure can be uttered in out-of-the-blue contexts, while a Topic-Comment structure cannot, see Rizzi 2005b. Once the subject Criterion is introduced, the representations of (1a,b) become the following, for the relevant parts:

(9) a. * Qui crois-tu [que [t_{qui} Subj va gagner]]?

b. Qui crois-tu [que [Paul Subj va aider t_{qui}]]?

(9a) is ruled out by Criterial Freezing, whereas no problem arises for object extraction in (9b), as there is no Object Criterion (that is, no object equivalent of classical EPP.)[3]

This perspective on the subject object extraction asymmetry provides a principled account of the immovability of subjects. Of course, the analysis must be modulated to account for the fact that languages do have ways of forming questions and other A'-constructions which target (embedded) subjects. Such strategies fall into two broad categories:

A. Fixed subject strategies: The subject doesn't move, it remains in its freezing position in Spec/Subj and a well-formed A'-construction involving the subject is obtained
1. with no movement at all (resumption), or
2. with movement of a larger constituent including the "frozen" subject (clausal pied-piping).

B. Skipping strategies: The subject moves, but it is allowed to skip the freezing position and is extracted directly from its thematic position or from some other predicate-internal position.

The most straightforward case of strategy A is the use of a resumptive pronoun for A'-constructions involving embedded subjects; a familiar case of strategy B is subject extraction from a lower position in Italian and other Null Subject languages, with the Subject Criterion satisfied by expletive *pro*. We would like to argue that variants of this strategy are more widespread than traditionally assumed. In sections 3 and 4, we discuss the two basic cases of strategy A, and in the rest of the paper we address various strategies amenable to the general pattern B.

[3]See also Boeckx's 2003 Principle of Unambiguous Chains, which requires chains to contain at most one strong occurrence, i.e. one EPP-position in the sense of Chomsky (2000), (which roughly corresponds to a criterial position in our sense) as well as Richards 2001, chapter 4. We will not attempt to compare the empirical consequences of these approaches here.

3 Unmovable embedded subjects: Resumptive pronouns in Hebrew relatives

Some subject-object asymmetries reported in the literature are not naturally amenable to classical ECP, but follow rather straightforwardly from a criterial freezing account.

A case in point is the following subject-object asymmetry in Hebrew restrictive relative clauses with resumptive pronouns, first discussed in Borer 1984:249-250: An object resumptive pronoun can appear either in-situ or be fronted to any higher topic or topic-like position in CP, but a subject resumptive pronoun can only remain in situ. Contrast the examples in (10) with those in (11) (resumptive pronouns in bold).

(10) a. kaniti et ha-šulxan še xana amra še dalya ma'amina
(I).bought ACC the-table that Hannah said that Dalya believes
še Kobi raca **oto**.
that Kobi wanted him

'I bought the table that Hannah said that Dalya believes that Kobi wanted.'

b. kaniti et ha-šulxan še Xana amra še dalya ma'amina
(I).bought ACC the-table that Hannah said that Dalya believes
še **oto** Kobi raca ___.
that him Kobi wanted

c. kaniti et ha-šulxan še xana amra še **oto** dalya
(I).bought ACC the-table that Hannah said that him Dalya
ma'amina še kobi raca ___.
believes that Kobi wanted

d. kaniti et ha-šulxan še **oto** xana amra še dalya
(I).bought ACC the-table that him Hannah said that Dalya
ma'amina še Kobi raca ___.
believes that Kobi wanted

'I bought the table that Hannah said that Dalya believes that Kobi wanted.'

(11) a. kaniti et ha-šulxan še xana amra še dalya ta'ana
(I).bought ACC the-table that Hannah said that Dalya claimed
še **hu** ya'ale harbe kesef.
that he will.cost a.lot money

b. * kaniti et ha-šulxan še xana amra še **hu** dalya
(I).bought ACC the-table that Hannah said that he Dalya
ta'ana še ___ ya'ale harbe kesef.
claimed that will.cost a.lot money

c. * kaniti et ha-šulxan še **hu** xana amra še dalya
(I).bought ACC the-table that he Hannah said that Dalya
ta'ana še ___ ya'ale harbe kesef.
claimed that will.cost a.lot money

'I bought the table that Hannah said that Dalya claimed that will cost a lot of money.'

The data in (10) were interpreted by Borer as evidence for the successive cyclic nature of wh-movement. The data in (11), however, posed a problem which she resolved by attributing to subject relative operators a language-specific lexical property.

In the approach developed in this article, a more general solution is available. We want to claim that (11b,c) are ungrammatical because the resumptive pronoun satisfies the Subject Criterion in Spec/Subj and is consequently frozen in this position. No problem arises for object topicalization (10), as there is no Object Criterion.

The complete picture is slightly more complex, though. There is no ban as such in Hebrew on subject relativization, as shown by the full acceptability of (12):

(12) kaniti et ha-šulxan še xana amra še dalya ta'ana še
(I).bought ACC the-table that Hannah said that Dalya claimed that
ya'ale harbe kesef.
will.cost a.lot money

'I bought the table that Hannah said that Dalya claimed that will cost a lot of money.'

The grammaticality of (12) shows that Hebrew must possess some device for extracting a subject without moving it first to Spec/Subj, a device of the 'skipping' kind that we will discuss in section 6 below.

The question then arises why the mechanism operative in (12) is not available for the resumptive relatives of (11): why is the subject resumptive pronoun forced to move to Spec/Subj, where it gets frozen, while the null

relative operator (or the relative clause head under a raising analysis: Vergnaud 1974, Kayne 1994, Bianchi 1999) can skip it?

We believe that the answer is provided in part by the weak nature of resumptive pronouns. Although Hebrew does not morphologically distinguish weak from strong pronouns (in the sense of Cardinaletti & Starke 1999 and related work, see Laenzlinger & Shlonsky 1997, Shlonsky 1997), it can be plausibly argued that both the object and the subject pronouns in (10) and (11) are weak, since they can be associated with an inanimate relative head (strong pronouns being typically restricted to animate referents.) The status of weak pronoun is not per se incompatible with occurrence in the peculiar pronominal topic position involved in Hebrew resumptive relatives, as the object case shows, but we would like to argue that its weak character forces the subject resumptive to move to Spec/Subj, thus triggering the freezing effect.[4]

Although (some[5]) weak subject pronouns are possible in postverbal position in Hebrew, as in (14b), they may not remain below adverbs such as 'usually' in a post verbal or inverted position, as in (14a). Compare with the freer distribution of full DP's in (13):

(13) a. matai yocet be-derex klal Rina la-sadot?
when goes.out usually Rina to-the.fields

b. matai yocet Rina be-derex klal la-sadot?
when goes.out Rina usually to-the.fields

'When does Rina usually go out to the fields?'

(14) a. * matai yocet be-derex klal **hi** la-sadot?
when goes.out usually she to-the.fields

b. matai yocet **hi** be-derex klal la-sadot?
when goes.out she usually to-the.fields

'When does she usually go out to the fields?'

[4]Weak pronouns in Italian - unlike those of Hebrew - are morphologically distinct from strong ones. As the contrast below indicates, a strong pronoun like *lui* 'he', can appear both pre- and postverbally while its weak counterpart *egli* is restricted to preverbal position:

(i) a. Egli/lui/Gianni ha parlato
He-weak/he-strong/Gianni has spoken

b. Ha parlato Gianni/LUI/*egli

[5]First and second person pronouns, for example, can only appear in a postverbal position under specific circumstances, see Doron 1988, Shlonsky 1997, 2000.

The positional constraints on weak pronouns are usually dealt with by assuming that a weak pronoun is licensed in a Spec/head configuration with a designated head, plausibly Subj for weak subject pronouns. So, subject weak pronouns must move to Spec/Subj to satisfy this requirement, and they can't stay in the lower position below the adverbial in (14a). Ordinary subject DP's do not have to meet this requirement, as shown in (13a). (The VS order in (14b) is presumably obtained by further leftward movement of the inflected verb, after the pronoun has moved to Spec/Subj; see Shlonsky & Doron 1992, Shlonsky 1997, 2006.)

In conclusion, subject resumptive pronouns in Hebrew are weak pronouns. This is not per se incompatible with the pronominal topicalization found in Hebrew resumptive relatives, as the case of object resumptive relatives shows, but it forces subject resumptives to move to Spec/Subj, where they get frozen. Whatever designated licensing head there may be for weak object pronouns, it has no freezing effect, as there is no object criterion; relative operators (or moved relative heads) are not weak pronouns and can therefore skip the freezing position in cases like (12), through the technique(s) discussed later in this article.

4 Clausal Pied-Piping in Imbabura Quechua

Criterial Freezing precludes satisfaction of criteria "in passing", so that the same element cannot satisfy two or more criteria in distinct positions. Sometimes, minimal use of already available mechanisms is made to rule in structures involving multiple criterial satisfaction without violating the freezing constraint. We have seen that clausal pied-piping can marginally solve the problem raised by the simultaneous satisfaction of Q and Rel Criteria by elements of the same complex phrase in Italian (section 2).

Imbabura Quechua (IQ) employs a similar strategy for solving the problem of subject extraction.[6] Consider the contrast between object and subject extraction from an embedded clause. Object extraction can take two forms. The first is straightforward wh-movement to Comp, illustrated in (15) (see Cole 1985, Cole & Hermon 1981 and Hermon 1984, from where the data is taken. See also Richards 2001.)

(15) ima-ta-taj Maria-ka Juzi miku-shka-ta kri-n?
what-ACC-Q Maria-TOP José eat-NOMINALIZER-ACC believe-AGR
'What does Maria believe that José ate?'

[6]Thanks to Gabriela Hermon for discussion of the Quechua data in this section.

Note that the wh object lands to the left of the particle *-taj*, which we assume realizes the criterial Q head.[7]

The second strategy for wh-movement in IQ involves movement of the wh object to the embedded Comp-system and subsequent pied-piping of the whole embedded CP to the left of the criterial head *-taj*.[8]

(16) ima-ta wawa miku-chun-taj Maria kri-n?
what-ACC child eat-FINITE-Q Maria believe-AGR

'What does Maria believe (that) the child eat?'
Lit. '[What the child eat] does Maria believe?'

In contrast to the object question, a wh question on the embedded subject can only utilize the pied piping strategy; compare the ungrammatical wh-extraction in (17a) and clausal pied-piping in (17b):

(17) a. * pi-taj Maria-ka chayamu-shka-ta kri-n?
who-Q Maria-TOP arrive-NOMINALIZER-ACC believe-AGR

'Who does Maria believe (that) has arrived?'

b. [pi chayamu-shka-ta-taj] Maria ___ kri-n?
who arrive-NOMINALIZER-ACC-Q Maria believe-AGR

'Who does Maria believe (that) has arrived?'
Lit. '[Who has arrived] does Maria believe?'

The ungrammaticality of (17a) is immediately captured as a violation of Criterial Freezing: the wh subject moves to Spec/Subj to satisfy the Subject Criterion in the embedded clause, it is frozen there, and further movement is blocked. Since there is no Object Criterion, objects can be freely extracted out

[7] Although some superficial resemblance might be found between *taj* and the scope-determining Q-particle in Sinhala (see Kishimoto 2005), there are substantial differences in the strategies of question formation in Imbabura Quechua, where wh-movement to the left periphery is obligatory, and Sinhala, a wh in situ system, which suggest that the two should not be conflated. A reviewer points out that in cases of multiple wh movement to the left periphery in Imbabura Quechua, each wh-word is followed by *taj*, as Cole 1985 notes. This might suggest that the projection housing wh-elements in the the Imbabura Quechua left periphery can be recursive, with potential consequences for the characterization of multiple wh fronting, which we do not pursue here.

[8] This kind of pied-piping strategy is reminiscent of Basque, where some islands can be circumvented by pied piping the entire island (Ortiz de Urbina 1989; see also Richards 2000 on Japanese).

of IP, as in (15). The pied-piping option allows the subject to bypass criterial freezing (while remaining one of the two options available to objects.) In (17b), the whole embedded clause is pied piped (recall that only the head and the specifier are frozen, but not the XP containing them: movement remains possible as long as the criterial configuration is not undone,) and moved to Spec/Q in the matrix clause. Both the Subject and Q criteria are thus satisfied without violating Criterial Freezing, much as the Q and Rel criteria in (6c).

A question arises as to the categorial nature of the pied-piped constituents. In (16), the object is first moved to the embedded C system, and then the category containing it, the whole CP, is pied piped to the main C-system. The category pied-piped in (17b) would seem to be smaller than the one moved in (16). If the wh subject is frozen in Spec/Subj and the pied-piped phrase must have the wh element in its Spec, the pied-piped projection in (17b) would seem to be SubjP, rather than CP. This may be so: it is, after all, well-known that the size of the pied-piped constituent can vary considerably. On the other hand, movement of an IP-like constituent stranding C is rather unprecedented (and precluded, if IP-like constituents are not Phases in the sense of Chomsky 2001, and non-phase categories are unmovable, or at least inaccessible to long-distance movement). So, it is worthwhile to consider a (minimal) alternative to this analysis.

We have argued that subject raising to Spec/Subj is the only mechanism available in IQ to satisfy the Subject Criterion. This turns out to be too strong. Local wh-movement seems to be possible, as evidenced by (18). Since the subject wh precedes the Q head *taj*, it must have been moved.

(18) pi-taj shamu-rka?
who-Q left-AGR

'Who left?'

How can this be reconciled with the Subject Criterion and Criterial Freezing? This is a particular case of the general issue of local subject movement to C, an issue we discuss at length in section 7, where we introduce a mechanism which enables local movement.

With respect to (17b), we can adopt this mechanism to allow local movement of the embedded subject to the embedded C-system, except that the question feature in the embedded clause is not criterial (*believe* does not select for an indirect question), but the purely formal counterpart of Q in the sense of Rizzi 2003, namely, the formal feature which drives successive cyclic movement of wh operators to eventually reach the criterial Q position. That this is

legitimate from the perspective of the Last Resort guideline is evidenced by the acceptability of (16) and more generally, by the crosslinguistic evidence for 'internal movement', in the sense of van Riemsdijk 1984 - the strategy of moving a wh operator to the edge of a pied-piped constituent, transforming it into a complex wh phrase. A case of legitimate internal movement in a familiar construction is the following: local movement of the wh expression to the edge of DP in (19b) is rendered legitimate by the interface effect of creating a complex operator which then moves to the criterial position. In the absence of this effect, such local, DP-internal movement violates the Last Resort guideline, (19c).

(19) a. He's prepared to buy [a [very expensive] car]

b. [[How expensive] a t car] is he prepared to buy t?

c. * He's prepared to buy [[very expensive] a t car]

The local subject movement which makes the clausal pied-piping possible in (17b) seems to be of a similar kind.

5 Skipping Spec/Subj: Null Subject Languages

The second family of strategies for extracting subjects in A'-constructions permit the thematic subject to skip the Spec/Subj position and obviate criterial freezing.

An obvious manifestation of such strategy is the filling of the criterial position with another element, e.g. an expletive.

(20) . . . Subj is [what in the box]

For instance, in English copular constructions, when a derivation reaches point (20), if expletive *there* is inserted to satisfy the Subject Criterion, then the thematic argument *what* remains available for further movement, ultimately yielding (21b); if no expletive is used and *what* is raised to Spec/Subj in (20), further wh-movement as in (21a) is barred by Criterial Freezing.

(21) a. * What do you think that t_{what} is in the box?

b. What do you think that there is t_{what} in the box?

This, in essence, is the strategy used by Null Subject languages to avoid ECP violations with subject extraction, according to the analysis of Rizzi 1982: the insensitivity to complementizer-trace effects, illustrated by the well-formedness of (22a) in Italian, is explained by assuming a representation like (22b), with the preverbal subject position (in current terms, Spec/Subj) filled by pro and the wh subject extracted from the thematic (or some other low) position (see Rizzi 1990 for discussion and review of the cross-linguistic evidence for this analysis):

(22) a. Chi credi che vincerà?
'Who do think that will win?

b. Chi credi [che [pro Subj vincerà t_{chi}]]
'Who do you think that will win'

In the original proposal, the expletive *pro* in Null Subject languages had the role of fending off an ECP violation by permitting subject extraction from a properly governed position. In the current framework, expletive *pro* is instrumental in formally satisfying the Subject Criterion, hence in allowing the thematic subject to escape the effects of Criterial Freezing.

This analysis, and, more generally, the fact that expletives exist, raises an important question for the criterial approach. Why can an expletive satisfy the Subject Criterion? The very existence of expletives, originally taken as the major piece of evidence for the EPP, is also commonly interpreted as providing critical evidence against attempts to link the obligatoriness of subjects to some kind of special interpretive property associated to the subject position. As an expletive is devoid of interpretive (referential) content, the argument goes, the obligatoriness of subjects must be treated as a purely formal principle.

Nevertheless, the conclusion that expletive-like elements cannot be involved in the satisfaction of genuine criterial (scope-discourse) properties seems to be too strong. Consider, for instance, the so called "partial wh movement" construction, possible in colloquial German and several other languages:

(23) Was glaubst du welchen Mantel Jakob heute angezogen hat?
What believe you which coat Jakob today put.on has?

'Which coat to you believe Jakob put on today?'

According to one familiar analysis, (McDaniel 1989), the substantive wh phrase is moved to the embedded C-system which is not criterial (a verb

like *glauben* ('believe') does not select an indirect question), while the Spec relevant for the Q Criterion is filled by an invariable expletive-like wh element *was* ('what'), and acts as a kind of scope-marker for the substantive wh phrase. Under this analysis, partial movement looks like the A'-equivalent of an expletive construction in the A-system. So, the use of expletive-like elements is not inherently incompatible with the system of Criteria.

True, the use of expletives seems to be more widespread in subject position than in A'-constructions, in that many languages lack any kind of partial A'-movement, while some form of overt or null subject expletive is presumably available in all languages. Still, this state of affairs is not difficult to understand if we think of the special status that the Subject Criterion must inevitably be assumed to have in the system of Criteria. The Subj layer defines a structural zone connecting the CP and the IP-systems. As such, it may be assumed to share properties with both systems. The CP zone is specialized in creating dedicated positions to express scope-discourse properties, topicality, focus, scope of different kinds of sentential operators; such positions are formally optional, in the sense that they are activated in a structure when the discourse conditions and communicative intentions require them. Otherwise, they remain inert. On the other hand, a notable characteristic of the IP zone is obligatoriness, at least the obligatoriness of the heads forming the backbone of the 'functional' IP hierarchy; tense in the first place (Cinque 1999). So, we may think of the Subj layer as sharing properties of the two systems it connects: on a par with the CP-system, it is dedicated to a scope-discourse property and on a par with the IP-system, it is obligatorily expressed.

There is a certain tension between these two properties, as formal optionality is characteristic of the expression of scope-discourse properties. We may think of expletives as a way to resolve this tension: when discourse conditions, communicative intentions or the thematic structure of the verb require a presentational structure, in which a certain event is not described as being "about" a certain argument, an expletive is used to formally satisfy the Subject Criterion. The interpretive systems, receiving a representation in which no argument is expressed in the aboutness position, interpret the structure presentationally.

The case of expletives is similar to many cases in natural language syntax, where a formal device has a core interface function and a somewhat larger domain of formal application. The device acquires, as it were, a formal life of its own, extending its scope beyond the core interface effects which functionally motivate it. One example of this mode of functioning is grammatical gender, extending from natural gender to an arbitrary classification of the entire

nominal system in many languages. Other cases may include the obligatory focus position in Old Italian, which, according to Poletto 2005, can be filled with an expletive element when not used to express focus, and perhaps even the V2 constraint in Germanic can be looked at as a formal generalization of interpretively determined Spec/head requirements. The linguistic representation of tense also manifests a formal life going beyond its core interpretative function: mathematical and logical truths are atemporal, nevertheless when we express them through language we don't use an untensed sentence, but a sentence with unmarked tense, to comply with the formal requirement which makes T obligatory in syntactic structures. Similar considerations may hold for uses of the subjunctive, extending from core cases of irrealis to numerous other configurations.

The original analysis of subject extraction in Null Subject languages linked the apparent insensitivity to the *that*-trace effect to so called 'free subject inversion', namely, to the option of VS order. Both were taken to be contingent on the possibility of filling the preverbal subject position with expletive *pro* and leaving the thematic subject in a lower, predicate-internal position. In such a configuration, it was argued, the thematic subject is accessible to movement directly from this (properly governed) position.

A problem for this analysis was pointed out by Chao 1981, who observed that the process of inversion in Brazilian Portuguese is less free than in other Romance languages, compare (24a) with its Spanish equivalent in (24b). Nevertheless, Brazilian Portuguese is insensitive to complementizer-trace effects, as in (25).

(24) a. *(João disse que) saíram eles [BP]

b. (Juan dijo que) salieron ellos [Sp]
'(J. said that) left they'

(25) Quem o João disse que vai chegar tarde? [BP]
'Who J. said that is going to arrive late?'

Chao argued that the apparent violability of the complementizer-trace effect is not necessarily contingent upon free inversion. Her alternative analysis was to assume a resumptive *pro* in subject position in structures like (25), an approach which would be consistent with the present framework: the case would reduce to another instance of the Fixed Subject Strategy A (see section 2).

Yet Menuzzi 2000 provides interesting evidence suggesting that in BP, as in Italian, wh extraction of the subject takes place from a position lower than

what in our terms is Spec/SubjP. He observes that extracted wh phrases can launch floated quantifiers which can appear in different lower positions, but not in preverbal position. In terms of Sportiche's 1988 analysis of Q-float as Q-stranding, this suggests that subject extraction skips Spec/SubjP:

(26) a. Que rapazes o Paulo desconfia que tenham beijado *todos* a Maria?
'Which boys Paulo suspects that have kissed all Maria?'

b. Que rapazes o Paulo desconfia que tenham *todos* beijado a Maria?

c. * Que rapazes o Paulo desconfia que *todos* tenham beijado a Maria?

Notice that BP has an expletive *pro*, as illustrated by the following constructions:

(27) a. *pro* parece que o José passou por aqui.
'seems that J. came by here'

b. *pro* choveu a noite inteira
'rained all night'

So, it appears that the language can use the Italian strategy: the subject Criterion is satisfied by expletive *pro* and the thematic subject is extracted from a lower position. The impossibility of (24a) must, then, be due to some other factor. Belletti 2001, 2004 reanalyzes free subject inversion as subject focalization, which involves movement of the thematic subject to a low, predicate-internal focal position. The ungrammaticality of (24a) might then be related to the unavailability of this kind of subject focalization in BP. This option is partially independent from the insensitivity to complementizer-trace effects (but not completely unrelated, the availability of expletive *pro* being instrumental for both properties). See Nicolis 2005 for a cross-linguistic appraisal of the issue along these lines.

6 *Que-qui* phenomena: the expletive approach

The following paradigm illustrates the much-debated *que-qui* phenomenon in French. When the relativized element is the local subject, the complementizer obligatorily assumes the form *qui*, an option which is excluded when the relativized element is the object:

(28) a. * L'homme [Op que [t est venu]]
'The man QUE has come'

b. L'homme [Op qui [t est venu]]
'The man QUI has come'

(29) a. L'homme [Op que [tu as vu t]]
'The man QUE you have seen'

b. * L'homme [Op qui [tu as vu t]]
'The man QUI you have seen'

The phenomenon is not a specific property of the relative clause complementizer-system. The alternation is observed with simple questions in varieties which allow the co-occurrence of the wh element and the overt complementizer, such as Québec French:

(30) a. Qui que [tu as vu t]?
Who QUE you have seen?'

b. Qui qui [t est venu]?
'Who QUI has come?'

It is also observed in cases of subject extraction from an embedded clause, in relatives and questions, even though here the acceptability of the *qui* variant, as in (31b), appears to be dialect-specific, while straight subject extraction with *que* (31a) and non-subject extraction with *qui* (32b) are uniformly rejected:

(31) a. * Quelle étudiante crois-tu [t' que [t va partir]]?
'Which student do you believe that is going leave?'

b. % Quelle étudiante crois-tu [t' qui [t va partir]]?
'Which student do you believe QUI is going to leave?

(32) a. Quelle étudiante crois-tu [t que [Marie va aider t]]?
'Which student do you believe that Marie is going to help?'

b. * Quelle étudiante crois-tu [t' qui [Marie va aider t]]?
'Which student do you believe QUI Marie is going to help?'

The analysis in Rizzi (1990), capitalizing on previous proposals by Taraldsen 1978, Pesetsky 1982, among others, ran as follows: *qui* is the "agreeing variant" of *que*; the wh element passing through the Spec/C can trigger agreement of C, which is morphologically manifested by the form *qui*:

(33) qui = que+Agr (Rizzi 1990)

In an ECP-based framework, the activation of agreement turns C into a proper governor for the subject trace. The agreeing form cannot occur with object movement, as in (32b), because, if the Spec of an agreeing head is (of the same type as) an A position, the chain (t, t') in (32b) crosses another A position, the subject position, in violation of Relativized Minimality.

This analysis is rather straightforward, but it has to face a problem of morphological plausibility: given the nominal or verbal morphological expression of Phi-features found elsewhere in French, it is not very plausible to think of *-i* as an agreement marker, as nothing similar appears in the verbal or nominal agreement paradigms. A similar objection can be raised against an agreement analysis of the analogous *dat - die* alternation in West Flemish, Bennis & Haegeman 1984, Haegeman 1992, with the aggravating factor that a genuine agreeing form of the C-system is found in the language, giving rise to a completely different morphological alternation (*dat - dan*, the latter form expressing plural agreement with the subject).

Taraldsen 2001 proposes a different approach to *qui* which has more morphological plausibility. He argues that the form should be analyzed as *que + -i*, where *-i* is an expletive-like element akin to the standard French expletive *il*, which appears with weather verbs, in subject extraposition and presentational sentences in French, as in (35).

(34) Taraldsen (2001): qui = que+Expl

(35) Il est arrivé trois filles
'It arrived three girls'

Under this analysis, the *que-qui* alternation is immediately traceable to the analysis of Null Subject languages, except that here it is the overt expletive *-i*, not *pro*, which fills the subject position - as in (36) - and permits extraction of the thematic subject from a lower position (either the thematic position *t*, as *venir* is an unaccusative verb, or some higher position *t'* in the functional structure). In our terms, *-i* satisfies the Subject Criterion in (36), hence it allows the relative operator corresponding to the thematic subject to be moved to the relevant position in the left periphery:

(36) L'homme Rel Op qu'[-i Subj est t' venu t]

As for the impossibility of *qui* with object extraction, sentences like (32b) are simply not derivable because the structure does not provide enough room

for the expletive *-i* and the subject DP. This analysis is immediately compatible with the Criterial Freezing approach, and in fact it reduces the French case to a variant of the device used in Null Subject languages.

Nevertheless, the analysis expressed by (36) must be refined, as there are a number of significant differences between *il* and *-i* which make a complete assimilation of the two impossible.

6.1 Number agreement

Il has its own number specification, singular, which triggers agreement on the verb (whether or not the nominal associate is moved,) as in (37a-b), while *-i* is compatible with whatever number specification the thematic subject has, as in (38):

(37) a. Il est (* sont) arrivé trois filles
'It is (are) arrived three girls'

b. Combien de filles est-ce qu'il est (* sont) arrivé?
'How many girls is it that it is (are) arrived?'

(38) Les filles qui sont arrivées
'The girls QUI are arrived'

6.2 Definiteness and no TEC

Expletive *il* requires an indefinite associate, and is limited to occur with specific verb classes. It occurs most naturally with unaccusative verbs, it has an intermediate status with unergative verbs and it is excluded with transitive verbs (i.e., French disallows the Transitive Expletive Construction):

(39) a. Il est arrivé une fille / * la fille
'It arrived a girl / the girl

b. ?(?) Il a téléphoné beaucoup d'étudiants
'It telephoned many students'

c. * Il a acheté ce livre une fille
'It bought this book a girl'

-i, on the contrary, is compatible with a definite associate (at least in the sense that the head of the relative clause can be definite), and is not sensitive to any verb-class restriction:

(40) a. La fille qui est arrivée
'The girl QUI is arrived'

b. La fille qui a téléphoné
'The girl that telephoned'

c. La fille qui a acheté ce livre
'The girl QUI bought this book'

6.3 Position

Il, both referential and expletive, must be adjacent to the inflected verb.

(41) a. * Il, la semaine prochaine, partira en Italie
'He, next week, will leave to Italy'

b. * Il, la semaine prochaine, viendra trois filles
'It, next week, will come three girls'

c. * Je crois qu'il, la semaine prochaine, partira en Italie
/ viendra trois filles
'I believe that he / it, next week, will leave to Italy
/ will come three girls'

d. Je crois que, la semaine prochaine, il partira en Italie
/ viendra trois filles
'I believe that, next week, he / it will leave to Italy
/ will come three girls

i, on the other hand, can be separated from the inflected verb by an adverbial, whereas it must remain agglutinated to *que*:

(42) a. L'homme qui, la semaine prochaine, partira en Italie
'The man qui, next week, will leave to Italy'

b. * L'homme que, la semaine prochaine, -i partira en Italie
'The man that, next week *-i* will leave to Italy'

The last set of observations is particularly revealing as to the impossibility of fully assimilating *il* and *-i*. *-i* clearly occupies a higher position than *il*, since it precedes adverbs and forms a word with the complementizer.

This positional property may also be responsible for *-i*'s insensitivity to verb classes. *Il* is a clitic hosted in the inflectional system, presumably externally merged in a position sufficiently low to make the insertion sensitive to the lexical properties of the verb. *-i*, however, may be thought of as a weak, clitic-like pronominal element externally merged in the complementizer-system, too high in the structure to be sensitive to verbal properties. More precisely, we would like to propose that *-i* is externally merged under Fin, the lowest head of the complementizer-system (Rizzi 1997), as a particular, nominal realization of this head. As such it can precede a fronted adverbial in (42a), which, following Benincà 2001 and Benincà & Poletto 2004, we assume can be positioned in the higher part of the IP-system.

Not only do *il* and *-i* vary positionally, they also differ in inherent constitution. Il is intrinsically marked as [Plural], whereas *-i* has an unvalued number feature which is valued when the subject moves to its Spec. Straightforward morphological evidence for the postulation of such a number feature is provided by the substandard variety of French, discussed in Laenzlinger 1997, in which the 'plural' *qui* is pronounced *qui[z]* in liaison contexts, i.e., it carries the standard plural morpheme *[z]* of the French nominal system:

(43) Moi, qui ai fait ça / Toi, qui as fait ça
Lui, qui a fait ça / Nous, qui[z] avons fait ça
Vous, qui[z] avez fait ça / Eux, qui[z] ont fait ça
'I, you, he, . . . QUI have/has done this'

We take *-i*, therefore, to be listed in the French lexicon with the following specification:

(44) -i : [+Fin], [+N], [aPl]

Let us now consider how *qui* can permit subject movement and extraction. Consider the derivation of *L'homme qui va partir*, 'the man qui is going to leave', starting at the level at which the thematic subject, here a relative operator, has been moved to Spec/Agr (or whatever head takes care of the Case-agreement properties), and the Subj head is merged as an obligatory component of the clausal structure:

(45) Subj [Rel Op Agr [va [t partir t]]]

At this point, if the relative operator is moved to Spec of Subj, it would satisfy the Subject Criterion there and would be stuck in that position under

Criterial Freezing; it could never reach the criterial position for relative operators in the left periphery and the structure would crash. But the derivation can continue on from (45) by directly merging the next higher head in the clausal structure, Fin, which can be selected in its 'nominal' variant, *-i*:

(46) Fin-i [Subj [Rel Op Agr [va [t partir t]]]]

Here, *-i* is a nominal element in a local configuration with Subj; as such, it satisfies the Subject Criterion (the configuration is not Spec-head here, but head-head, see below.) The relative operator, therefore, does not move to Spec/Subj, as the Subject Criterion is already satisfied by *-i*, and remains available for movement to the position where the Relative Criterion must be satisfied in the left periphery; we will further assume that, on its way to the relative position, the operator passes through Spec/Fin in order to value the number feature on *-i*. We thus obtain (47).

(47) L'homme [Rel Op que [t''' [Fin-i] [Subj [t'' Agr [va [t' partir t]]]]]]

A similar analysis can be proposed for the dialect-specific case of *que-qui* which permits subject extraction from an embedded declarative, as in (31b): *Quelle étudiante crois-tu qui va partir?*

(48) a. Subj [quelle étudiante Agr [va [t partir t]]] ⇒

[$_{Fin}$ -i] is merged and satisfies the Subj Criterion ⇒

b. [$_{Fin}$ -i] Subj [quelle étudiante Agr [t' va partir t]] ⇒

the wh phrase moves to Spec/[$_{Fin}$-i]
and values the number feature ⇒

c. Quelle étudiante [$_{Fin}$ -i] Subj [t'' Agr [t' va partir t]] ⇒

que is merged, the main clause structure is merged,
and then wh moves to the main clause C-system ⇒

d. Quelle étudiante crois-tu [que t''' [$_{Fin}$ -i] Subj [t'' Agr [t' va partir t]]]

As for the variation on the judgment on (48d), it can be assumed that the selection of *-i* is generally available in the relative C-system, while the

extension of this option to a declarative C-system is dialect specific (see section 8 for further discussion).

One important property of *que-qui* is the fact that it is only triggered when the (local) subject is moved. Object movement is incompatible with it:

(49) * Quelle étudiante crois-tu qui Marie va aider t?
'Which student do you believe QUI Marie will help?'

Consider the relevant derivational stage, when the nominal Fin head *-i* is merged immediately above the Subj layer, where it satisfies the Subject Criterion:

(50) [$_{Fin}$ -i] Subj Marie Agr [t va aider quelle étudiante]

At this point, the number feature in *-i* must be valued by attracting a nominal element. But the nominal cannot be the thematic subject, *Marie*: if it was attracted to Spec/[$_{Fin}$ -i], it would end up in a non-criterial position, in violation of movement as last resort (in this configuration, [$_{Fin}$ -i] and Subj are the two elements involved in criterial satisfaction; Spec/[$_{Fin}$ -i] is not.) The wh object would not run into that problem: as a wh element, it would eventually move to a criterial position, a Q position in the main complementizer-system. But the object cannot be attracted to Spec/[$_{Fin}$ -i] in (49), if the attractor is the unvalued number feature, because of locality / Relativized Minimality. The closest potential attractee in (49) is the subject, *Marie*. So, no grammatical output is derivable from (49), and the only case in which selection of *-i* in Fin can lead to a well-formed structure is when the wh element is the local subject.

Along similar lines, one can exclude selection of *-i* when no A'-movement to the left periphery takes place:

(51) * Je crois qui Marie va aider l'étudiante
'I believe QUI Marie will help the student'

Here *Marie* cannot be attracted to Spec/Fin$_{-i}$ for the same reason as in (50): the DP would end up in a non-criterial position, (* *je crois Marie qui va aider l'étudiante* in the interpretation 'I believe that Marie will help the student' is ruled out by the last resort principle), and if no attraction takes place, the number feature on *-i* would remain unvalued, and the derivation would crash.

It should be remembered here that in the system of Chomsky 2001, the valuation of an unvalued feature does not necessarily require movement: it

can be implemented through a simple probe-goal AGREE relation. AGREE is a necessary prerequisite for movement in that system, but does not require movement to take place. So, couldn't the valuation of [Plural] in *-i* be achieved by AGREE with the subject without movement of the latter, a derivational option which would incorrectly permit a structure like (51)?

We can observe, in this connection, that in some clear cases, feature valuation does indeed require movement. It is so in the case which provided the empirical basis for the first detailed model of a generative theory of agreement, French past participle agreement, (Kayne 1989). The participle cannot agree with the object in situ (as in (52a)), nor can it just attract the object to its Spec as in (52b), a derivation which would violate the Last Resort guideline on movement, because no scope-discourse interface effect is associated with the participial specifier. Agreement is possible when the object must move for independent reasons to a higher position, and triggers participial agreement in passing:

(52) a. * Jean a repeinte la chaise
'Jean has repainted the chair'

b. * Jean a la chaise repeinte t
'Jean has the chair repainted'

c. La chaise que Jean a t' repeinte t
'The chair that Jean has repainted'

In some cases, UG clearly requires that the valuation of an unvalued feature be executed via movement. It could be the case that movement is always required, or that it may be required or not, as a matter of parameterization. A plausible case of valuation via pure AGREE, without movement, for a feature in the C-system may be the phenomenon of number agreement of C with the subject which is observed in various Germanic varieties (Haegeman 1990, 1992, Carstens 2003.) In any event, as in some clear cases valuation requires movement, we may assume that this mechanism is involved in *-i* valuation, so that the observed structural properties follow.

The core of our analysis of *que-qui* is the assumption that *-i* in Fin can satisfy the Subject Criterion. This is not literally compatible with the format for criteria assumed in Rizzi 2003 and repeated here for convenience:

(53) For [+F] a criterial feature, X_{+F} is in a Spec-head configuration with A_{+F}.

[$_{Fin}$ -i] is not in a Spec-head configuration with Subj in (46). Rather, it is the immediately superordinate head to Subj in the clausal hierarchy. So, what we need is a more general characterization of the criterial configuration, one which encompasses both Spec-head and local head-head configurations. What the two configurations have in common is locality: nothing intervenes between the criterial head Subj and the element which satisfies the criterion, be it a Spec or a head. We can therefore restate (53) as follows:

(54) For [+F] a criterial feature, X_{+F} is locally c-commanded by A_{+F}.

A final problem that must be addressed by the analysis of *que-qui* is raised by contrasts like the following, which we introduced earlier as providing critical evidence for the different positions of *-i*, and *il*:

(55) a. L'homme qui, la semaine prochaine, partira en Italie
'The man who, next week, will leave to Italy'

b. * Il, la semaine prochaine, partira en Italie
'He, next week, will leave to Italy'

c. La semaine prochaine, il partira en Italie
'Next week, he will leave to Italy'

The problem is this : if *-i* and *il* occupy different positions, one necessarily higher and the other necessarily lower than the adverbial phrase, how can they both satisfy the Subject Criterion? The possibility that *-i* may satisfy the Criterion in the same position as il, and then raise to Fin across the adverbial element is precluded, given Criterial Freezing.

The first relevant observation here is that a lexical subject can occur to the left of the adverbial in such structures as (56).

(56) Jean, la semaine prochaine, partira en Italie
'Jean, next week, will leave to Italy'

So there must be a way to satisfy the Subject Criterion from that position, and this way may be used in (55a) as well. The problem now reduces to accounting for the contrast between (55b) and (56). Such pairs are discussed by Cardinaletti 2004 as providing evidence for (at least) two subject positions in the higher part of the inflectional field, i.e., in informal notation,

(57) Subject1 - Adverbial - Subject2 - Agr . . .

French subject clitics (and other kinds of weak pronouns) are specialized to occur in Subject2, while nonpronominal DPs can occur in either Subject1 or Subject2.

From our perspective, there can only be a single criterial Subj position (although there are surely a number of distinct positions which house subjects, see, e.g., Shlonsky 2000.) We therefore propose that in this subfield in-between the CP field, demarcated by Fin, and the IP field, classically assumed to be closed by the Agr-T system, the two heads Subj and Mod (attracting to its Spec a highlighted adverbial, Rizzi 2001, 2004) are freely ordered (much as, say, Mod and Top appear to be freely ordered in the left periphery). So, (56) must really be split into the two cases arising from the ordering options:

(58) a. Subj [AdvP Mod [Agr . . .]]

b. AdvP Mod [Subj [Agr . . .]]

In section 3, we exploited the requirement that weak subject pronouns must move to Spec/Subj in order to explain the freezing (in particular, the non-topicalizability) of Hebrew resumptive pronouns. French subject clitics are subject to a more stringent requirement - they must, in addition, end up in a position adjacent to an Agr specification (Cardinaletti & Starke 1999). This restricts their occurrence to configuration (58b) (if Agr locally raises to Subj, the weak pronoun in cases like (55c) would be in a Spec/head configuration with a head bearing Agr). Configuration (58a) is precluded for French subject clitics, as Agr is too far away (Agr to Subj being presumably blocked here by the intervening Mod head, under the Head Movement Constraint); but the configuration would be accessible to other kinds of subjects, as in (56). The nominal Fin strategy (clearly not a 'subject clitic' in the same sense as *il*, etc.) would exploit (58a), along the lines we have discussed.

7 Local Subject Questions

How can one derive local subject questions such as (59a,b)?

(59) a. Who came?

b. Qui est venu?

It is sometimes assumed that the wh-subject does not move to the C-system at all and remains in subject position, Spec/SubjP in our terms. If this is so, local subject questions don't raise any particular problem for our approach: the

Subject Criterion is satisfied and Criterial Freezing is operative, as in simple declaratives.

We find this line of analysis dubious on various grounds, though. First of all, if the normal scope site of wh-elements is in the C-system of the clause (i.e., we typically don't find moved wh-elements in the periphery of the vP-system or in some other position in the inflectional system), subject questions would be an exception.

Second, in languages with an overt Foc or Q head, subject wh elements are overtly moved to the Spec/Foc, on a par with other wh-elements (e.g., in Gungbe, the order is *who Foc came*, see Aboh 2004; or see the Imbabura Quechua example (18)).

Third, in languages in which the wh-element can (or must) co-occur with an overt complementizer, we typically have the sequence wh-subject - complementizer. This demonstrates that wh-movement to the left periphery has taken place.

(60) Quem que vai chegar? [Brazilian Portuguese]
'Who QUE will come ?'

Fourth, Null subject languages provide direct evidence that also in local movement the subject is moved directly from a predicate internal position to the left periphery, without passing through Spec/Subj (a position filled by expletive *pro* in our analysis). All the evidence supporting movement from a postverbal position with subject extraction (*ne*-cliticization, agreement patterns, etc.: Rizzi 1982, 1990) extends to the case of local movement. If UG permitted wh-elements to remain in the IP-initial subject position in simple subject wh-questions, it would not be clear why Null Subject languages could not use this option and must resort to movement from a lower position to the left periphery.

Fifth, subject questions are possible in indirect questions (*I wonder who came*). Under the IP analysis of (59), it would not be clear how to state selectional requirements in full generality (verbs like *wonder* should sometimes select a CP, sometimes an IP).

For these reasons, we will assume that the wh-element must be extracted from IP in cases like (59) and moved to the C-system, the natural scope domain of wh-operators. But if this is correct, how can it satisfy the Subject Criterion and escape the effect of Criterial Freezing, which would freeze it in Spec/Subj?

The dialectal varieties of French overtly manifesting *que-qui* in simple main questions directly show that the nominal Fin-strategy may be used for local subject movement as well, as in (61b):

(61) a. Quel garçon que tu as vu? [Québec French]
which boy QUE you have seen

'Which boy have you seen?'

b. Quel garçon qui est venu?
which boy QUI has come

'Which boy has come?'

So, one approach that this observation immediately suggests is that languages may use a variant of the quasi-expletive Fin-device with no overt morpho-phonological effects for local subject movement. In other words, it could be that Standard French (59b) and Québec French (61b) have in essence the same representation, except that standard French has an unpronounced occurrence of *qui*. In turn, this difference could be connected to the language-specific sensitivity to "doubly filled C" effects, permitting or excluding the occurrence of overt C-material with an overt wh-operator. Québec French permits this option, as (61a) shows, while Standard French does not. Remember that Standard French can use an overt *qui* with local subject movement in relatives like (28b), where the Spec is filled by a null operator, or, with dialectal variation, in cases of extraction from a declarative like (31b), where the Spec is a trace: in these cases, the wh-element is not pronounced (not locally, at least), hence the nominal Fin manifesting *i* can be pronounced in compliance with the "doubly filled C"-constraint. In local questions, where the operator is an overt wh-element, an overt *qui* is banned. Our proposal is that here the language resorts to the null variant of the nominal Fin. In French subject relatives and in extraction from embedded questions, Fin cannot be null due to the requirement that embedded finite clauses express the C-system in French. In order to comply with this requirement, the language always resorts to the overt nominal Fin in this context, and *qui* is always pronounced.

Even Standard French has an overt reflex of the *que-qui*-alternation in simple wh-questions: the complex wh *est-que* form, whatever its exact analysis, appears to be immune from "doubly filled C"-effects and manifests an alternation between *est-ce que* and *est-ce qui*:

(62) a. Quel garçon est-ce que tu as vu?
Which boy EST-CE QUE you have seen

b. Quel garçon est-ce qui est venu?
Which boy EST-CE QUI has come

So, the hypothesis that simple subject questions in standard French may involve a null-*qui*-strategy is made immediately plausible by a comparative analysis with regional varieties and related constructions, where an overt *qui* actually appears. The choice between the overt or null variant of Fin endowed with the required nominal qualities seems to be by and large determined by the status of the "doubly filled C"-constraint in the particular variety.

Going back to the derivation of local subject questions (59a,b), and directly extending the analysis to English, let us go through the relevant derivational steps, starting from the merger of Subj with the rest of the clause:

(63) Subj [. . . [wh_{subj} . . .]]

If, at this point, the wh-subject is internally merged (moved) into Spec/Subj, the Subject Criterion is satisfied but the element is frozen in place, hence it will be unable to reach its scope position in the C-system.

Suppose that normal Fin is directly merged with structure (63), without any movement to Spec/Subj:

(64) Fin [Subj [. . . [wh_{subj} . . .]]]

The Subject Criterion is evidently not satisfied in (64) and the structure is doomed. But we have another option, namely that of externally merging Fin endowed with the relevant nominal quality and unvalued Phi-features (henceforth Fin+Phi), with (63), yielding the following:

(65) Fin+Phi [Subj [. . . [wh_{subj} . . .]]]

The Subject Criterion is now satisfied by Fin+Phi, much as in the French *qui* case (remember that we have defined the criterial configuration in a way that encompasses local head-head relations, see (54).) The wh-element corresponding to the thematic subject is now free to move to its final scope position, presumably the Spec of the Focus-head endowed with Q in the left periphery. This movement takes place with an intermediate transit through Spec/Fin to value the Phi-features on this position, much as in the derivation of French sentences with overt *qui*. We thus end up with the following representation:

(66) Who Foc [t' Fin+Phi [Subj . . . [t . . .]]]

So, Fin+Phi offers a kind of bypassing device for the thematic subject, by satisfying the Subject Criterion in an interpretively vacuous manner (in fact

like an expletive), and allowing the thematic subject endowed with the wh- (or some other A'-)feature to move higher.

One significant consequence of this analysis is that simple subject questions involve direct movement from a lower, predicate internal position, even in languages like French, in a manner at least partially analogous to the one proposed for Null Subject languages over twenty years ago (modulo the unavailability of *pro* and the role of Fin in these languages).

It is natural to extend this analysis to English, even though the comparative evidence across dialects and related constructions which makes the proposal immediately plausible for French is not available. Indirect evidence for this analysis can be found for English as well, though. McCloskey 2000 provides an interesting argument in support of the view that in local subject questions in English, the wh-subject does not transit through its canonical IP-initial position. He studies a variety of regional English spoken in West Ulster in which wh-elements are allowed to launch floated quantifiers. These question formed in this manner differ from normal wh-questions in ". . . implicating that the answer is a plurality and in insisting on an exhaustive, rather than a partial, listing of the members of the answer set" (McCloskey. Op.cit., p. 58). The quantifier can be stranded in first (external) merge position, as in (67c), or in the position of an intermediate trace, as in (67b):

(67) a. What *all* did he say (that) he wanted?

b. What did he say *all* (that) he wanted?

c. What did he say (that) he wanted *all*?

(West Ulster English: McCloskey 2000)

Subject questions can also launch a floated quantifier, which may appear after the verb-object sequence, an order that McCloskey interprets as manifesting the external merge position of the subject, followed by scrambling to the left of the V-object constituent:

(68) Who was throwing stones *all* around Butchers' Gate?

What makes this option especially relevant to our discussion is that this Q-float structure is not possible with an ordinary, non wh-subject:

(69) * They were throwing stones *all* around Butchers' Gate

The ill-formedness of (69) is not particularly surprising; it is just a particular case of the general fact that a floated quantifier cannot be stranded in first-merge position, i.e., the quantifier cannot be stranded in object position with an unaccusative or passive structure:

(70) a. * The students left all

b. * The students have been contacted all (by the advisor)

Something like condition (71) seems to hold of quantifier float:

(71) In subject chains, floated Q's can't be stranded in first-merge (thematic) position.

The nature of this generalization is not clear, but the facts support the conclusion that a subject chain is not formed in the course of the derivation of (68). In other words, in the derivation of (68), *who* does not pass through the subject position occupied by *they* in (69). If it did, the two structures would be indistinguishable for the relevant part, and the contrast between them would remain unexplained. If, on the other hand, the subject wh does not transit through Spec/Subj, as we have been arguing, and is moved to the left periphery from a lower position, possibly from its first merge position, constraint (71) does not apply to case (68), and the contrast is captured.[9]

We should now make sure that the proposed 'bypassing'-system does not overgenerate. For instance, one should not be able to freely generate subjectless sentences by formally satisfying the Subject Criterion through Fin+Phi in the absence of movement:[10]

(72) * Bill said that seems that Mary is sick

So, the system should be able to rule out a configuration like the following:

(73) * Fin+Phi [Subj seems that]

[9]Holmberg & Hroarsdottir 2004 develop a different kind of argument in support of the view that subject wh-movement does not pass through the EPP position. Also, if Fitzpatrick (2005) is correct in arguing that Quantifier Float in Sportiche's sense is only possible under A'-movement then the contrast between (68) and (69) follows straightforwardly, provided that the chain link to Spec/Fin+Phi is both A and A', while the chain link to Spec/Subj is just A (see discussion after ex. (88)).

[10]The impossibility of a null-expletive structure of this sort should be evaluated in an embedded context (like the one in (72)) because registers of spoken English allow 'Root Expletive Drop' of the kind analyzed in Rizzi 2005a.

Clearly, the expletive-like function of Fin+Phi must be tied to subject extraction, a problem which we have already addressed in connection with French *qui*. We can envisage the same solution: the Phi-features in Fin are unvalued in Chomsky's (2001) sense, and valuation is achieved when the local subject moves to Spec/Fin, on its way to its final criterial landing site in the left periphery. Cases like (73) are therefore excluded because Phi in Fin remains unvalued in the absence of movement.

Consider now the case in which a wh-element different from the local subject moves to the left periphery of the main clause, e.g., an embedded clause subject or object. Why couldn't these elements pass through the matrix Spec/Fin+Phi and value Phi, thus yielding an ungrammatical subjectless sentence:

(74) a. * Who Foc t' Fin+Phi Subj seems (that) Bill met t?

b. * Who Foc t' Fin+Phi Subj seems (that) t met Bill?

Again, the point is analogous to the impossibility of licensing -i in French through any movement other than that of the local subject. If positions defined by agreement in Phi-features are A positions, Rizzi 1990, the chain link terminating in t' is a link of type A. (74a,b) are thus excluded by whatever principle proscribes the continuation of an A chain from a tensed complement:

(75) a. *John seems (that) Bill met t

b. * John seems (that) t met Bill

The 'expletive capacity' of Fin+Phi is thus made entirely contingent on movement of the subject, as desired.

8 Subject extraction across a null C and the **for*-trace effect.

Consider a successful case of subject extraction from an apparently C-less embedded clause in English and other Germanic languages:

(76) Who do you think came?

If we extend the 'silent' Fin+Phi idea discussed in connection with local subject movement to this case, (76) would have the following representation (irrelevant traces and other details omitted; see section 9 for further discussion on the complete C-structure in this case):

(77) Who do you think [t' Fin+Phi [Subj [t came]]]?

Suppose that Fin+Phi is merged directly after Subj is merged, formally satisfying the Subject Criterion. The wh-element corresponding to the thematic subject then moves to Spec/Fin+Phi to value the unvalued features on this position, and then it is extracted to the main clause. Criterial Freezing raises no problem here, as the element satisfying the Subject Criterion in the embedded clause is the Fin+Phi head, so that the thematic subjects remains available for extraction.

This approach is very close to the Agr in C approach of Rizzi (1990), according to which an Agr-(Phi-)morpheme "properly governs" the subject trace in the highest Spec position in the IP structure, satisfying the ECP. The two approaches differ in the presence vs. the absence of a subject trace in the clausal subject position (Spec/Subj, in current terms), and in the role of the device permitting subject extraction: In the old approach, the role of Agr in C was to provide a "proper governor" for the subject trace, so that the ECP could be satisfied; in the new approach, the role of Fin+Phi is to provide an expletive-like element to formally satisfy the Subject Criterion, so that the thematic subject can avoid the effect of Criterial Freezing.

A definite advantage to the new approach is that it immediately explains **for*-trace effects:

(78) * Who would you prefer [for [t_{who} Subj to win]]?

This case is notoriously problematic for an ECP approach, as the prepositional complementizer clearly governs the subject DP in a GB-type analysis (it licenses Case on the subject, it precludes *PRO*, etc.), so that the artificial distinction between 'government' and 'proper government' must be invoked. The analysis is rendered even more problematic by the fact that the minimally different preposition for normally allows preposition stranding, hence must function as a "proper governor": *Who did you work for t?*

The Fin+Phi approach fares better here. The complementizer for plausibly retains the categorial status of a preposition, hence it lacks the nominal featural endowment which qualifies an element to be a candidate for satisfaction of the Subject Criterion. The only way to satisfy the criterion is then to move the thematic subject to Spec/Subj, where it gets frozen, whence the impossibility of (78). Extraction of the complement of the preposition for is not problematic because no Subject (or other) Criterion is involved.

One may object that P seems to be able to carry Phi-features in some languages (see, e.g., agreeing prepositions in Celtic, McCloskey & Hale

1983,) so one could imagine a *for*+Phi able to satisfy the Subject Criterion, much as Fin+Phi in (77). Why is this option excluded? The requirement that the element satisfying the subject criterion be endowed with Phi is, in a sense, a corollary of the substantive requirement that the subject be a nominal element. Remember that we have assumed that the element satisfying the Criterion must be [+N]. Fin may normally satisfy this requirement, but the particular realization of Fin as *for* cannot if the complementizer *for* retains the specification [-N] of the preposition *for* (this is plausible, as it retains the Case-assigning capacity of the preposition). As a [-N]-element, *for* in Fin cannot function as a quasi-expletive in the proposed way, hence it does not help to satisfy the Subject Criterion in (78), and subject extraction is impossible.

9 *That*-trace

Consider now a standard *that*-trace effect. The empirical assumption made by the traditional analysis is that the *that*-trace configuration should be banned in general. Nevertheless, Sobin 1987, 2002, among others, argues that the ban against this structure is dialect-specific, and provides evidence to the effect that subject extraction over that is acceptable in some varieties of English.

(79) % Who did you say that t_{who} came?

From this viewpoint, the *that*-trace configuration in English is analogous to the *qui*-trace configuration in cases of embedded subject extraction in French ((31b), etc.), which also appears to manifest variable acceptability across dialects. The analogy is further strengthened by the fact that just as *qui* is invariably acceptable in subject relatives in French (*l'homme qui est venu*), *that* is invariably acceptable in subject relatives in English.

(80) The man that came

It seems, therefore, that the invariably acceptable strategy in subject relatives (*qui, that*) may, with dialectal variability, be extended to subject extraction from embedded declaratives.

Starting from the restrictive variety of English which excludes (79), we must capture the fact that the overt complementizer *that* is incompatible with the Fin+Phi strategy. Consider the following possibility. *That* expresses both finiteness (it is incompatible with non-finite IP's) and (declarative) Force.

(81) Force Fin IP

So if the complete CP-system involves the structure, the normal derivation of a *that*-clause is one in which *that* is first merged in Fin, to express finiteness, and then moves to Force to check the Force feature (the result, from our perspective, would be the same if that expressed both properties syncretically, as in Rizzi 1997, Shlonsky 2006). So, we end up with a representation like the following:

(82)

Force	Fin	IP
|	|	
that	t_{that}	

Under the natural assumption that expletive and argument functions cannot be performed by the same element, *that* in (82) cannot simultaneously be the head of the declarative - a clausal argument - and function as an expletive-like surrogate subject to formally satisfy the Subject Criterion.

What about the varieties in which (79) is possible? Pursuing the analogy with French, we entertain the hypothesis that these varieties may separate Force and Fin, much as French *que* and *-i*: Fin is expressed by our quasi expletive nominal filler endowed with unvalued Phi-features (always silent in English), and *that* is merged higher up (expressing force and presumably also carrying a finiteness feature which is checked under Agree with Fin):

(83)

Force	Fin	IP
|	|	
that	Phi	

Here, the functions of head of the clausal argument and of quasi-expletive are performed by separate elements, and no conflict arises.

Such a system is invariably available for subject relatives, presumably because of the functional need to have a device to express relatives on all major argument positions, subjects in the first place. What is dialect-specific in both English and French is the possibility of extending the strategy involving a more complex C-system from the domain of subject relatives to that of subject extraction from embedded declaratives.

We can now go back to the structure involving subject extraction with a null C in English, (77), and sharpen the analysis proposed in section 8. Since such sentences are uniformly acceptable across dialects, it is unlikely that they may simply involve an unpronounced variant of the complex-C strategy of (83). If (77) were modeled on (83) - modulo a silent *that* - we would expect a variability of judgments, parallel to what Sobin describes for (79).

A more promising possibility is that finite declaratives lacking an overt C in English do not involve a complete Force-Fin structure, but a reduced, or 'truncated' one. Suppose that there is no structurally expressed Force head in structures with so called 'C-deletion', and that the declarative interpretation is assigned by default, much as in ECM-clauses, in which the C-system is radically absent and the embedded clause is interpreted as a declarative. So, if (77) is truncated at Fin, no conflict arises in the role(s) of Fin, and the quasi-expletive nominal Fin strategy can be deployed, as discussed in section 8.

Treating C-less sentences as truncated structures raises another more radical possibility. One could envisage that sentences like *Who do you think came?* involve a deeper truncation, affecting not only the whole CP-structure but also the SubjP-layer, hence closing off IP with the AgrP-projection responsible for the Case-agreement system. This would immediately predict the absence of any criterial freezing effect in such truncated structures, due to the radical absence of the criterial SubjP-layer, with no need for the quasi-expletive skipping device.

We believe this 'deep truncation' approach may, indeed, be the appropriate analysis of another major case of absence of freezing effects. In subject-raising environments, the embedded subject is able and obliged to move to the main subject position.

(84) Mary seems [t to be happy]

It is well known that Raising is incompatible with any C-structure (i.e., we never find Raising out of infinitival indirect questions), an observation which led to the traditional "S'-deletion" analysis in the GB framework. We can adopt this analysis, with the additional proviso that the truncation process in Raising-infinitives includes the embedded SubjP-layer, so that Raising can apply without the need for any special skipping device.

But deep truncation may be too radical for wh-extraction from finite clauses like (76). At least a minimal vestige of the C-system may be required in finite clauses to permit a proper temporal interpretation, with the speech time somehow structurally expressed in C, as many have suggested (see Bianchi 2003 for recent discussion). Potential evidence for restricted truncation in cases of English C-deletion is provided by the well-known observation that C-deletion is incompatible with the activation of the left periphery of the clause for a topic and focus.[11]

[11] Sobin (2002) observes that embedded declaratives with a null complementizer and a preposed adverbial are quite acceptable for many speakers. This is not surprising if preposed

(85) a. She thought *(that) this book, you should read

b. She thought *(that) never in her life would she accept this solution

(adapted from Grimshaw 1997)

If (76) involves a defective C-system which expresses finiteness, and if the expression of the argumental status of the clause is a prerogative of the Force specification, as seems natural (if the clause is a declarative, a question or an exclamative is the crucial information that a higher selector looks for,) no conflict of function arises for the Fin-layer of the defective C in (77), which is allowed to function as a subject surrogate, as we have proposed.

Consider now the systematic non-extractability of subjects across the C-system of indirect questions:

(86) a. * The man who I wonder if t_{who} will leave

b. * The man who I wonder when t_{who} will leave t_{when}

c. * The man who I wonder what t_{who} will say t_{what}

Here, according to Sobin's variation study, we observe no dialect split, subject-extraction being uniformly excluded (and judged worse than the variably degraded object-extraction in this environment). Why don't we seem to find dialects of English using a strategy which would allow Fin to satisfy the Subj Criterion in indirect questions?

Notice that the interrogative interpretation cannot be assigned by default, and requires a structurally specified head expressing interrogative force, i.e., there are no ECM indirect questions. So, a 'truncation' approach could not help here.

Given the availability of merging Fin and Force separately, as in the varieties of French and English which allow subject-extraction over a declarative *que* or *that*, one wonders why this option cannot be extended to indirect non-subject interrogatives.

An extension of the complex (or split) Comp-system of (83) to interrogatives would involve the merge of a nominal Fin, able to function as a quasi expletive, followed by the external merger of a higher head in the C-system expressing the interrogative character of the clause. Perhaps this is not possible. More precisely, perhaps Fin in interrogatives may be able to function

adverbials may be part of the IP space and thus compatible with a truncated Comp-system (see the discussion at the end of section 6).

as an expletive-like element in a highly selective manner, i.e., only when the subject is locally moved, under the analysis of local movement developed for indirect subject-questions in section 7 and schematized in (87).

(87) I wonder [who Foc [Fin Subj [t will leave]]]

Presumably, this strategy is not available for subject-extraction from an indirect question, when the locally moved phrase is distinct from the subject:

(88) * Who do you wonder [when Foc [Fin Subj [t_{who} will leave t_{when}]]]

It is plausible that Fin in indirect questions carries in some form the specification of its clausal type. For instance, Fin in embedded questions in German is unable to attract the inflected verb, and V2 is systematically banned in this environment (but not in embedded declaratives). Suppose that this "memory" of the clausal type is expressed as the possession of a formal *q*-feature, which characterizes Fin as "Fin of a question", and must enter into an agree relation with the wh-element. In cases of local subject-movement like (87), Fin will carry this *q*-feature, and also Phi features, according to our analysis in section 7. Fin+Phi+q can produce a well-formed structure in (87), where the local subject is moved, and both featural requirements are satisfied by the same element. Fin formally satisfies the Subject Criterion, permitting the thematic subject to be moved, and attracts the thematic wh-subject to its Spec, where the subject satisfies both its *q*- and Phi-features (with the relevant chain link to Spec/Fin+Phi+q being both A and A', see note 10.)

Consider now (88). In order to allow subject-extraction, Fin should also carry both specifications Phi and *q*; but here, Fin+Phi+q would have to have its featural requirements satisfied by distinct elements, Phi by the subject and *q* by the locally moved wh-phrase (*when* here). This would not be possible under a plausible uniqueness assumption: A single head can have its featural requirements satisfied by a single phrase, not by two separate phrases. For example, a Phi-feature set on a single head could not agree in person with one nominal and in number with another nominal. If this is a general property of feature checking, then a configuration like (88) is banned in principle.[12]

[12]The more uniform ban on the wh - trace configuration as compared with the *that* - trace configuration is confirmed cross-linguistically. For example, Shlonsky 1988, 1990 observed that Modern Hebrew permits the latter (see (12) above), presumably derivable through the mechanism discussed in connection with (79), but not the former. Nonetheless, the possibility of a well-formed structure corresponding to (88) may still arise in some language at the price of further complicating the C-system, e.g. of having the *q* and the Phi specifications on separate

10 Conclusion

If the EPP is to be restated as a Subject Criterion, the difficulty of moving subjects can be ascribed to Criterial Freezing, a principle which interrupts a movement-chain as soon as a position dedicated to a scope-discourse property is reached. Subjects are more difficult to move than objects or other arguments, but are not unmovable: languages invent strategies which make subject movement possible at the price of introducing special formal devices to circumvent the freezing effect. Some such strategies consist in acknowledging the immovability of subjects and forming A'-chains on subjects through resumption, or pied-piping of the whole embedded clause. Other languages use special devices, expletives of various sorts, to formally fulfill the Subject Criterion, thus allowing the thematic subject to skip the freezing position and be moved from a lower, predicate internal position, a legitimate extraction site. Some languages systematically use regular, bona fide expletives in this function, as is the case for Null Subject Languages, according to a traditional analysis. We have proposed that other language specific devices, traditionally analyzed in different terms, are essentially reducible to variants of the same skipping strategy.

Some of the strategies of subject extraction come for free, given the general parametric properties of the languages. This is the case, for instance, of Null Subject Languages, in which the independent availability of the null expletive offers a systematic "skipping" device. In such cases, as is expected, we do not observe variation, i.e. we do not seem to find a dialectal variety of Italian or Spanish manifesting the *that*-t effect. Other strategies involve special devices whose purpose seems to be limited to permit subject movement in particular environments. Here we expect, and find, variation also in closely related languages and dialects, variation having to do with the existence and scope of the special device: French *que-qui* is a case in point.

A basic tenet of our analysis, which we have directly adopted from the classical ECP approach, is that we need a strong, cross-linguistically uniform explanation of the difficulty of subject extraction. According to this line of analysis, the variation does not result from a parametrisation of the relevant principle, but rather from the different language-particular devices used to circumvent a general prohibition. This indirect approach seems more restrictive,

---

heads. This may be the strategy used in the Scandinavian varieties in which the equivalent of (88) appears to be possible. We will not address the relevant mechanism in this paper, nor other special strategies of subject extraction like the 'anti-adjacency' or 'adverb effects, on which see Culicover 1992, Browning 1996, Rizzi 1997, Sobin 2002.

and better suited to predict the observed patterns of invariance and variation, than one which would directly weaken the prohibition on language extraction by making it a language specific property.

References

Aboh, Enoch
2004 *The morphosyntax of complement-head sequences: Clause structure and word order patterns in Kwa.* Oxford University Press, New York

Baltin, Mark and Chris Collins (eds.)
2001 *The handbook of contemporary syntactic theory.* Blackwell, Oxford.

Barlow, Michael and Charles A. Ferguson (eds.)
1988 *Agreement in natural language: Approaches, theories, descriptions.* CSLI Publications, Stanford CA.

Bates, Dawn (ed.)
1992 *The Proceedings of the Tenth West Coast Conference on Formal Linguistics.* CSLI Publications, Stanford CA.

Batllori, Montserrat et al. (eds.)
2005 *Grammaticalization and parametric variation.* Oxford University Press, Oxford.

Belletti, Adriana
2001 'Inversion' as focalization. In Hulk & Pollock (2001), pp. 60-90.
2004 Aspects of the low IP area. In Rizzi (2004), pp. 16-51.

Belletti, Adriana (ed.)
2004 *Structures and beyond.* Oxford University Press, Oxford.

Belletti, Adriana and Luigi Rizzi (eds.)
1996 *Parameters and functional heads.* Oxford University Press, Oxford.

Benincà, Paola (ed.)
1989 *Dialect variation and the theory of grammar.* Foris, Dordrecht.

Benincà, Paola
2001 The position of Topic and Focus in the left periphery. In Cinque & Salvi (2001), pp. 39-64.

Benincà, Paola and Cecilia Poletto
2004 Topic, Focus and V2: Defining the CP sublayers. In Rizzi (2004), pp. 52-75.

Bennis, Hans, Martin Everaert & Eric Reuland (eds.)
2000 *Interface strategies.* Royal Netherlands Academy of Arts and Sciences, Amsterdam.

Bennis, Hans and Liliane Haegeman
1984 On the status of agreement and relative clauses in West Flemish. In Geest & Putseys (1984), pp. 33-53.

Bianchi, Valentina
1999 *Consequences of antisymmetry: Headed relative clauses.* Mouton de Gruyter, Berlin.
2003 On finiteness and logophoric anchoring. In Guerón & Tamowski (2003), pp. 213-246.

Boeckx, Cedric
2003 *Islands and chains. Resumption as stranding.* John Benjamins, Amsterdam.
Borer, Hagit
1984 Restrictive relatives in Modern Hebrew. *Natural Language & Linguistic Theory*, 2: 219-260.
Browning, Maggie
1996 CP Recursion and *that*-t Effects. *Linguistic Inquiry*, 27: 237-255.
Brugé, Laura et al. (eds.)
2005 *Proceedings of the XXX Incontro di Grammatica Generativa.* Cafoscarina, Venezia.
Cardinaletti, Anna
2004 Towards a cartography of subject positions. In Rizzi (2004), pp. 115-165.
Cardinaletti, Anna and Michal Starke
1999 The typology of structural deficiency: On the three grammatical classes. In van Riemsdijk (1999), pp. 145-233.
Carstens, Vicki
2003 Rethinking Complementizer Agreement: Agree with a Case-Checked Goal. *Linguistic Inquiry*, 34: 393-412.
Chao, Wynn
1981 PRO-drop languages and nonobligatory control. In Chao & Wheeler (1981)
Chao, Wynn and Deirdre Wheeler (eds.)
1981 *University of Massachusetts Occasional Papers in Linguistics 7.*
Cheng, Lisa and Luigi Corver (eds.)
2006 *Wh-Movement: Moving On.* MIT Press, Cambridge MA.
Chomsky, Noam
2000 Minimalist inquiries. In Martin, Michaels & Uriagereka (2000), pp. 89-155
2001 Derivation by phase. In Kenstowicz (2001), pp. 1-52.
Cinque, Guglielmo
1999 *Adverbs and functional heads: A cross-linguistic perspective.* Oxford University Press, Oxford.
Cinque, Guglielmo and Giampaolo Salvi (eds.)
2001 *Current studies in Italian syntax.* Elsevier, Amsterdam.
Cole, Peter
1985 *Imbabura Quechua.* Croom Helm, London.
Cole, Peter and Gabriella Hermon.
1981 Subjecthood and Islandhood: Evidence from Quechua. *Linguistic Inquiry*, 12: 1-30.
Culicover, Peter W.
1992 Topicalization, Inversion and Complementizers in English. In Delfitto et al. (1992), pp. 1-43.

Deen, Kamil et al. (eds.)
2005 *Proceedings of the inaugural conference on generative approaches to language acquisition - North America (GALANA).* UCONN / MIT Working Papers in Linguistics, Cambridge MA.

Delfitto, Denis et al. (eds.)
1992 *Going Romance and Beyond.* OTS Occasional Publications Series, Utrecht.

Doron, Edit
1988 On the complementarity of subject and subject-verb agreement. In Barlow & Ferguson (1988), pp. 201-218.

Fitzpatrick, Justin
2006 Two Types of Floating Quantifiers and their A/A'-Properties. *Proceedings of NELS*, 36: 253-265.

Frascarelli, Mara (ed.)
2006 *Phases of interpretation.* Mouton de Gruyter, Berlin.

Grimshaw, Jane
1997 Projection, Heads, and Optimality. *Linguistic Inquiry*, 28: 373-422.

Guerón, Jacqueline and Tamowski, L. (eds.)
2003 *Temps et Point de Vue / Tense and Point of View.* Université de Paris X, Paris.

Geest, Wim de and Putseys, Yvan (eds.)
1984 *Sentential complementation.* Foris, Dordrecht.

Haegeman, Liliane
1990 Subject pronouns and subject clitics in West Flemish. *The Linguistic Review*, 7: 333-363.
1992 *Theory and description in generative grammar, a case study of West Flemish.* Cambridge University Press, Cambridge.

Haegeman, Liliane (ed.)
1997 *Elements of grammar: A handbook of generative syntax.* Kluwer, Dordrecht.

Hermon, Gabriela
1984 *Syntactic modularity.* Foris, Dordrecht.

Holmberg, Anders and Thorbjörg Hróarsdóttir
2004 Agreement and Movement in Icelandic Raising Constructions. *Lingua*, 114: 651-673.

Hulk, Aafke and Jean-Yves Pollock (eds.)
2001 *Inversion in Romance and the theory of universal grammar.* Oxford University Press, Oxford.

Kayne, Richard S.
1989 Facets of Romance past participle agreement. In Benincà (1989), pp. 85-103.
1994 *The antisymmetry of syntax.*. MIT Press, Cambridge MA.

Kenstowicz, Michael (ed.)
2001 *Ken Hale: A life in language*. MIT Press, Cambridge MA.
Kishimoto, Hideki
2005 Wh-in-situ and movement in Sinhala questions. *Natural Language & Linguistic Theory*, 23: 1-51.
Landau, Idan
to appear EPP extensions. *Linguistic Inquiry*
Laenzlinger, Christopher
1997 *Comparative studies in word order variations: pronouns, adverbs and German clause structure*. John Benjamins, Amsterdam.
Laenzlinger, Christopher and Ur Shlonsky
1997 Weak Pronouns as LF Clitics: Clustering and Adjacency Effects in the Pronominal Systems of German and Hebrew. *Studia Linguistica*, 51: 154-185.
Martin, Roger, David Michaels and Juan Uriagereka (eds.)
2000 *Step by step: Minimalist syntax in honor of Howard Lasnik*. MIT Press, Cambridge MA.
McCloskey, James
2000 Quantifier Float and Wh-Movement in an Irish English. *Linguistic Inquiry*, 31: 57-84.
McCloskey, James and Kenneth Hale
1983 On the Syntax of Person-Number Inflection in Modern Irish. *Natural Language & Linguistic Theory*, 1: 487-533.
McDaniel, Dana
1989 Partial and Multiple Wh-Movement. *Natural Language & Linguistic Theory*, 7: 565-604.
Menuzzi, Sergio
2000 *That*-trace effects in Portuguese. *Forum Lingüístico*, 2: 13-39.
Nicolis, Marco
2005 *On pro drop*. Ph.D. thesis, University of Siena.
Ortiz de Urbina, Jon
1989 *Parameters in the grammar of Basque*. Foris, Dordrecht.
Pesetsky, David
1982 Complementizer-trace Phenomena and the Nominative Island Condition. *The Linguistic Review*, 1: 297-343.
Poletto, Cecilia
2005 *Sì* and *e* as CP expletives in Old Italian. In Batllori et al. (2005), pp. 206-235.
Postal, Paul M.
1972 On some rules that are not successive cyclic. *Linguistic Inquiry*, 3: 211-222.
Richards, Norvin
2000 An island effect in Japanese. *Journal of East Asian Linguistics*, 9: 187-205.

2001 *Movement in language. Interactions and architectures.*. Oxford University Press, Oxford.

Rizzi, Luigi

1982 *Issues in Italian Syntax.* Foris, Dordrecht.

1990 *Relativized Minimality.* MIT Press, Cambridge MA.

1996 Residual verb second and the Wh-Criterion. In Belletti & Rizzi (1996), pp. 63-90.

1997 The fine structure of the left periphery. In Haegeman (1997), pp. 281-337.

2001 Relativized minimality effects. In Baltin & Collins (2001), pp. 89-110.

2004 Locality and left periphery. In Belletti (2004), pp. 223-251.

2005a Grammatically-based target-inconsistencies in child language. In Deen et al. (2005), pp. 19-49.

2005b On some properties of subjects and topics. In Brugé et al. (2005)

2006 On the Form of Chains: Criterial Positions and ECP Effects. In Cheng & Rizzi (2006), pp.97-134.

Rizzi, Luigi (ed.)

2004 *The Structure of CP and IP.* Oxford University Press, Oxford.

Rizzi, Luigi and Ur Shlonsky

2006 Satisfying the Subject Criterion by a Non Subject: English Locative Inversion and Heavy NP Shift. In Frascarelli (2006), pp. 341-361.

Roussou, Anna

2002 C, T, and the Subject: That-t Phenomena Revisited. *Lingua*, 112: 13-52.

Shlonsky, Ur

1988 Complementizer-Cliticization in Hebrew and the Empty Category Principle. *Natural Language & Linguistic Theory*, 6: 191-205.

1990 Pro in Hebrew Subject Inversion. *Linguistic Inquiry*, 21: 263-275.

1997 *Clause structure and word order in Hebrew and Arabic: An essay in comparative Semitic syntax.*. Oxford University Press, Oxford.

2000 Subject positions and copular constructions. In Bennis, Everaert & Reuland (2000), pp. 325-347.

2006 Projection étendue et cartographie de SC. *Nouveaux cahiers de linguistique franaise*, 2: 83-93.

Shlonsky, Ur and Edit Doron

1992 Verb-Second in Hebrew. In Bates (1992), pp. 431-446.

Sobin, Nicholas

1987 The Variable Status of Comp-Trace Phenomena. *Natural Language & Linguistic Theory*, 5: 33-60.

2002 The Comp-trace effect, the adverb effect and minimal CP. *Journal of Linguistics*, 38: 527-560.

Sportiche, Dominique
1988 A Theory of Floating Quantifiers and Its Corollaries for Constituent Structure. *Linguistic Inquiry*, 19: 425-449.
Taraldsen, Knut T.
1978 *On the NIC, Vacuous Application and the That-Trace Filter*. Indiana University Linguistics Club, Bloomington.
2001 *Subject extraction, the distribution of expletives and stylistic inversion.*. In Hulk & Pollock (2001), pp. 163-182.
Toman, Jindrich (ed.)
1984 *Studies in German grammar*. Foris, Dordrecht.
van Riemsdijk, Henk
1984 *On pied-piped infinitives in German relative clauses.*. In Toman (1984), pp. 165-192.
van Riemsdijk, Henk (ed.)
1999 *Clitics in the Languages of Europe*. Mouton de Gruyter, Berlin.
Vergnaud, Jean-Roger
1974 *French relative clauses*. Ph.D. thesis, MIT.

Some Remarks on Locality Conditions and Minimalist Grammars*

Hans-Martin Gärtner and Jens Michaelis

Introduction

In this paper we undertake a study of syntactic *locality conditions (LCs)* within Stablerian *minimalist grammars (MGs)* (Stabler 1997, 1998, 1999 and elsewhere). We show that the "restrictiveness" of LCs measured in terms of weak generative capacity depends on how they are combined. Thus, standard MGs incorporating just the *shortest move condition (SMC)* are mildly context-sensitive. Adding the *specifier island condition (SPIC)* to such grammars either reduces complexity or, interestingly, it increases complexity. This depends on the co-presence or absence of the SMC, respectively. Likewise, the effect of adding the *adjunct island condition (AIC)* to an extended MG is either trivial (without co-presence of the SMC) or, apparently, crucial in preserving mild context-sensitivity. The point of this exercise is to demonstrate that LCs as such - intuitions to the contrary notwithstanding - are not automatically restrictive where a formal notion of restrictiveness is applied. Independent motivation for our work comes from a recent convergence of two research trends. On the one hand, appeal has been made to the formal complexity of natural languages in work on language evolution (Hauser et al. 2002, Piattelli-Palmarini and Uriagereka 2004) and to *computational efficiency* in mainstream minimalism (Chomsky 2005). On the other hand, the formally well-understood Stablerian MGs provide enough descriptive flexibility to be taken seriously as a syntactic theory by the working linguist. A more comprehensive study of the complexity of constraint interaction is still outstanding.

*For comments, criticisms and suggestions we are greatful to audiences at FG 2003 Vienna, ZAS Berlin, Symposium "Recursion + Interfaces = Language?," LACL 2005 Bordeaux, FGMoL 2005 Edinburgh, INRIA/LaBRI Bordeaux, SfS Tübingen. Special thanks go to an anonymous reviewer. Finally we would like to thank Tom Cornell, Ed Keenan, Greg Kobele, Hap Kolb, Marcus Kracht, Uwe Mönnich, Christian Retoré, Jim Rogers, Ed Stabler, Peter Staudacher, Wolfgang Sternefeld, Craig Thiersch for many hours of stimulating discussion. The usual disclaimers apply.

1 Locality Conditions

Generative grammar took one of its more important turns when *locality conditions (LCs)* were established in work by Ross (1967) and Chomsky (1973, 1977). As is well-known, this led to a period of intense research into the proper formulation of LCs, as documented in work by Huang (1982), Chomsky (1986), Rizzi (1990), Cinque (1990), Manzini (1992), Müller and Sternefeld (1993), and Szabolcsi and Zwarts (1997), among others.

Formally LCs can be separated into two types, *intervention-based LCs (ILCs)* and *containment-based LCs (CLCs)*. ILCs are often characterized in terms of minimality constraints, such as the minimal link, minimal chain, shortest move, or attract closest condition. In the framework of *minimalist grammars (MGs)* (Stabler 1997, 1999), which we are adopting in this paper, intervention-based locality is captured by the shortest move condition (SMC). CLCs are often characterized in terms of (generalized) grammatical functions. Familiar conditions define adjunct islands, subject islands, and specifier islands. MGs have integrated versions of a *specifier island condition (SPIC)* (Stabler 1999) and an *adjunct island condition (AIC)* (Frey and Gärtner 2002; Gärtner and Michaelis 2003). In (1) we schematically illustrate the structure of these LC-types.

(1) a. $[\ldots\alpha\ldots[\ldots\beta\ldots\gamma\ldots]]$

b. $[\ldots\alpha\ldots[_{\beta}\ \ldots\ \gamma\ \ldots\ ]]$

An ILC, (1a), prevents establishing dependencies between constituents α and γ across an intervening β. Intervention is typically defined in terms of c-command or similar notions. A CLC, (1b), on the other hand, prevents establishing dependencies between constituents α and γ into or out of a containing β. Containment is usually defined in terms of dominance.

It is also well-known that LCs have been central in the quest for achieving the "Goals of Generative Linguistic Theory." Thus, consider the following statement by Chomsky (1973, p. 232):

> From the point of view that I adopt here, the fundamental empirical problem of linguistics is to explain how a person can acquire knowledge of language. [...] To approach the fundamental empirical problem, we attempt to restrict the class of potential human languages by setting various conditions on the form and function of grammars.

Quite uncontroversially, LCs have been taken to serve as restrictions in this sense. However, the important underlying notion of restrictiveness is much

less easy to pin down in a principled manner. In particular, it is difficult to answer the following two questions in any satisfactory way.

Q1: How do we know that we have restricted the class of potential human languages?

Q2: Could we measure the degree of restriction, and if so, how?

Researchers are fundamentally divided over how to deal with these questions. Currently, at least two major approaches coexist. The first one, which we will call "formalist," is rooted in formal complexity theory as discussed in Chomsky (1956, 1959). The alternative, which we call "cognitivist," is built on the prospects of establishing a theory of "relevant cognitive complexity." For this distinction we rely on Berwick and Weinberg (1982, p. 187), who emphasized that "[t]here is a distinction to be drawn between *relevant cognitive complexity* and the *mathematical complexity* of a language."

Interestingly, Chomsky (1977) may be understood as having sided with the cognitivists, interpreting locality conditions as part of such a theory, as the following quote indicates.[1]

> Each of these conditions [subjacency, SSC, PIC] may be thought of as a limitation on the scope of the processes of mental computation. (Chomsky 1977, p. 111)

Now, a standard criticism raised by formalists against cognitivists concerns the inability of the latter of answering questions Q1 and Q2. In particular, cognitivist notions of restrictiveness have been found inadequate for defining classes of languages. Formalism, on the other hand, is typically criticized especially for employing the measure of weak generative capacity, which, it is felt, requires abstractions too far removed from the grammars found useful by linguists.

However, recent developments, taking their outset from "The Minimalist Program" (Chomsky 1995) have created a situation where formalism and cognitivism have begun to converge on common interests again.

In particular, work on language evolution by, i.a., Hauser et al. (2002) and Piattelli-Palmarini and Uriagereka (2004) has raised the interest of cognitivists in formalist concerns.[2]

---

[1] *SSC* stands for the *specified subject condition*, and *PIC* stands for the *propositional island condition*.

[2] According to Kolb (1997, p. 3) the same trend toward formalism characterizes Chomsky's minimalist revision of *principles and parameters (PP) theory*: "PP theory often gives the im-

At the same time, work on minimalist grammars (MGs), as defined by Stabler (1997), has led to a realignment of "grammars found 'useful' by linguists" and formal complexity theory. MGs are capable of integrating (if needed) mechanisms such as: head movement (Stabler 1997, 2001), (strict) remnant movement (Stabler 1997, 1999), affix hopping (Stabler 2001), adjunction and scrambling (Frey and Gärtner 2002), and late adjunction and extraposition (Gärtner and Michaelis 2003).

In addition to this descriptive flexibility, Michaelis (1998 [2001a]) has shown MGs to provide a *mildly context-sensitive grammar (MCSG)* formalism in the sense of Joshi (1985).[3] This class of formalisms, which is shown in Fig. 15 (Appendix C), has repeatedly been argued to be of exactly the right kind when it comes to characterizing the complexity of human languages. MCSGs combine conditions on weak generative capacity with the condition of polynomial time parsability[4] and the so-called *constant growth* property. Constant growth informally means that "if the strings of a language are arranged in increasing order of length, then two consecutive lengths do not differ in arbitrarily large amounts" (Joshi et al. 1991, p. 32).

Given the two properties just outlined, MGs are an ideal tool for studying the complexity and/or restrictiveness of LCs. Such a study is what the remainder of this paper is devoted to. Concretely we are going to look at the behavior and interaction of the SMC, the SPIC and the AIC. It will turn out that different LCs have different effects on complexity. The original complexity result has been shown to hold for standard MGs incorporating the SMC. Now, importantly, adding the SPIC to standard MGs has non-monotonic consequences in the sense that whether complexity goes up or down depends on the absence or (co-)presence of the SMC, respectively. Thus, if we interpret (and measure) growing restrictiveness in terms of complexity reduction, we must conclude that adding a constraint like the SPIC as such does not - intuitions to the contrary notwithstanding - lead to more restrictive grammars automatically.

---

pression of a mere collection of 'interesting' facts which is largely data driven and where every new phenomenon may lead to new (ad hoc) formal devices, often incompatible, and without a measure to compare and/or decide between conflicting analyses meaningfully—in short: As a formal system it looks largely incoherent. [...] In what amounts to just about a U-turn, [in] its latest version, chapter 4 of Chomsky (1995) [...] [c]omplexity considerations are reintroduced into theory formation, and the non-recursiveness assumption is (implicitly) retracted." The trend has gained full momentum in Chomsky's more recent writings, where *computational efficiency* is counted among the crucial (sub-)factors of *language design* (Chomsky 2005, p. 6).

[3]See also Michaelis (2001b, 2005) and references cited therein for further details.

[4]This is the dimension that underlies the formal study of island conditions in Berwick (1992). For psycholinguistic studies, see Pritchett (1992) and Gibson (1991).

For the AIC, the picture is slightly more complicated. First of all, the AIC only makes sense if (base-)adjunction and adjunction by scrambling/extraposition is added to MGs. Even more specifically, the AIC seems to make a difference if adjunction is allowed to occur countercyclically or *late*, i.e. if it is allowed to target a non-root constituent. Under these conditions, adding the AIC together with the SMC guarantees that the resulting grammars stay within the class of MCSGs. Without the AIC there are configurations that appear to go beyond that boundary. In MGs without the SMC, on the other hand, it is plausible to assume that the AIC does not change complexity at all, i.e. it is void. Again we can conclude that the restrictiveness of a constraint is not inherently given but depends on the structure it interacts with.

Before we can present these results, we give a brief introduction to standard MGs and the relevant extensions. This will be done in Section 2. Section 3 contains our main results. Section 3.1 illustrates how an MG including the SPIC but without the SMC goes beyond MCSGs. Section 3.2 shows a case where an MG with the SMC but without the AIC appears to lose its status as MCSG. Section 4 is devoted to conclusions and a further outlook. Appendix A provides formal definitions and Appendix B sketches our approach to multiple wh-movement. We show there how to remove a prima facie conflict between this phenomenon and the SMC.

2 Minimalist Grammars

The objects specified by a *minimalist grammar (MG)* are so-called *minimalist expressions* or *minimalist trees*, which straightforwardly translate into the usual aboreal picture from syntactic theory as depicted in Fig. 1.[5]

A *simple expression* is given as a list of feature instances (technically: a single-noded tree labeled by that list) to be checked from left to right, where the intervening marker # is used to separate the checked part of feature instances from the non-checked one. A minimalist tree is said to *have*, or likewise, *display feature* f if its head-label is of the form $\alpha \# f \alpha'$.

Starting from a finite set of simple expressions (the *lexicon*), minimalist expressions can be built up recursively from others by applying structure building functions. The applications of these functions are triggered by particular instances of syntactic features appearing in the trees to which the functions are applied. After having applied a structure building function, the triggering

[5] Stabler's minimalist expressions are closely related to but not to be confused with Chomskyan *linguistic* expressions, the latter defined as pairs, $\langle \pi, \lambda \rangle$, of PF- and LF-representations (Chomsky 1995, p. 170).

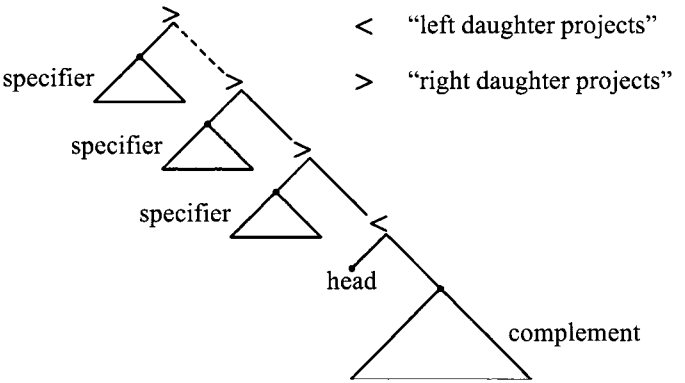


Figure 1. A typical minimalist expression

feature instances get marked as checked. Different structure building operations are triggered by different types of syntactic features. The standard ones are given by the following list:

(basic) categories:	`x, y, z, ...`
m(erge)-selectors:	`=x, =y, =z, ...`
m(ove)-licensees:	`-x, -y, -z, ...`
m(ove)-licensors:	`+x, +y, +z, ...`

Instances of (basic) category features and m-selectors trigger the *merge*-operator mapping a pair of trees to a single tree if the selecting tree ϕ displays m-selector `=x` and the selected tree χ displays the corresponding category `x`. χ is selected as a complement in case ϕ is simple, and as a specifier, otherwise. In both cases, the triggering feature instances get marked as checked in the resulting tree (see Fig. 2).

Instances of m-licensors and m-licensees trigger applications of the *move*-operator by which—without imposing the *shortest move condition (SMC)*—a single tree displaying m-licensor `+x` is mapped to a finite set of trees, consisting of every tree which results from moving a maximal projection displaying the corresponding m-licensee `-x` into a specifier position. Again the feature instances triggering the application of the operator get marked as checked in the resulting tree (see Fig. 3).

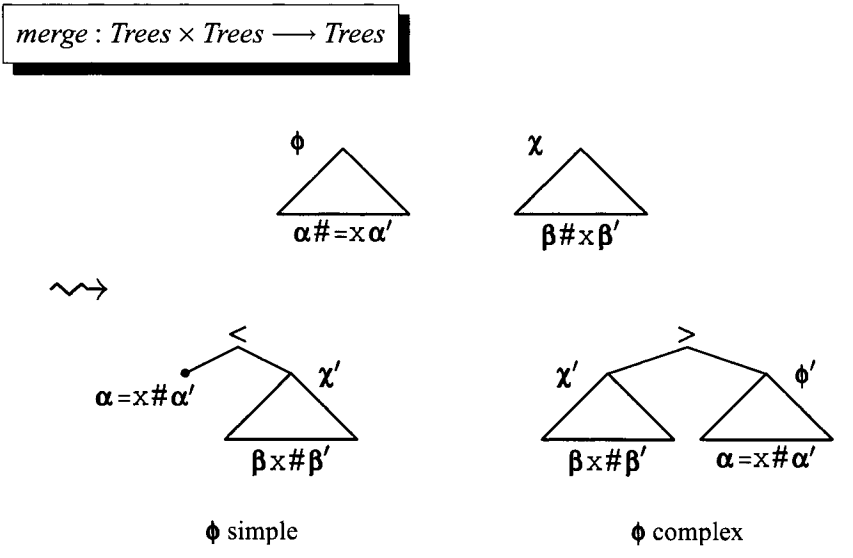


Figure 2. The merge-operator.

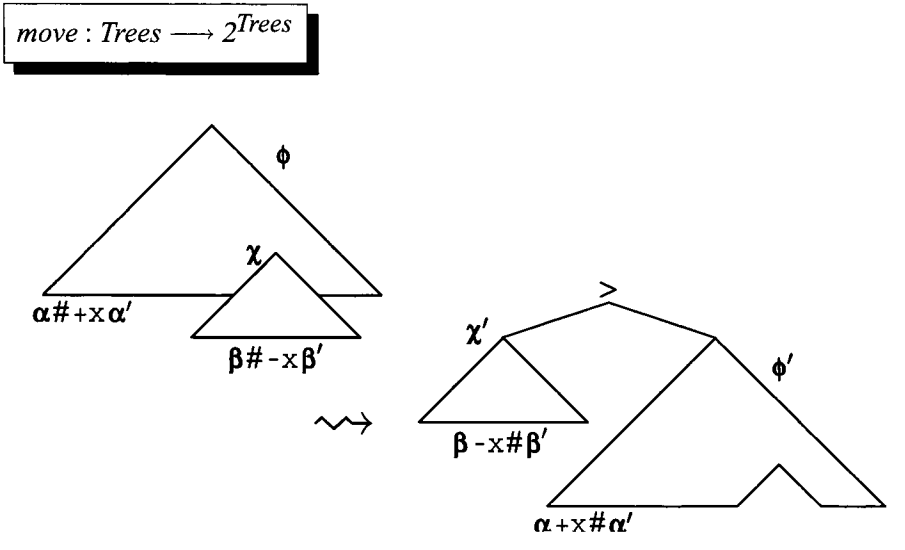


Figure 3. The move-operator.

The *tree language of an MG* is the set of trees of category c (the root category "complete" or "complementizer") each of which with essentially no unchecked syntactic features left after having been derived from a finite number of (possibly multiple) instances of lexical items by successively applying structure building operators. The *string language of an MG* is the set of strings each of which resulting from concatenating "left-to-right" the terminal leaf-labels of some tree belonging to the tree language.

Standard MGs usually come with a specific implementation of the *shortest move condition (SMC)*: for each MG there is an absolute (finite) upper bound n on the number of competing, i.e. simultaneously displayed, licensee features triggering an application of the move-operator. In the most radical version we have $n = 1$. As shown in Fig. 4, in the standard case this excludes both crossing and nesting dependencies involving multiple licensees of one and the same type.[6] Note also that, in this sense, absence of the SMC (– *SMC*, for short) means that no absolute upper bound on simultaneously displayed licensee features exists.

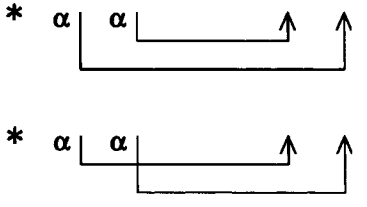


Figure 4. The shortest movement condition (SMC) (Stabler 1997, 1999)

The MG-variant proposed by Stabler (1999) also includes an implementation of the *specifier island condition (SPIC)* which essentially demands that proper extraction from specifiers is blocked (see Fig. 5).

Structurally similar to the SPIC, MGs can be endowed with an implementation of the *adjunct island condition (AIC)* demanding that, if at all, only full adjuncts but none of their proper subparts can extract (see Fig. 6).

Talk of adjuncts and the AIC presupposes extending MGs with additional syntactic features and structure building functions. To the list of syntactic

[6]See Section 3.2 for an exploitation of the dynamic character of the SMC. See Michaelis (2001b) and Stabler (1999) for the MG-treatment of cross-serial dependencies. See Appendix B for our approach to multiple wh-dependencies.

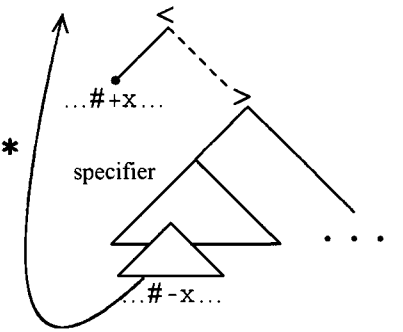


Figure 5. The specifier island condition (SPIC) (Stabler 1999).

features we add:

a(djoin)-selectors: $\approx x, \approx y, \approx z, \ldots$

s(cramble)-licensees: $\sim x, \sim y, \sim z, \ldots$

Then we introduce an *adjoin*-operator and *extraposition/scramble*-operator, which in contrast to the merge- and move-operator do not function as a bilateral checking mechanism but a unilateral one. This implements type-preservingness and iterability of adjunction, as is familiar from categorial grammar.

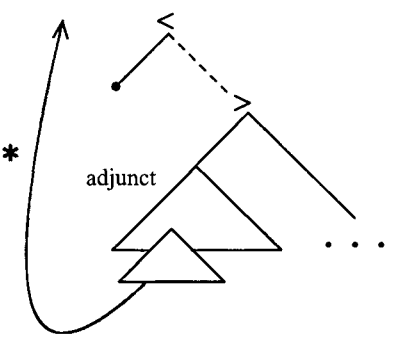
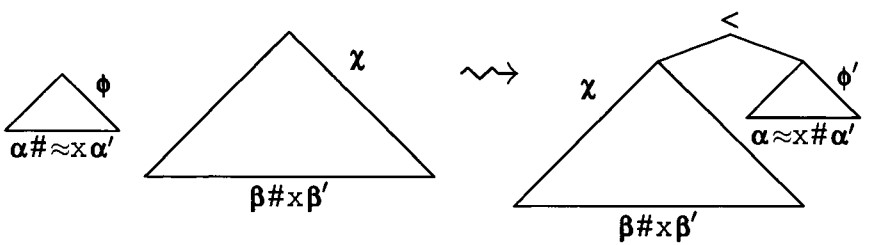


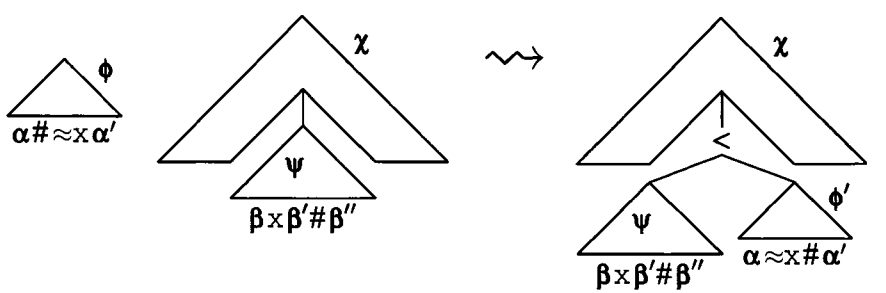
Figure 6. The adjunct island condition (AIC) (Frey and Gärtner 2002).

Instances of (basic) category features and a-selectors trigger the adjoin-operator mapping a pair of trees, $\langle \phi, \chi \rangle$, to a finite set of trees, consisting of every tree which results from adjoining the tree ϕ if it displays the a-selector $\approx x$ to the tree χ: cyclically in case χ displays the corresponding category x, or acyclically to a maximal projection ψ properly contained in χ in case the head-label of ψ contains a checked instance of the category x. In both cases, the a-selector feature instance triggering the application of the operator gets marked as checked in the resulting tree, while the other head-label of χ, respectively ψ, remains unchanged (cf. Fig. 7).

$$adjoin : Trees \times Trees \longrightarrow 2^{Trees}$$



cyclic adjunction (Frey and Gärtner 2002)



acyclic/late adjunction (Gärtner and Michaelis 2003)

Figure 7. The operator *adjoin*.

Instances of (basic) categories and s-licensees trigger applications of the scramble-operator which—without imposing the SMC—maps a single tree displaying category x to a finite set of trees, consisting of every tree which results from extraposing/scrambling a maximal projection displaying the corresponding s-licensee ~x into an adjoined position. Again, only the s-licensee feature instance triggering the application of the operator gets marked as checked in the resulting tree, while the corresponding head-label displaying category x remains unchanged (cf. Fig. 8).

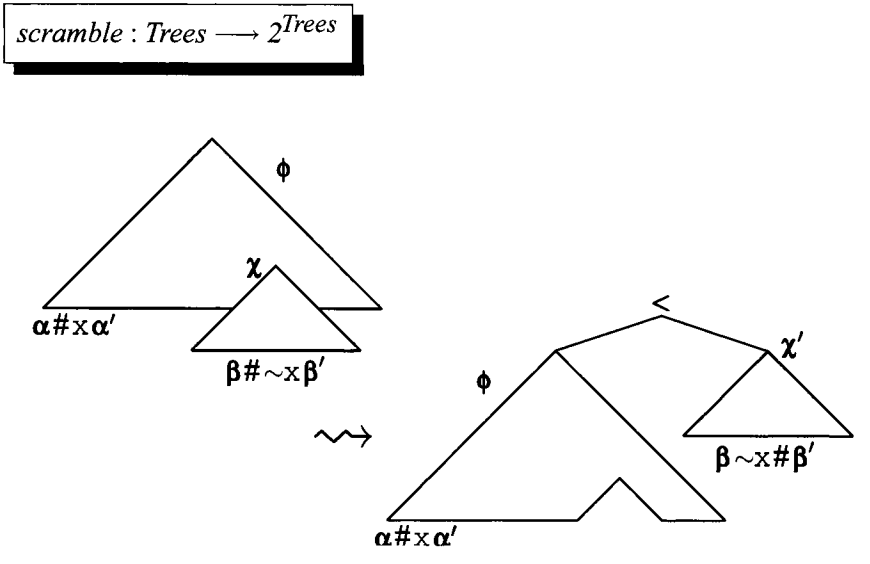


Figure 8. The operator *scramble*.

3 Locality Conditions and Complexity Results

As indicated in Section 1, our complexity results concern the interaction of locality constraints. In Section 3.1 we look at the interaction of the SPIC and the SMC within standard MGs. In Section 3.2 we introduce MGs with late adjunction and discuss the interaction of the AIC and the SMC within such extended grammars. The *MG-diamonds* in Fig. 9 provide a systematic picture for our study. The ultimate task is to establish complexity results for each corner and to reflect on their relation.

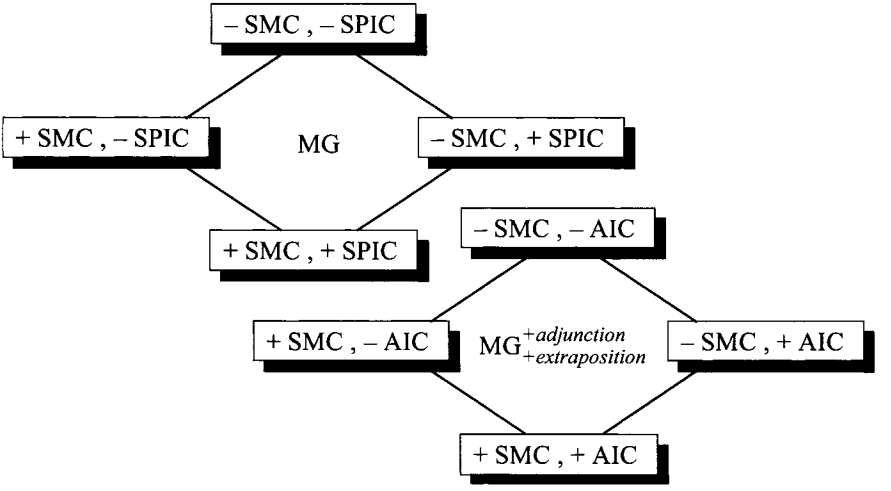
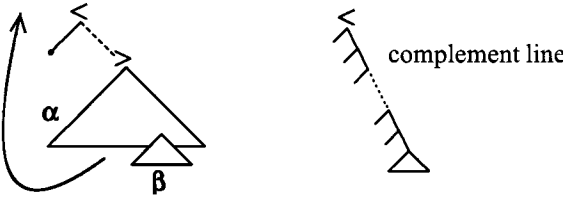


Figure 9. MG-diamonds — Towards complexity results concerning LCs

3.1 The Specifier Island Condition

Fig. 10 presents an example of a non-mildly context-sensitive MG not fulfilling the SMC but the SPIC, and deriving a language without constant growth property, namely, $\{a^{2^n} \mid n \geq 0\} = \{a, aa, aaaa, aaaaaaaa, \ldots\}$. The central column shows the lexical items as they are drawn from the lexicon, i.e., with all features unchecked. Arrows show the possible orders of interaction among lexical items and resulting constituents in terms of *merge*. Intermediate steps of *move* are left implicit.

As shown by Kobele and Michaelis (2005), not only this language, but in fact every language of type 0 can be derived by some MG not fulfilling the SMC but the SPIC for essentially two reasons: a) because of the SPIC,



movement of a constituent α into a specifier position freezes every proper subconstituent β within α, and b) without the SMC, therefore, the complement

line of a tree (in terms of the successively embedded complements) can technically be employed as a queue. As is well-known, systems able to simulate queues are able to generate arbitrary type 0-languages.

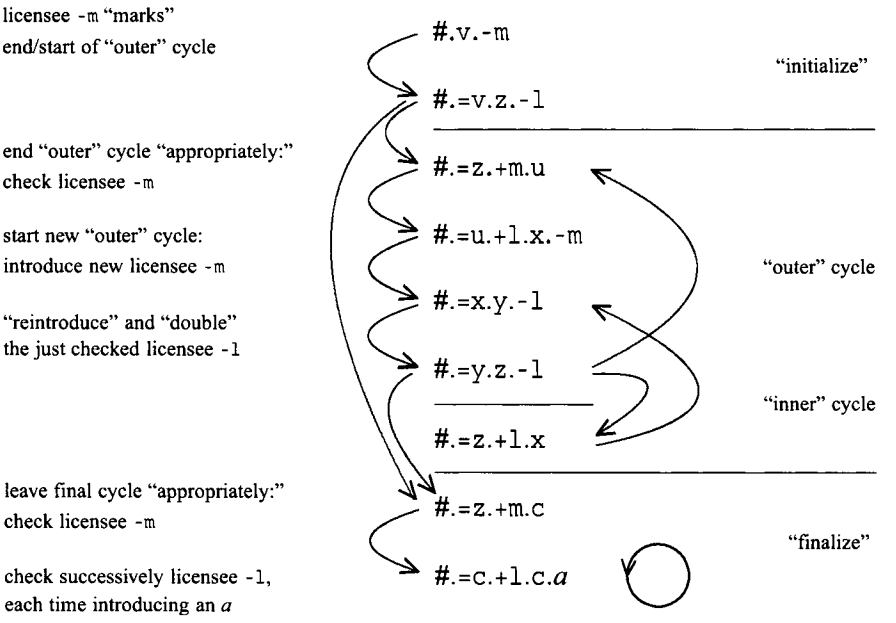


Figure 10. MG-example — Complexity results concerning LCs

Starting the "outer" cycle of our example in Fig. 10, the currently derived tree shows 2^n+1 successively embedded complements on the complement line, all with an unchecked instance of `-l`, except for the lowest one, which displays `-m`. (n equals the number of cycles already completed.) The initializing selecting head `#.=v.z.-l` introduces an additional m-licensee `-l` to create string a on a cycleless derivation. Going through the cycle provides a successive bottom-to-top "roll-up" of those complements in order to check the displayed features. Thereby, $2^{n+1}+1$ successively embedded complements on the complement line are created, again all displaying feature `-l` except for the lowest, which displays feature `-m`. Leaving the cycle procedure after a cycle has been completed leads to a final checking of the displayed licensees, where for each checked `-l` an a is introduced in the structure. This is the only way to create a convergent derivation.[7] Fig. 11 shows the result of a cycleless derivation creating string a, and a one-cycle derivation creating string aa.

[7]For further details see Gärtner and Michaelis (2005).

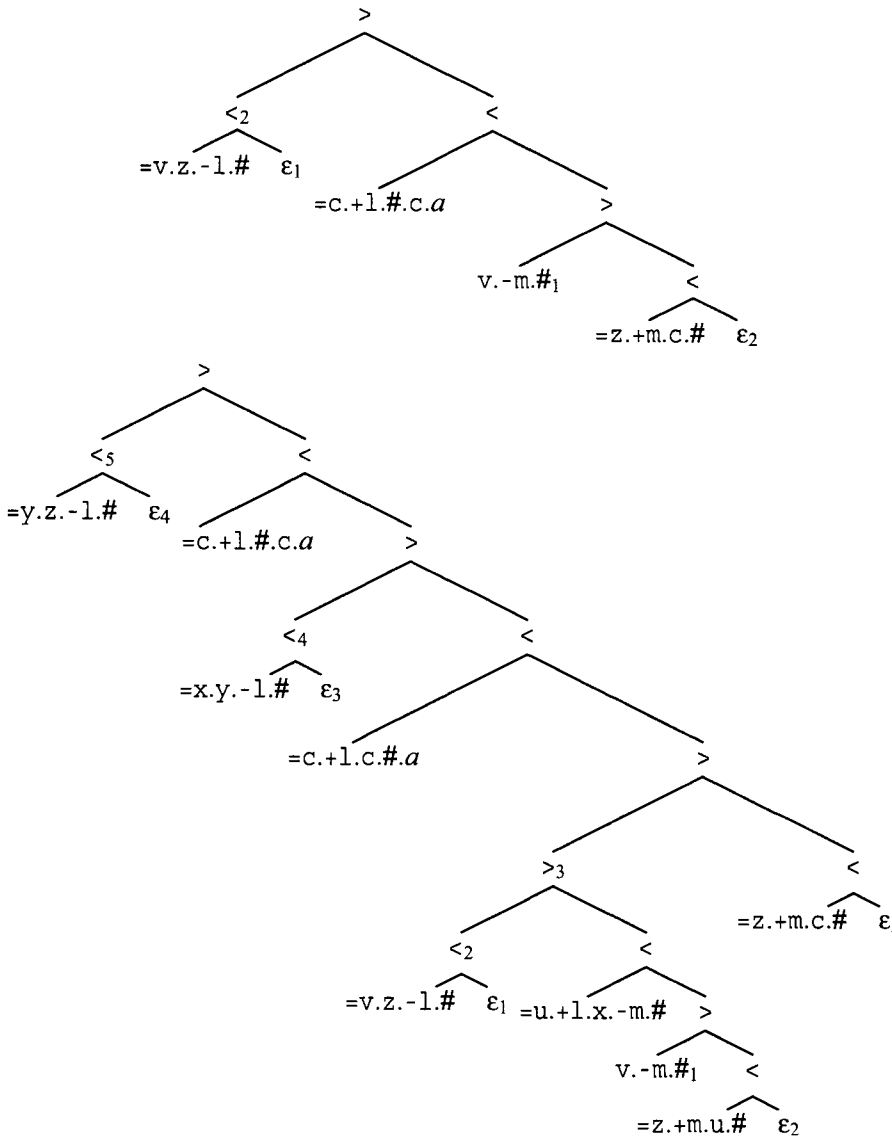


Figure 11. MG-example — Complexity results concerning LCs
(Numerical indices indicate antecedent-trace relations)

Fig. 12 summarizes what we know about the interaction of SMC and SPIC,[8] where $\mathcal{L}_1 \searrow \mathcal{L}_2$, respectively $\mathcal{L}_2 \swarrow \mathcal{L}_1$, means "language class $\mathcal{L}_2$ is lower in generative capacity than language class $\mathcal{L}_1$" while $\mathcal{L}_1 \nearrow \mathcal{L}_2$, respectively $\mathcal{L}_2 \nwarrow \mathcal{L}_1$, means "language class $\mathcal{L}_2$ is higher in generative capacity than language class $\mathcal{L}_1$." Crucially, adding the SPIC can either properly reduce complexity (lower left side) or properly increase complexity (upper right side). What the SPIC does depends on the presence or absence of SMC. Its behavior is thus non-monotonic.

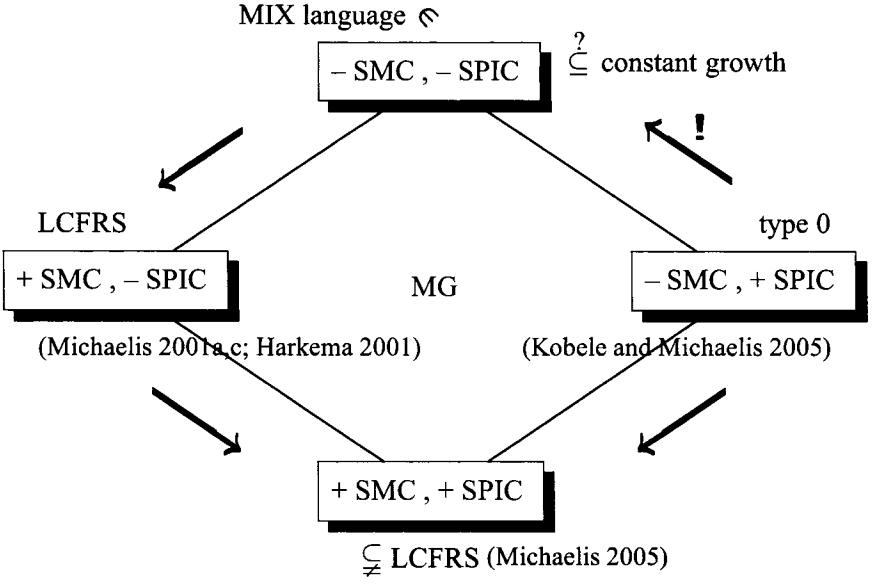


Figure 12. MG-diamond — Shortest move (SMC) and specifier islands (SPIC)

3.2 The Adjunct Island Condition

In this section we look at MGs with *(late) adjunction* and *scrambling/extraposition* and study the effects of imposing the AIC in a situation where the SMC alone appears to be too weak to guarantee mild context-sensitivity. As

[8] In Fig. 12 *LCFRS* stands for *Linear Context-Free Rewriting System*. For a more comprehensive picture of how these systems fit into the MCSG landscape see Appendix C. The MIX language is the language of all finite strings consisting of an equal number of *a*'s, *b*'s, and *c*'s appearing in arbitrary order.

in Section 3.1, the task is to fill in complexity relations between the corners of our MG-diamond, shown with relevant changes made in Fig. 13.

Late or countercyclic adjunction has already been introduced in Section 2 (cf. Fig. 7). One of its main linguistic motivations, going back (at least) to Lebeaux (1991), has to do with the possibility of avoiding standardly predicted but unattested violations of binding principle C. This is done by adjoining a constituent containing an R-expression after the constituent adjoined to has moved out of the c-command domain of a potentially offensive binder for that R-expression. (2) gives an example with a modifying relative clause.

(2) $[_{DP}\ [_{DP}$ which book $]_j\ [_{CP}$ that Mary$_i$ read $]\ ]$ did she$_i$ like t_j

For the complexity issue we are interested in here it is important to note that, as already briefly indicated by Gärtner and Michaelis (2003), late adjunction is capable of circumventing the SMC. (3) presents a case where this is actually welcome.

(3) [[[[Only those papers t_i]$_k$ did [everyone t_j] read t_k] [who was on the committee]$_j$] [that deal with adjunction]$_i$]

We assume for simplicity that both relative clauses in (3) are extraposed by an application of rightward scrambling and are adjoined to CP. This is very roughly sketched in (4).

(4) * $[_{CP}$ CP_2 CP_1] α α

This violates the SMC (see above) if α is instantiated as $\sim$C. However, as sketched in (5), a derivational sequence of (first) extraposition, late adjunction and (second) extraposition voids this problem.

(5)	$[_{CP}$		CP_1^{α}]			start here
	$[_{CP}$		--]	$CP_1^{\not\alpha}$		move CP_1, check α
	$[_{CP}$	CP_2^{α}	--]	$CP_1^{\not\alpha}$		late adjoin CP_2
	$[_{CP}$	--	--]	$CP_1^{\not\alpha}$	$CP_2^{\not\alpha}$	move CP_2, check α

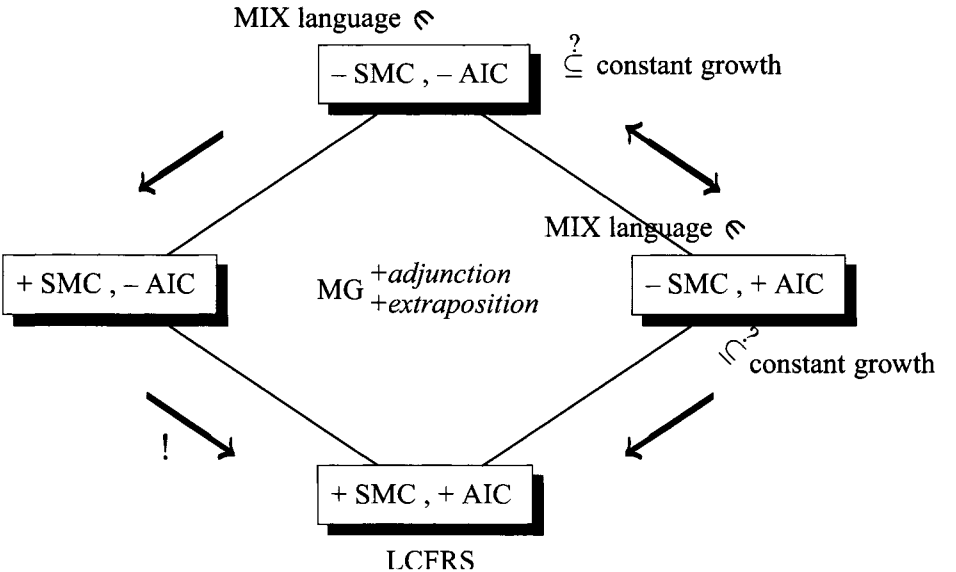


Figure 13. MG-diamond — Shortest Move (SMC) and Adjunct Islands (AIC)

The proof that MGs without late adjunction are mildly context-sensitive rests on the technical possibility of removing checked features from the structures.[9] Formally, late adjunction creates a situation where in order to locate the individual adjunction sites, an a priori not bounded amount of (categorial) information has to be stored during a derivation, i.e., adjunction sites have to

[9]See Stabler and Keenan (2003) for a reduced MG-format that cashes this out representationally. Chomsky (2005, p. 11) characterizes his "... 'no-tampering' condition of efficient computation" in almost the same way. Speaking of "operations forming complex expressions" Chomsky notes that it "sharply reduces computational load" if "what has once been constructed can be 'forgotten' in later computations." Without noting the tension created, Chomsky (2005, p. 12) introduces the "internal Merge" implementation of movement. This operation in fact requires an a priori not bounded amount of structure to remain available for copying and displacement. This undoes the effect of whatever structure may be 'forgotten' otherwise. Introducing the notion of "edge of a phase" (Chomsky 2001) as container for "still active" constituents does not essentially improve the situation, as long as there is no upper bound on the material inside such an edge. As far as we can see, this also negatively affects attempts by Chesi (2004) at providing any "measurable" complexity reductions in terms of phase-based locality. The point made by Berwick (1992) is closely related. Thus, "computational intractability" results if syntactic traces or "variables" are allowed to preserve an arbitrary amount of information (full copying being the extreme case).

be kept accessible. Therefore it is unclear whether, in general, MGs allowing late adjunction still belong to the same complexity class. If, however, the AIC is imposed, we can apply a specific reduction method in proving that for the resulting MGs the old complexity result holds. Under this reduction, however, late adjunction can only be simulated if the adjunct does not properly contain constituents bearing unchecked m- or s-licensees. But, this is exactly the situation where the AIC comes in. From a linguistic point of view it is rather natural to exclude extraction from adjuncts as Huang (1982) argued. This means that the weak generative capacity of MGs with late adjunction and extraposition can be kept within the bounds of standard MGs, i.e. mild context-sensitivity, if the AIC is imposed in addition to the SMC. Fig. 13 summarizes our results for SMC/AIC-interaction. Again, addition of an LC does not automatically restrict the grammar, as the upper right side shows. We conjecture that the AIC is a formal restriction only where it complements the SMC.

4 Conclusion and Further Outlook

Let us take a step back and summarize what we have found out about LCs within Stablerian MGs. Taking restrictiveness to be defined as weak generative capacity, we have illustrated how imposition of an LC can have either:

(A) restrictive effects, or

(B) no restrictive effects, or

(C) anti-restrictive effects.

Thus, adding the SPIC to standard MGs raises them to type 0 grammars if the SMC is absent, while together with the SMC it induces a genuine restriction (Section 3.1). Adding the AIC to an MG extended with the operations of late adjunction and extraposition (via rightward scrambling) is without effects unless the SMC is co-present. In the latter case, the AIC guarantees mild context-sensitivity, which the extended MGs without it are likely to go beyond (Section 3.2). We think that these non-monotonic effects of LCs should be of interest to everyone caring about formal (and measurable) notions of restrictiveness. In a nutshell, the message is that "constraints do not always constrain." Our result for MGs without SMC, but obeying the SPIC can be seen in the light of what Rogers (1998, p. 3f) concludes about a famous similar case:

> The significance of the [Peters & Ritchie-]results is [...] that, by itself, the hypothesis that natural languages are characterized by *Aspects*-style TGs [...] has no non-trivial consequences with respect to the class of natural languages.

Equally, by itself, the hypothesis that natural languages are characterized by the said MGs has no non-trivial consequences with respect to the class of natural languages.

As pointed out in Section 1, all of these issues have regained relevance due to the recent emergence of "cognitivist" studies of language evolution (Hauser et al. 2002, Piattelli-Palmarini and Uriagereka 2004) that reintroduce notions of classical formal complexity theory. Likewise, Chomskyan minimalism conceives of *computational efficiency* as contributing to the *design factors of language* (Chomsky 2005). This comes at a time where more and more grammar types have begun to converge on the mildly context-sensitive format (Joshi et al. 1991), Stablerian MGs among them.

There are some obvious ways to pursue the work begun here further. First of all, we have not looked at the interaction of SPIC and AIC. This is particularly relevant for MGs with late adjunction and extraposition for the following reasons. First, it is unclear whether the SPIC should constrain extraposition as much as it would in our current formalization. Secondly, the dynamics of late adjunction call for greater care to be taken in distinguishing static from dynamic formulations of LCs, i.e. LCs that put absolute bans on output structures vs. LCs that constrain individual derivational steps. On a more speculative note, it also remains to be seen whether a different division of labor between competence and performance aspects of grammars, as envisioned by Sternefeld (1998), could reorganize the (complexity) landscape of grammar formalisms in a fruitful fashion.

Appendix A

Throughout we let $\neg Syn$ and Syn be a finite set of *non-syntactic features* and a finite set of *syntactic features*, respectively, in accordance with (F1)–(F3) below. We take *Feat* to be the set $\neg Syn \cup Syn$.

(F1) $\neg Syn$ is disjoint from Syn and partitioned into the sets *Phon* and *Sem*, a set of *phonetic features* and a set *semantic features*, respectively.

(F2) Syn is partitioned into six sets:[10]

[10]Elements from Syn will usually be typeset in `typewriter font`.

Base		a set of *(basic) categories*
M-Select	$= \{ =x \mid x \in Base \}$	a set of *m(erge)-selectors*
A-Select	$= \{ \approx x \mid x \in Base \}$	a set of *a(djoin)-selectors*
M-Licensors	$= \{ +x \mid x \in Base \}$	a set of *m(ove)-licensors*
M-Licensees	$= \{ -x \mid x \in Base \}$	a set of *m(ove)-licensees*
S-Licensees	$= \{ \sim x \mid x \in Base \}$	a set of *s(cramble)-licensees*

(F3) *Base* includes at least the category c.

We use *Licensees* as a shorthand denoting the set *M-Licensees* $\cup$ *S-Licensees*.

Definition 4.1 An *expression (over Feat)*, also referred to as a *minimalist tree (over Feat)*, is a 6-tuple $\langle N_\tau, \lhd_\tau^*, \prec_\tau, <_\tau, label_\tau \rangle$ obeying (E1)–(E3).

(E1) $\langle N_\tau, \lhd_\tau^*, \prec_\tau \rangle$ is a finite, binary (ordered) tree defined in the usual sense: N_τ is the finite, non-empty set of *nodes*, and $\lhd_\tau^*$ and $\prec_\tau$ are the respective binary relations of *dominance* and *precedence* on N_τ.[11]

(E2) $<_\tau \subseteq N_\tau \times N_\tau$ is the asymmetric relation of *(immediate) projection* that holds for any two siblings, i.e., for each $x \in N_\tau$ different from the root of $\langle N_\tau, \lhd_\tau^*, \prec_\tau \rangle$ either $x <_\tau sibling_\tau(x)$ or $sibling_\tau(x) <_\tau x$ holds.[12]

(E3) $label_\tau$ is the *leaf-labeling function* from the set of leaves of $\langle N_\tau, \lhd_\tau^*, \prec_\tau \rangle$ into $Syn^*\{\#\}Syn^*Phon^*Sem^*$.[13]

We take *Exp(Feat)* to denote the class of all expressions over *Feat*.

Let $\tau = \langle N_\tau, \lhd_\tau^*, \prec_\tau, <_\tau, label_\tau \rangle \in Exp(Feat)$.[14]

[11] Thus, $\lhd_\tau^*$ is the reflexive-transitive closure of $\lhd_\tau \subseteq N_\tau \times N_\tau$, the relation of *immediate dominance* on N_τ.

[12] $sibling_\tau(x)$ denotes the (unique) sibling of any given $x \in N_\tau$ different from the root of $\langle N_\tau, \lhd_\tau^*, \prec_\tau \rangle$. If $x <_\tau y$ for some $x, y \in N_\tau$ then x is said to *(immediately) project over* y.

[13] For each set M, M^* is the Kleene closure of M, including ε, the empty string. For any two sets of strings, M and N, MN is the product of M and N w.r.t. string concatenation. Further, # denotes a new symbol not appearing in *Feat*.

[14] Note that the leaf-labeling function $label_\tau$ can easily be extended to a total labeling function ℓ_τ from N_τ into $Feat^*\{\#\}Feat^* \cup \{<, >\}$, where < and > are two new distinct symbols: to each non-leaf $x \in N_\tau$ we can assign a label from $\{<, >\}$ by ℓ_τ such that $\ell_\tau(x) = <$ iff $y <_\tau z$ for $y, z \in N_\tau$ with $x \lhd_\tau y, z$, and $y \prec_\tau z$. In this sense a concrete $\tau \in Exp(Feat)$ is depictable in the way indicated in Fig. 1.

For each $x \in N_\tau$, the *head of x (in* τ*)*, denoted by $head_\tau(x)$, is the (unique) leaf of τ with $x \lhd_\tau^* head_\tau(x)$ such that each $y \in N_\tau$ on the path from x to $head_\tau(x)$ with $y \neq x$ projects over its sibling, i.e. $y <_\tau sibling_\tau(y)$. The *head of* τ is the head of τ's root. τ is said to be a *head* (or *simple*) if N_τ consists of exactly one node, otherwise τ is said to be a *non-head* (or *complex*).

An expression $\phi = \langle N_\phi, \lhd_\phi^*, \prec_\phi, <_\phi, label_\phi \rangle \in Exp(Feat)$ is a *subexpression of* τ in case $\langle N_\phi, \lhd_\phi^*, \prec_\phi \rangle$ is a subtree of $\langle N_\tau, \lhd_\tau^*, \prec_\tau \rangle$, $<_\phi = <_\tau \restriction_{N_\phi \times N_\phi}$, and $label_\phi = label_\tau \restriction_{N_\phi}$. Such a subexpression ϕ is a *maximal projection (in* τ*)* if its root is a node $x \in N_\tau$ such that x is the root of τ, or such that $sibling_\tau(x) <_\tau x$. $MaxProj(\tau)$ is the set of maximal projections in τ.

$comp_\tau \subseteq MaxProj(\tau) \times MaxProj(\tau)$ is the binary relation defined such that for all $\phi, \chi \in MaxProj(\tau)$ it holds that $\phi\, comp_\tau\, \chi$ iff $head_\tau(r_\phi) <_\tau r_\chi$, where r_ϕ and r_χ are the roots of ϕ and χ, respectively. If $\phi\, comp_\tau\, \chi$ holds for some $\phi, \chi \in MaxProj(\tau)$ then χ is a *complement of* ϕ *(in* τ*)*. $comp_\tau^+$ is the transitive closure of $comp_\tau$. $Comp^+(\tau)$ is the set $\{\phi \mid \tau\, comp_\tau^+\, \phi\}$.

$spec_\tau \subseteq MaxProj(\tau) \times MaxProj(\tau)$ is the binary relation defined such that for all $\phi, \chi \in MaxProj(\tau)$ it holds that $\phi\, spec_\tau\, \chi$ iff both $r_\chi = sibling_\tau(x)$ and $x <_\tau r_\chi$ for some $x \in N_\tau$ with $r_\phi \lhd_\tau^+ x \lhd_\tau^+ head_\tau(r_\phi)$, where r_ϕ and r_χ are the roots of ϕ and χ, respectively. If $\phi\, spec_\tau\, \chi$ for some $\phi, \chi \in MaxProj(\tau)$ then χ is a *specifier of* ϕ *(in* τ*)*. $Spec(\tau)$ is the set $\{\phi \mid \tau\, spec_\tau\, \phi\}$.

A $\phi \in MaxProj(\tau)$ is said to *have*, or *display*, *(open) feature* f if the label assigned to ϕ's head by $label_\tau$ is of the form $\beta \# f \beta'$ for some $f \in Feat$ and some $\beta, \beta' \in Feat^*$.[15]

τ is *complete* if its head-label is in $Syn^*\{\#\}\{c\}Phon^*Sem^*$, and each of its other leaf-labels is in $Syn^*\{\#\}Phon^*Sem^*$. Hence, a complete expression over *Feat* is an expression that has category c, and this instance of c is the only instance of a syntactic feature which is preceded by an instance of # within its local leaf-label, i.e. the leaf-label it appears in.

The *phonetic yield of* τ, denoted by $Y_{Phon}(\tau)$, is the string which results from concatenating in "left-to-right-manner" the labels assigned via $label_\tau$ to the leaves of $\langle N_\tau, \lhd_\tau^*, \prec_\tau \rangle$, and replacing all instances of non-phonetic features with the empty string, afterwards.

For any $\phi, \chi \in Exp(Feat)$, $[<\phi, \chi]$ (respectively, $[>\phi, \chi]$) denotes the complex expression $\psi = \langle N_\psi, \lhd_\psi^*, \prec_\psi, <_\psi, label_\psi \rangle \in Exp(Feat)$ for which ϕ and

[15] Thus, e.g., the expression depicted in (3) has feature +x, while there is a maximal projection which has feature -x.

χ are those two subexpressions such that $r_\psi \lhd_\psi r_\phi$, $r_\psi \lhd_\psi r_\chi$ and $r_\phi \prec_\psi r_\chi$, and such that $r_\phi <_\psi r_\chi$ (respectively $r_\chi <_\psi r_\phi$), where r_ϕ, r_χ and r_ψ are the roots of ϕ, χ and ψ, respectively.

For any $\phi, \chi, \psi \in \mathit{Exp}(\mathit{Feat})$ such that χ is a subexpression of ϕ, $\phi\{\chi/\psi\}$ is the expression which results from substituting ψ for χ in ϕ.

In the following we write *MG* as a shorthand for *minimalist grammar.*

Definition 4.2 An *MG without both SMC and SPIC (MG$^{/-,-/}$)* is a 5-tuple of the form $\langle \neg\mathit{Syn}, \mathit{Syn}, \mathit{Lex}, \Omega, \mathrm{c}\rangle$ where Ω is the operator set consisting of the structure building functions $\mathit{merge}^{/-/}$ and $\mathit{move}^{/-,-/}$ defined as in (me$^{\text{-SPIC}}$) and (mo$^{\text{-SMC,-SPIC}}$) below, respectively, and where *Lex* is a *lexicon (over Feat)*, a finite set of simple expressions over *Feat*, and each item $\tau \in \mathit{Lex}$ is of the form $\langle \{r_\tau\}, \lhd_\tau^*, \prec_\tau, <_\tau, \mathit{label}_\tau\rangle$ such that $\mathit{label}_\tau(r_\tau)$ is an element in $\{\#\}(\mathit{M\text{-}Select} \cup \mathit{M\text{-}Licensors})^* \mathit{Base}\, \mathit{M\text{-}Licensees}^* \mathit{Phon}^* \mathit{Sem}^*$.

The operators from Ω build larger structure from given expressions by successively checking "from left to right" the instances of syntactic features appearing within the leaf-labels of the expressions involved. The symbol # serves to mark which feature instances have already been checked by the application of some structure building operation.

(me$^{\text{-SPIC}}$) $\mathit{merge}^{/-/}$ is a partial mapping from $\mathit{Exp}(\mathit{Feat}) \times \mathit{Exp}(\mathit{Feat})$ into $\mathit{Exp}(\mathit{Feat})$. For any $\phi, \chi \in \mathit{Exp}(\mathit{Feat})$, $\langle \phi, \chi\rangle$ is in $\mathit{Dom}(\mathit{merge}^{/-/})$ if for some category $\mathrm{x} \in \mathit{Base}$ and $\alpha, \alpha', \beta, \beta' \in \mathit{Feat}^*$, conditions (me.i) and (me.ii) are fulfilled:[16]

(me.i) the head-label of ϕ is $\alpha\#{=}\mathrm{x}\alpha'$ (i.e. ϕ has m-selector =x), and

(me.ii) the head-label of χ is $\beta\#\mathrm{x}\beta'$ (i.e. χ has category x).

Then,

(me.1) $\mathit{merge}^{/-/}(\phi, \chi) = [_< \phi', \chi']$ if ϕ is simple, and

(me.2) $\mathit{merge}^{/-/}(\phi, \chi) = [_> \chi', \phi']$ if ϕ is complex,

where ϕ' and χ' result from ϕ and χ, respectively, just by interchanging the instance of # and the instance of the feature directly following the instance of # within the respective head-label (cf. Fig. 2).

[16] For a partial function f from a class A into a class B, $\mathit{Dom}(f)$ is the domain of f, i.e., the class of all $x \in A$ for which $f(x)$ is defined.

(mo$^{\text{-SMC,-SPIC}}$) $move^{/-,-/}$ is a partial mapping from $Exp(Feat)$ into the class $\mathcal{P}_{\text{fin}}(Exp(Feat))$.[17] A $\phi \in Exp(Feat)$ is in $Dom(move^{/-,-/})$ if for some $-x \in$ *M-Licensees* and $\alpha, \alpha' \in Feat^*$, (mo.i) and (mo.ii) are true:

(mo.i) the head-label of ϕ is $\alpha\#{+}x\alpha'$ (i.e. ϕ has licensor $+x$),

(mo.ii) there exists a $\chi \in MaxProj(\phi)$ with head-label $\beta\#{-}x\beta'$ for some $\beta, \beta' \in Feat^*$ (i.e. $\chi \in MaxProj(\phi)$ exists displaying feature $-x$).

Then,

$$move^{/-,-/}(\phi) = \left\{ [{>}\chi', \phi'] \;\middle|\; \begin{array}{l} \chi \in MaxProj(\phi) \text{ with head-label } \beta\#{-}x\beta' \\ \text{for some } \beta, \beta' \in Feat^* \end{array} \right\},$$

where ϕ' results from ϕ by interchanging the instance of # and the instance of $+x$ directly following it within the head-label of ϕ, while the subtree χ is replaced by a single node labeled ε. χ' arises from χ by interchanging the instance of # and the instance of $-x$ immediately to its right within the head-label of χ (cf. Fig. 3).

Definition 4.3 An *MG without SMC, but with SPIC* ($MG^{/-,+/}$) is a five-tuple of the form $\langle \neg Syn, Syn, Lex, \Omega, c \rangle$ where Ω is the operator set consisting of the structure building functions $merge^{/+/}$ and $move^{/-,+/}$ defined as in (me$^{\text{+SPIC}}$) and (mo$^{\text{-SMC,+SPIC}}$) below, respectively, and where *Lex* is a lexicon over *Feat* defined as in Definition 4.2.

(me$^{\text{+SPIC}}$) $merge^{/+/}$ is a partial mapping from $Exp(Feat) \times Exp(Feat)$ into $Exp(Feat)$. For any $\phi, \chi \in Exp(Feat)$, $\langle \phi, \chi \rangle$ is in $Dom(merge^{/+/})$ if for some category $x \in$ *Base* and $\alpha, \alpha', \beta, \beta' \in Feat^*$, conditions (me.i) and (me.ii) above and (me.spic) are fulfilled:

(me.spic) if ϕ is complex then there is no $\psi \in MaxProj(\chi)$ with head-label $\gamma\#y\gamma'$ for some $y \in$ *Licensees* and $\gamma, \gamma' \in Feat^*$ (i.e. the selected specifier does not properly contain a maximal projection with an unchecked syntactic feature instance).

Then, $merge^{/+/}(\phi, \chi) = merge^{/-/}(\phi, \chi)$.

[17] $\mathcal{P}_{\text{fin}}(Exp(Feat))$ is the class of all finite subsets of $Exp(Feat)$.

($\text{mo}^{\text{-SMC,+SPIC}}$) $move^{/-,+/}$ is a partial mapping from $Exp(Feat)$ into the class $\mathcal{P}_{\text{fin}}(Exp(Feat))$. A $\phi \in Exp(Feat)$ is in $Dom(move^{/-,+/})$ if for some $-x \in$ *M-Licensees* and $\alpha, \alpha' \in Feat^*$, (mo.i) and (mo.ii) given above and (mo.spic) are true:

(mo.spic) there is no $\psi \in MaxProj(\chi)$ different from χ, and with head-label $\gamma\#y\gamma'$ for some $y \in$ *Licensees* and $\gamma, \gamma' \in Feat^*$ (i.e. the maximal projection moved to the specifier does not itself properly contain itself a maximal projection displaying an unchecked syntactic feature instance).

Then, $move^{/-,+/}(\phi) = move^{/-,-/}(\phi)$.

The formulation of the SPIC as presented here, could be seen as an "active" variant, preventing the creation of expressions which include specifiers from which proper extraction could potentially take place. The MG-version presented in Stabler 1999 allows derivation of such expressions, but prevents these expressions to enter a convergent derivation by explicity stating a "passive" formulation of the SPIC, demanding that the maximal projection $\chi \in MaxProj(\phi)$ which has feature $-x$ can only move in order to check the licensee, if there exists a $\psi \in Comp^+(\phi)$ with $\chi = \psi$ or $\chi \in Spec(\psi)$.

Definition 4.4 An *MG with SMC, but without SPIC* ($MG^{/+,-/}$) is a five-tuple of the form $\langle \neg Syn, Syn, Lex, \Omega, c \rangle$ where Ω is the operator set consisting of the structure building functions $merge^{/-/}$ and $move^{/+,-/}$ defined as in ($\text{me}^{\text{-SPIC}}$) above and ($\text{mo}^{\text{+SMC,-SPIC}}$) below, respectively, and where *Lex* is a lexicon over *Feat* defined as in Definition 4.2.

($\text{mo}^{\text{+SMC,-SPIC}}$) $move^{/+,-/}$ is a partial mapping from $Exp(Feat)$ into the class $\mathcal{P}_{\text{fin}}(Exp(Feat))$. A $\phi \in Exp(Feat)$ belongs to $Dom(move^{/+,-/})$ if for some $-x \in$ *M-Licensees* and $\alpha, \alpha' \in Feat^*$, (mo.i) and (mo.ii) above and (mo.smc) are true:

(mo.smc) exactly one $\chi \in MaxProj(\phi)$ exists with head-label $\gamma\#-x\gamma'$ for some $\gamma, \gamma' \in Feat^*$ (i.e. exactly one $\chi \in MaxProj(\phi)$ has $-x$).[18]

Then, $move^{/+,-/}(\phi) = move^{/-,-/}(\phi)$.

[18] Note that condition (mo.smc) implies (mo.ii).

Definition 4.5 An *MG with both SMC and SPIC ($MG^{/+,+/}$)* is a five-tuple of the form $\langle \neg Syn, Syn, Lex, \Omega, c \rangle$ where Ω is the operator set consisting of the structure building functions $merge^{/+/}$ and $move^{/+,+/}$ defined as in (me^{+SPIC}) above and ($mo^{+SMC,+SPIC}$) below, respectively, and where *Lex* is a lexicon over *Feat* defined as in Definition 4.2.

($mo^{+SMC,+SPIC}$) $move^{/+,+/}$ is a partial mapping from $Exp(Feat)$ into the class $\mathcal{P}_{fin}(Exp(Feat))$. A $\phi \in Exp(Feat)$ is in $Dom(move^{/+,+/})$ if for some $-x \in$ *M-Licensees* and $\alpha, \alpha' \in Feat^*$, (mo.i), (mo.ii), (mo.spic) and (mo.smc) above are true. Then, $move^{/+,+/}(\phi) = move^{/-,-/}(\phi)$.[19]

Let $G = \langle \neg Syn, Syn, Lex, \Omega, c \rangle$ be an $MG^{/-,-/}$, $MG^{/-,+/}$, $MG^{/+,-/}$, respectively $MG^{/+,+/}$. For the sake of convenience, we refer to the corresponding merge- and move-operator in Ω by *merge* and *move*, respectively. Then the *closure of G*, $CL(G)$, is the set $\bigcup_{k \in \mathbb{N}} CL^k(G)$, where $CL^0(G) = Lex$, and for $k \in \mathbb{N}$,[20] $CL^{k+1}(G) \subseteq Exp(Feat)$ is recursively defined as the set

$$CL^k(G) \cup \{merge(\phi, \chi) \mid \langle \phi, \chi \rangle \in Dom(merge) \cap CL^k(G) \times CL^k(G)\} \cup \bigcup_{\phi \in Dom(move) \cap CL^k(G)} move(\phi).$$

The set $\{\tau \mid \tau \in CL(G) \text{ and } \tau \text{ complete}\}$, denoted by $T(G)$, is the *minimalist tree language derivable by G*. The set $\{Y_{Phon}(\tau) \mid \tau \in T(G)\}$, denoted by $L(G)$, is the *minimalist (string) language derivable by G*.

In the following we will use the notation $MG_{adj,ext}$ as a shorthand for *minimalist grammar with generalized adjunction and extraposition.*

Definition 4.6 An $MG_{adj,ext}$ *without both SMC and AIC* ($MG^{/-,-/}_{adj,ext}$) is a 5-tuple $G = \langle \neg Syn, Syn, Lex, \Omega, c \rangle$ where Ω is the operator set consisting of the functions $merge^{/-/}$, $move^{/-,-/}$, $adjoin^{/-/}$ and $scramble^{/-,-/}$ defined as in (me^{-SPIC}) and ($mo^{-SMC,-SPIC}$) above, and (ad^{-AIC}) and ($sc^{-SMC,-AIC}$) below, respectively, and where *Lex* is a *lexicon (over Feat)*, i.e., a finite set of simple expressions over *Feat*, and each lexical item $\tau \in Lex$ is of the form $\langle \{r_\tau\}, \lhd_\tau^*, \prec_\tau, <_\tau, label_\tau \rangle$ such that $label_\tau(r_\tau)$ is an element belonging to $\{\#\}(M\text{-}Select \cup M\text{-}Licensors)^*(Base \cup A\text{-}Select)Licensees^*Phon^*Sem^*$.

[19] Note that the the sets $move^{/+,-/}(\phi)$ and $move^{/+,+/}(\phi)$ in ($mo^{+SMC,-SPIC}$) and ($mo^{+SMC,+SPIC}$), respectively, both are singleton sets because of (SMC). Thus, these functions can easily be identified with one from $Exp(Feat)$ to $Exp(Feat)$.

[20] $\mathbb{N}$ is the set of all non-negative integers.

(ad$^{\text{-AIC}}$) *adjoin*$^{/-/}$ is a partial mapping from $Exp(Feat) \times Exp(Feat)$ into the class $\mathcal{P}_{\text{fin}}(Exp(Feat))$. A pair $\langle \phi, \chi \rangle$ with $\phi, \chi \in Exp(Feat)$ belongs to $Dom(adjoin^{/-/})$ if for some category $x \in Base$ and $\alpha, \alpha' \in Feat^*$, conditions (ad.i) and (ad.ii) are fulfilled:

(ad.i) the head-label of ϕ is $\alpha\#{\approx}x\alpha'$ (i.e. ϕ has a-selector $\approx$x), and

(ad.ii) there exists some $\psi \in MaxProj(\phi)$ with head-label of the form $\beta\#x\beta'$ or $\beta x\beta'\#\beta''$ for some $\beta, \beta', \beta'' \in Feat^*$

Then,

$$adjoin^{/-/}(\phi, \chi) = \left\{ \chi\{\psi/[{<}\psi, \phi']\} \;\middle|\; \begin{array}{l} \psi \in MaxProj(\chi) \text{ with head-label } \beta\#x\beta' \text{ or } \beta x\beta'\#\beta'' \text{ for some} \\ \beta, \beta', \beta'' \in Feat^* \end{array} \right\},$$

where ϕ' results from ϕ by interchanging the instances of # and $\approx$x, the latter directly following the former in the head-label of ϕ (cf. Fig. 7).

(sc$^{\text{-SMC,-AIC}}$) The function *scramble*$^{/-,-/}$ maps partially from $Exp(Feat)$ into the class $\mathcal{P}_{\text{fin}}(Exp(Feat))$. A $\phi \in Exp(Feat)$ is in $Dom(scramble^{/-,-/})$ if for some $x \in Base$ and $\alpha, \alpha' \in Feat^*$, (sc.i) and (sc.ii) are true:

(sc.i) the head-label of ϕ is $\alpha\#x\alpha'$ (i.e. ϕ has category x), and

(sc.ii) there is some $\chi \in MaxProj(\phi)$ with head-label $\beta\#{\sim}x\beta'$ for some $\beta, \beta' \in Feat^*$ (i.e. there is some $\chi \in MaxProj(\phi)$ displaying $\sim$x).

Then,

$$scramble^{/-,-/}(\phi) = \left\{ [{>}\chi', \phi'] \;\middle|\; \begin{array}{l} \chi \in MaxProj(\phi) \text{ with head-label} \\ \beta\#{\sim}x\beta' \text{ for some } \beta, \beta' \in Feat^* \end{array} \right\},$$

where $\phi' \in Exp(Feat)$ is identical to ϕ except for the fact that the subtree χ is replaced by a single node labeled ε. $\chi' \in Exp(Feat)$ arises from χ by interchanging the instance of # and the instance of $\sim$x immediately to its right within the head-label of χ (cf. Fig. 8).

Definition 4.7 An $MG_{adj,ext}$ *without SMC, but with AIC* ($MG^{/-,+/}_{adj,ext}$) is a five-tuple of the form $\langle \neg Syn, Syn, Lex, \Omega, c \rangle$ where Ω is the operator set consisting of the structure building functions $merge^{/-/}$, $move^{/-,-/}$, $adjoin^{/+/}$ and $scramble^{/-,+/}$ defined as in (me$^{-\text{SPIC}}$) and (mo$^{-\text{SMC,-SPIC}}$) above, and (ad$^{+\text{AIC}}$) and (sc$^{-\text{SMC,+AIC}}$) below, respectively, and where *Lex* is a lexicon over *Feat* defined as in Definition 4.6.

(ad$^{+\text{AIC}}$) $adjoin^{/+/}$ is a partial mapping from $Exp(Feat) \times Exp(Feat)$ into the class $\mathcal{P}_{\text{fin}}(Exp(Feat))$. A pair $\langle \phi, \chi \rangle$ with $\phi, \chi \in Exp(Feat)$ belongs to $Dom(adjoin^{/+/})$ if for some category $x \in Base$ and $\alpha, \alpha' \in Feat^*$, conditions (ad.i) and (ad.ii) above and (ad.aic) are fulfilled:

(ad.aic) there is no $\psi \in MaxProj(\phi)$ with head-label $\gamma \# y \gamma'$ for some $y \in Licensees$ and $\gamma, \gamma' \in Feat^*$ (i.e. the adjunct does not properly contain a maximal projection with an unchecked syntactic feature instance).

Then, $adjoin^{/+/}(\phi, \chi) = adjoin^{/-/}(\phi, \chi)$.

(sc$^{-\text{SMC,+AIC}}$) The function $scramble^{/-,+/}$ maps partially from $Exp(Feat)$ into the class $\mathcal{P}_{\text{fin}}(Exp(Feat))$. A $\phi \in Exp(Feat)$ is in $Dom(scramble^{/-,+/})$ if for some $x \in Base$ and $\alpha, \alpha' \in Feat^*$, (sc.i) and (sc.ii) above and (sc.aic) are true:

(sc.aic) there is no $\psi \in MaxProj(\chi)$ different from χ, and with head-label $\gamma \# y \gamma'$ for some $y \in Licensees$ and $\gamma, \gamma' \in Feat^*$ (i.e. the maximal projection scrambled/extraposed to an adjunct position does not itself properly contain itself a maximal projection displaying an unchecked syntactic feature instance).

Then, $scramble^{/-,+/}(\phi) = scramble^{/-,-/}(\phi)$.

Definition 4.8 An $MG_{adj,ext}$ *with SMC, but without AIC* ($MG^{/+,-/}_{adj,ext}$) is a five-tuple of the form $\langle \neg Syn, Syn, Lex, \Omega, c \rangle$ where Ω is the operator set consisting of the structure building functions $merge^{/-/}$, $move^{/+,-/}$, $adjoin^{/-/}$ and $scramble^{/+,-/}$ defined as in (me$^{-\text{SPIC}}$), (mo$^{+\text{SMC,-SPIC}}$) and (ad$^{-\text{AIC}}$) above and (sc$^{+\text{SMC,-AIC}}$) below, respectively, and where *Lex* is a lexicon over *Feat* defined as in Definition 4.6.

($\text{sc}^{+\text{SMC},-\text{AIC}}$) The function $scramble^{/+,-/}$ maps partially from $Exp(Feat)$ into the class $\mathcal{P}_{\text{fin}}(Exp(Feat))$. A $\phi \in Exp(Feat)$ is in $Dom(scramble^{/+,-/})$ if for some $x \in Base$ and $\alpha, \alpha' \in Feat^*$, (sc.i) and (sc.ii) above and (sc.smc) are true:

(sc.smc) exactly one $\chi \in MaxProj(\phi)$ exists with head-label $\gamma\#{\sim}x\gamma'$ for some $\gamma, \gamma' \in Feat^*$ (i.e. exactly one $\chi \in MaxProj(\phi)$ has ${\sim}x$).[21]

Then, $scramble^{/+,-/}(\phi) = scramble^{/-,-/}(\phi)$.

Definition 4.9 An $MG_{adj,ext}$ *with both SMC and AIC* ($MG^{/+,+/}_{adj,ext}$) is a five-tuple of the form $\langle \neg Syn, Syn, Lex, \Omega, c \rangle$ where Ω is the operator set consisting of the structure building functions $merge^{/-/}$, $move^{/+,-/}$, $adjoin^{/+/}$ and $scramble^{/+,+/}$ defined as in ($\text{me}^{-\text{SPIC}}$), ($\text{mo}^{+\text{SMC},-\text{SPIC}}$) and ($\text{ad}^{+\text{AIC}}$) above and ($\text{sc}^{+\text{SMC},+\text{AIC}}$) below, respectively, and where *Lex* is a lexicon over *Feat* defined as in Definition 4.6.

($\text{sc}^{+\text{SMC},+\text{AIC}}$) $scramble^{/+,+/}$ is a partial mapping from $Exp(Feat)$ into the class $\mathcal{P}_{\text{fin}}(Exp(Feat))$. A $\phi \in Exp(Feat)$ is in $Dom(scramble^{/+,+/})$ if for some $x \in Base$ and $\alpha, \alpha' \in Feat^*$, (sc.i), (sc.ii), (sc.aic) and (sc.smc) above are true. Then, $scramble^{/+,+/}(\phi) = scramble^{/-,-/}(\phi)$.

Consider an $\text{MG}^{/-,-/}_{\text{adj,ext}}$, $\text{MG}^{/-,+/}_{\text{adj,ext}}$, $\text{MG}^{/+,-/}_{\text{adj,ext}}$, respectively $\text{MG}^{/+,+/}_{\text{adj,ext}}$, G, of the form $\langle \neg Syn, Syn, Lex, \Omega, c \rangle$. For the sake of convenience, we refer to the corresponding merge-, move-, adjoin- and scramble-operator in Ω by *merge*, *move*, *adjoin* and *scramble*, respectively. The *closure of* G, $CL(G)$, is the set $\bigcup_{k \in \mathbb{N}} CL^k(G)$, where $CL^0(G) = Lex$, and for $k \in \mathbb{N}$, $CL^{k+1}(G) \subseteq Exp(Feat)$ is recursively defined as the set

$$CL^k(G) \cup \{merge(\phi,\chi) \mid \langle \phi,\chi \rangle \in Dom(merge) \cap CL^k(G) \times CL^k(G)\}$$
$$\cup \bigcup_{\phi \in Dom(move) \cap CL^k(G)} move(\phi)$$
$$\cup \bigcup_{\langle \phi,\chi \rangle \in Dom(adjoin) \cap CL^k(G) \times CL^k(G)} adjoin(\phi,\chi)$$
$$\cup \bigcup_{\phi \in Dom(scramble) \cap CL^k(G)} scramble(\phi)$$

The set $\{\tau \mid \tau \in CL(G) \text{ and } \tau \text{ complete}\}$, denoted by $T(G)$, is the *minimalist tree language derivable by* G. The set $\{Y_{Phon}(\tau) \mid \tau \in T(G)\}$, denoted by $L(G)$, is the *minimalist (string) language derivable by* G.

[21] Note that condition (sc.smc) implies (sc.ii).

Appendix B

One phenomenon that appears to challenge the SMC adopted here is multiple wh-fronting in Slavic languages. Take (6) from Bulgarian (Richards 2001, p. 249).

(6) *Koj$_i$ kogo$_j$ kakvo$_k$* t$_i$ *e pital* t$_j$ t$_k$
Who whom what AUX ask
'Who asked whom what?'

On standard assumptions, (6) requires three m-licensee instances of type -wh, which are successively checked in the C-domain. The required pre-movement representation, (7), is ruled out by the strictest version of the SMC (see above).

(7) [$_{IP}$ -wh.koj e [$_{VP}$ pital -wh.kogo -wh.kakvo]]

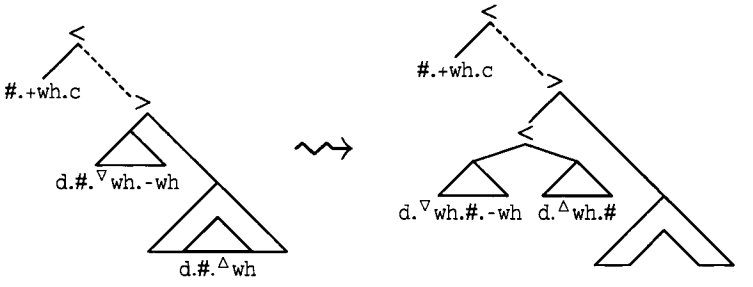
However, an SMC-violation can be circumvented if we adopt the wh-cluster hypothesis by Sabel (1998; 2001) and Grewendorf (2001). Under this perspective, wh-expressions undergo successive cluster-formation before the resulting cluster takes a single wh-movement step, in compliance with the SMC. For this we have to add the feature type of c(luster)-licensees and -licensors to MGs.

c(luster)-licensees: $^{\triangle}$x, $^{\triangle}$y, $^{\triangle}$z, ...
c(luster)-licensors: $^{\triangledown}$x, $^{\triangledown}$y, $^{\triangledown}$z, ...

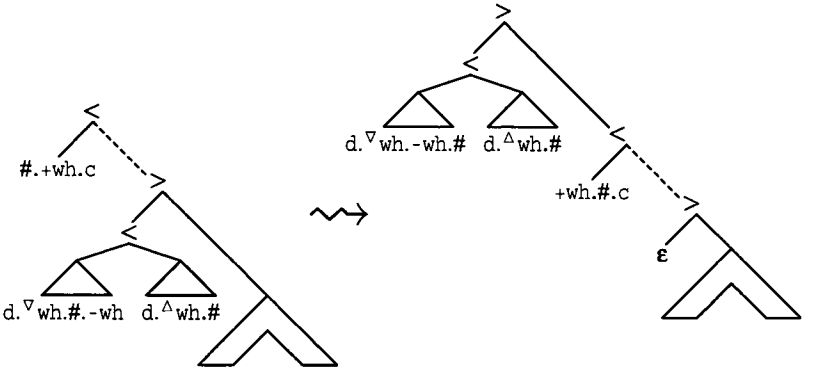
In Fig. 14 we show a derivation with two wh-phrases. For cases with three or more such phrases the intermediate ones have to be of type d.$^{\triangledown}$wh.$^{\triangle}$wh. Note that additional word order variation can be found in Bulgarian, as shown in (8) (Richards 2001, p. 249).

(8) *Koj kakvo kogo e pital*

This can be derived if cluster-formation is preceded by a scrambling-step of *kakvo* across *kogo* to VP, which requires it to be of type d.~v.$^{\triangledown}$wh. See Sabel (1998) for more discussion of wh- and focus-driven movements in multiple wh-configurations. Semantically, wh-cluster-formation can be interpreted as quantifier composition, a.k.a. "absorption" (Higginbotham and May 1981).



Wh-clustering, n = 2, crucial step 1



Wh-clustering, n = 2, crucial step 2

Figure 14. Wh-clustering involving c-licensors and c-licensees.

Appendix C

A general picture of the MCSG landscape is given in the next figure, where, in particular, we have the following abbreviations: TAG = tree adjoining grammars, LIG = linear indexed grammars, CCG = combinatory categorial grammars, HG = head grammars, LCFRS = linear context-free rewriting systems, MCTAG = (set local) multi-component tree adjoining grammars, IG = indexed grammars.

An arrow always points to a class which is less powerful in generative capacity. If there is a double-arrow between two classes their generative capacity is equal.

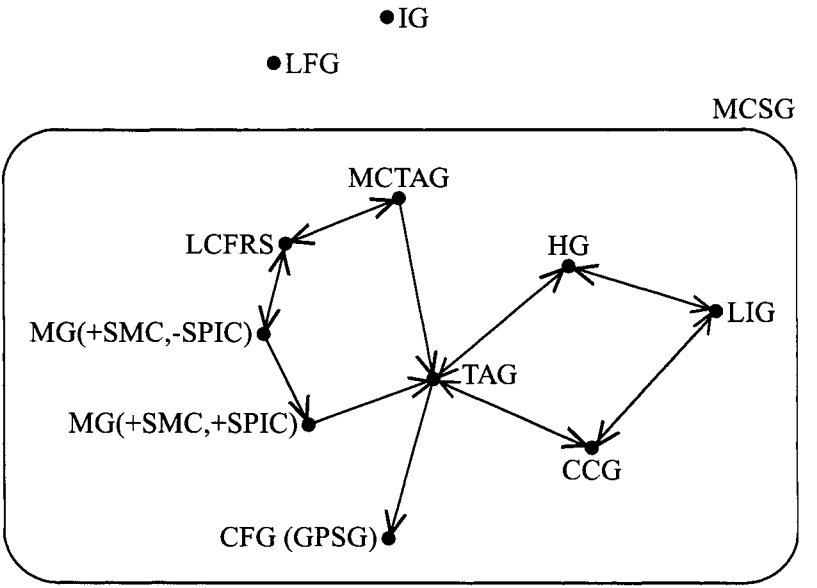


Figure 15. MCSG landscape

References

Berwick, Robert

1992 No variable is an island. Computational complexity and island constraints. In Goodluck and Rochemont (1992), pp. 35–59.

Berwick, Robert and Amy Weinberg

1982 Parsing efficiency, computational complexity, and the evaluation of grammatical theories. *Linguistic Inquiry*, 13: 165–191.

Chesi, Christiano

2004 *Complexity and Determinism in Linguistic Computation.* Manuscript, University of Siena, Siena.

Chomsky, Noam

1956 Three models for the description of language. In *IRE Transactions on Information Theory*, volume IT-2(3), pp. 113–124.

1959 On certain formal properties of grammars. *Information and Control*, 2: 137–167.

1973 Conditions on transformations. In S. Anderson and P. Kiparsky, (eds.), *A Festschrift for Morris Halle*, pp. 232–286. Holt, Rinehart and Winston, New York, NY.

1977 On wh-movement. In P. Culicover, T. Wasow, and A. Akmajian, (eds.), *Formal Syntax*, pp. 71–132. Academic Press, New York, NY.

1986 *Barriers.* MIT Press, Cambridge, MA.

1995 *The Minimalist Program.* MIT Press, Cambridge, MA.

2001 Derivation by phase. In Michael Kenstowicz, (ed.), *Ken Hale. A Life in Language*, pp. 1–52. MIT Press, Cambridge, MA.

2005 Three factors in language design. *Linguistic Inquiry*, 36: 1–22.

Cinque, Guglielmo

1990 *Types of A'-Dependencies.* MIT Press, Cambridge, MA.

de Groote, Philippe, Glyn Morrill, and Christian Retoré, (eds.)

2001 *Logical Aspects of Computational Linguistics (LACL '01)*, Lecture Notes in Artificial Intelligence Vol. 2099. Springer, Berlin, Heidelberg.

Frey, Werner and Hans-Martin Gärtner

2002 On the treatment of scrambling and adjunction in minimalist grammars. In *Proceedings of the Conference on Formal Grammar (FGTrento)*, Trento, pp. 41–52.

Gärtner, Hans-Martin and Jens Michaelis

2003 A note on countercyclicity and minimalist grammars. In *Proceedings of the Conference on Formal Grammar (FGVienna)*, Vienna, pp. 103–114.

2005 A note on the complexity of constraint interaction. Locality conditions and minimalist grammars. In P. Blache, E. Stabler, J. Busquets, and R. Moot, (eds.), *Logical Aspects of Computational Linguistics (LACL '05)*, Lecture Notes in Artificial Intelligence Vol. 3492, pp. 114–130. Springer, Berlin, Heidelberg.

Gibson, Edward
1991 Linguistic complexity. Locality of syntactic dependencies. *Cognition*, 68: 1–76.
Goodluck, Helen and Michael Rochemont, (eds.)
1992 *Island Constraints. Theory, Acquisition and Processing*. Kluwer, Dordrecht.
Grewendorf, Günther
2001 Multiple wh-fronting. *Linguistic Inquiry*, 32: 87–122.
Harkema, Henk
2001 A characterization of minimalist languages. In de Groote et al. (2001), pp. 193–211.
Hauser, Marc, Noam Chomsky, and Tecumseh Fitch
2002 The faculty of language. What is it, who has it, and how did it evolve? *Science*, 298: 1569–1579.
Higginbotham, James and Robert May
1981 Questions, quantifiers, and crossing. *The Linguistic Review*, 1: 41–79.
Huang, James C.-T.
1982 *Logical Relations in Chinese and the Theory of Grammar*. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.
Joshi, Aravind K.
1985 Tree adjoining grammars: How much context-sensitivity is required to provide reasonable structural descriptions? In D. R. Dowty, L. Karttunen, and A. M. Zwicky, (eds.), *Natural Language Parsing. Psychological, Computational, and Theoretical Perspectives*, pp. 206–250. Cambridge University Press, New York, NY.
Joshi, Aravind K. K. Vijay-Shanker, and David J. Weir
1991 The convergence of mildly context-sensitive grammar formalisms. In P. Sells, S. M. Shieber, and T. Wasow, (eds.), *Foundational Issues in Natural Language Processing*, pp. 31–81. MIT Press, Cambridge, MA.
Kobele, Gregory M. and Jens Michaelis
2005 Two type-0 variants of minimalist grammars. In *FG-MoL 2005. The 10th conference on Formal Grammar and The 9th Meeting on Mathematics of Language,* Edinburgh.
Kolb, Hans-Peter
1997 *GB Blues. Two Essays on Procedures and Structures in Generative Syntax*. Bericht Nr. 110, Arbeitspapiere des SFB 340, Universität Tübingen.
Lebeaux, David
1991 Relative clauses, licensing, and the nature of the derivation. In Susan D. Rothstein, (ed.), *Perspectives on Phrase Structure. Heads and Licensing*, pp. 209–239. Academic Press, New York, NY.
Manzini, Rita
1992 *Locality*. MIT Press, Cambridge, MA.

Michaelis, Jens

1998 Derivational minimalism is mildly context-sensitive. In *Proceedings of the Conference on Logical Aspects of Computational Linguistics (LACL '98)*, Grenoble, pp. 61–64.

2001a Derivational minimalism is mildly context-sensitive. In M. Moortgat, (ed.), *Logical Aspects of Computational Linguistics (LACL '98)*, Lecture Notes in Artificial Intelligence Vol. 2014, pp. 179–198. Springer, Berlin, Heidelberg.

2001b *On Formal Properties of Minimalist Grammars*. Linguistics in Potsdam 13. Universitätsbibliothek, Publikationsstelle, Potsdam.

2001c Transforming linear context-free rewriting systems into minimalist grammars. In de Groote et al. (2001), pp. 228–244.

2005 An additional observation on strict derivational minimalism. In *FG-MoL 2005. The 10th conference on Formal Grammar and The 9th Meeting on Mathematics of Language*, Edinburgh.

Müller, Gereon and Wolfgang Sternefeld

1993 Improper movement and unambiguous binding. *Linguistic Inquiry*, 24: 461–507.

Piattelli-Palmarini, Massimo and Juan Uriagereka

2004 Immune syntax. The evolution of the language virus. In Lyle Jenkins, (ed.), *Variation and Universals in Biolinguistics*, pp. 341–377. Elsevier, Oxford.

Pritchett, Bradley

1992 Parsing with grammar. Islands, heads, and garden paths. In Goodluck and Rochemont (1992), pp. 321–349.

Richards, Norvin

2001 *Movement in Language. Interactions and Architectures*. Oxford University Press, Oxford.

Rizzi, Luigi

1990 *Relativized Minimality*. MIT Press, Cambridge, MA.

Rogers, James

1998 *A Descriptive Approach to Language-Theoretic Complexity*. Studies in Logic, Language and Information. CSLI Publications, Stanford, CA.

Ross, John R.

1967 *Constraints on Variables in Syntax*. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Sabel, Joachim

1998 *Principles and Parameters of Wh-Movement*. Habilitationsschrift, Universität Frankfurt.

2001 Deriving multiple head and phrasal movement. The cluster hypothesis. *Linguistic Inquiry*, 32: 532–547.

Stabler, Edward P.

1997 Derivational minimalism. In C. Retoré, (ed.), *Logical Aspects of Computational Linguistics (LACL '96)*, Lecture Notes in Artificial Intelligence Vol. 1328, pp. 68–95. Springer, Berlin, Heidelberg.

1998 Acquiring languages with movement. *Syntax*, 1: 72–97.

1999 Remnant movement and complexity. In G. Bouma, G.-J. M. Kruijff, E. Hinrichs, and R. T. Oehrle, (eds.), *Constraints and Resources in Natural Language Syntax and Semantics*, pp. 299–326. CSLI Publications, Stanford, CA.

2001 Recognizing head movement. In de Groote et al. (2001), pp. 245–260.

Stabler, Edward P. and Edward L. Keenan

2003 Structural similarity within and among languages. *Theoretical Computer Science*, 293: 345–363.

Sternefeld, Wolfgang

1998 *Grammatikalität und Sprachvermögen*. SfS-Report-02-98, Seminar für Sprachwissenschaft, Universität Tübingen.

Szabolcsi, Anna and Frans Zwarts

1997 Weak islands and an algebraic semantics for scope taking. In Anna Szabolcsi, (ed.), *Ways of Scope Taking*, pp. 217–262. Kluwer, Dordrecht.

Flat Binding: Binding Without Sequences*

Uli Sauerland

1 Introduction

One of the most important properties of language is that it allows humans to combine basic concepts into novel, derived concepts by arranging words in a phrase structure. A central objective of linguistics is, therefore, to understand the relationship between the set of basic concepts that enter into a phrase structure and the concept derived from them. I assume with much recent work that at least two mechanisms underly this human ability (Montague 1974[1970]; Lewis 1970; Heim and Kratzer 1998, and others): One is a local mechanism that combines two concepts into one: local predication. For concreteness, I assume that two concepts combine into one by functional application. Local predication, however, is of subordinate concern to the present paper. The second mechanism is the non-local mechanism of binding.[1] The nature of

*This paper further develops a line of research I have been pursuing over the last decade. I have benefited from the advice of too many of my colleagues during this time to mention each one here, but remain grateful nevertheless. For useful comments on the present version, I thank Danny Fox, Hans-Martin Gärtner, Irene Heim, Chris Kennedy, Manfred Krifka, David Pesetsky, Chris Tancredi, Kazuko Yatsushiro, and Ede Zimmermann, the participants of classes on binding and pronouns that I taught at the University of Tokyo, Humboldt University Berlin, and the University of Vienna, and the audiences at shorter presentations at the University of Connecticut at Storrs, the Ecole Normale Supérieure at Paris, the University of Frankfurt and the *Interfaces + Recursion = Language?* Symposium in Berlin. Very detailed and much appreciated comments on the present version were given by Philippe Schlenker and Magdalena Schwager. Furthermore, I gratefully acknowledge the financial support of the German Research Council DFG (grant SA 925/1), which made this research possible, and the DFG (grant 925/2) and the EU-commission (FP7 project CHLaSC, contract 28395) for additional support. Of course, I still remain fully responsible for any errors in this paper.

[1] The term *Binding* has several uses in cognitive science. In this paper, *Binding* always refers to the semantic phenomenon in sentence interpretation that is described in the main text. I use the term *binding* as a descriptive label for a possible interpretation of sentences like (1), though the term is originally associated with variable binding in the sense of first order predicate logic. Note that accounts of binding based on combinatory logic also include specific principles (the duplicator of Curry and Feys (1958), the Geach-rule of Jacobson (1999)) to add binding to an otherwise functional interpretation system. Therefore predication and binding are two distinct

binding is the central concern of this paper. My goal is to motivate a new approach to binding that, at the same time, involves fewer language specific assumptions than existing accounts and solves several outstanding empirical problems in the theory of binding.

Binding is a pervasive phenomenon in language. The basic phenomenon of binding is illustrated by the sentence in (1). (1) has an interpretation that entails that *a* called *a*'s mother for any boy *a*. To derive this interpretation straightforward functional application is insufficient because the inner argument position of the predicate *mother* and the outer argument position of the predicate *call* must be applied to the same individual.

(1) Every boy called his mother.

Therefore, the phenomenon of binding requires an enrichment of the inventory of concept combination principles. This poses the question what additional mechanism makes binding possible. The present paper provides a new answer for this question.

Existing approaches to binding in language start from the perspective of logic. The question of how concepts can be combined arises also—perhaps independently of language—in the theory of rationality and mathematical thought, and is one of the central questions of logic in this case. In this domain, the first worked out systems of concept combination were developed. In particular, a property akin to binding is part of two such systems:[2] predicate logic (Frege 1879), and combinatory logic (Schönfinkel 1924; Curry 1930). The existing work on binding in language all derives from one of these two approaches.[3] In the following, I use the terms *Index Binding* for work employing indexed variables following predicate logic, and *combinatory Binding* for work employing the techniques of combinatory logic. I present the basic assumptions of index binding and also discuss combinatory binding in the following subsections. In particular, I show that underlying both accounts is a position-based system of memory: this amounts to the assumption that human memory contains a language dedicated sub-system that stores recursive sequences where stored elements can be referred to by their position in the sequence.[4] This assumption is one that needs strong justification since we

mechanisms in such a system as well.

[2]Of course, this property is actually called *binding* in the case of predicate logic.

[3]This may not be fully accurate for some work closely connected to syntax, specifically by Higginbotham (1983).

[4]This claim is not explicit in the linguistic work. However, if one compares the linguistic

do not generally refer to items only by an abstract position.[5] The approach I develop is not directly based on either one of the approaches of logic. Most significantly it does not assume anything language specific about human memory for binding: Rather than requiring a recursively structured sequence, it only assumes that memory holds a set of items. Since a set is a flatter structure than a sequence, I call the model of binding I develop *Flat Binding*. The difference between the position-based view of memory underlying index and combinatory binding, on the one hand, and the flat binding account I develop, on the other, is illustrated by the pictures in (2): where as the position-based view assumes that memory consists of an ordered set of storage positions, the flat view does not require such an assumption.

(2) a. Position-based memory sequence:

b. Flat memory set:

My proposal for binding builds on two insights from the previous literature: The first is that elements that can be bound have been shown to have silent content in many cases (Fox 1999b; Sauerland 1998, 2004, in print; Jacobson 2000). The second insight is that the elements pronouns primarily refer to are not individuals, but individual concepts (or guises) that contain information about the stored elements (Heim 1992, 1998; Aloni 2001, 2005; Percus and Sauerland 2003).[6] As we will see below, these two insights combine fruitfully: Intuitively, since one says that bound phrases contain more information than previously assumed, while the other says that the stored elements contain more information, the two insights make it possible for the bound elements to uniquely describe exactly one of the stored elements. The objective of this paper is to show that the combination of the two insights is really powerful

work with other work on memory such as that presented by Miyake and Shat (1999) the special status of the position-based memory is apparent.

[5]From this perspective, sign languages are interesting. I am inclined to assume with McBurney (2002) that some real modality effects are found with pronouns in sign-languages, but have not been able to study the matter in sufficient detail yet.

[6]See also (Montague 1974[1970]; Lasersohn 2005).

enough to render storage sequences superfluous. The elimination of sequences then also provides a rationale for the underlying empirical insights: If there are no sequences or other recursive storage structures, bound elements must be descriptions that pick out an element from a set of stored items, and the items stored must come with some information about their properties. Therefore the elimination of position-based storage structures is not just an end for itself, but has further explanatory benefits.

In the next subsection, I introduce the classic approach to binding based on predicate logic in some more detail, mainly to later contrast specific properties with the flat binding approach. As the foundation of both approaches I assume a model-theoretic account of natural language semantics. This approach assumes that humans form a *Semantic Model*—an internal scene representation that represents their beliefs. The formation of the semantic model is affected by sensory stimulation, prior linguistic input, recollection of memories, and perhaps other factors. Furthermore humans form an internal representation of a sentence based on linguistic stimulation—a logical form of a sentence. The semantic mechanisms then map constituents of the logical form onto elements of the semantic model.

1.1 Position-Based Approaches to Binding

I use the term *Index Binding* to refer to the dominant approach to binding in natural language. Index binding is based on the work of Frege (1879) and Tarski (1936) in logic. Quine (1960), Montague (1974[1970]), and Lewis (1970) first applied the approach to natural language, but here I follow the elegant exposition of the textbook of Heim and Kratzer (1998). The three basic assumptions of the index-binding approach are the following:

(3)
i. the semantic model contains an *assignment sequence*
ii. bound elements bear abstract indices referring to positions of the sequence
iii. indexed λ-operators can modify specific positions of the assignment sequence

The assignment sequence referred to in (i) is a language particular part of human memory not needed anywhere else in the study of memory as the papers collected by Miyake and Shat (1999) show.[7] It is a function from a

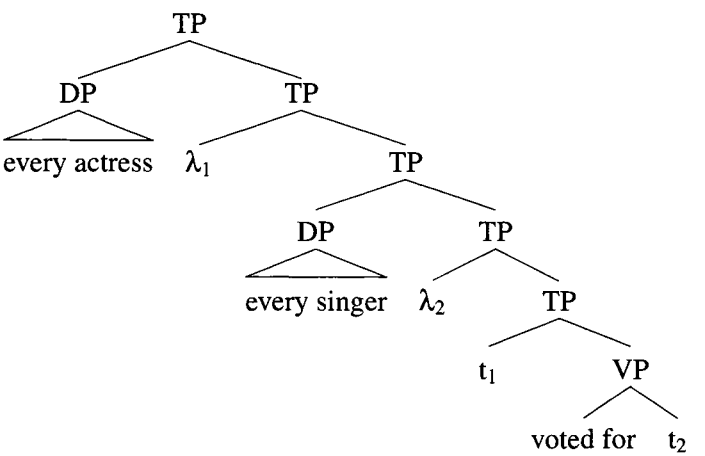
[7]Perhaps there would be another case of position-based memory if we understood the cognitive status of mathematical thought. However, I am sceptical about this and, in any case, it would not apply to non-mathematicians.

set of indices to entities. Bound elements such as pronouns and traces are represented as indices in logical form, and refer to a particular position of the assignment sequence. The abstract indices assigned to these are typically drawn from the set of natural numbers, though any infinite set would do. The assignment sequence maps these indices to values in the semantic model. The λ-operators also bear indices and extend or change the assignment function at the position of their index. Except for the application of the λ-rule, the assignment function is held constant during the evaluation of a sentence. The interpretation rules for bound elements and binders are given in (4):[8]

(4) a. $[\![pro_i]\!]^g = [\![t_i]\!]^g = g(i)$
b. $[\![\lambda_i\ X]\!]^g(a) = [\![X]\!]^{g[i \mapsto a]}$

To illustrate the index binding system, consider example (5). Phonological and syntactic processes form an abstract mental representation of a sentence, which the semantic mechanism applies to—the logical form. For concreteness, I assume the logical form depicted below (5) in this illustration of index binding.[9] In particular, the syntactic processes make sure that the λ-operators are co-indexed with their respective traces as shown in (5).

(5) Every actress voted for every singer.



[8]I use the notation $f[a \mapsto b]$ for a modified function. For a function f and any a and b, $f[a \mapsto b]$ is defined as function with domain($f[a \mapsto b]$) = domain(f) ∪ {a} and the value assignment $f[a \mapsto b](x)$ is equal to b if $x = a$ and equal to $f(x)$ otherwise.

[9]I assume here that quantifiers are never interpreted in argument positions, which is controversial but has some support (Heim and Kratzer 1998; Fox 2000; Yatsushiro 2001). If I did not make this assumption, it would not change anything substantial, but the example sentences would need to be longer.

Consider now how the interpretation mechanism applies to the representation in (5). To make things simple, assume that the assignment sequence is initially the empty sequence $\emptyset$. The quantifier *every actress* starts a set of computations as the complement is applied to each actress. The same holds for the second quantifier *every singer*. The two λ-operators both add one value to the assignment sequence. Index 1 is mapped to an actor a, and index 2 to a singer s, where the exact value of a and s depends on which of the computations the quantifiers initiated we are considering. The two traces, t_1 and t_2, are both evaluated relative to the assignment sequence $\{1 \mapsto a, 2 \mapsto s\}$, and therefore t_1 is assigned value a and t_2 value s. The indices ensure two properties of the interpretation of (5): For one, that the value stored by the same λ-operator is always assigned to a particular trace across the computations initiated by the quantifier. And secondly, that the value assigned is always that of the particular λ-operator the trace is coindexed with as a result of the syntactic processes. These two consequences of the indexation pattern in (5) ensure that the predicted interpretation corresponds to the bound construal of (5).

Index binding, as was already stated above, is presently by far the most popular approach to binding in language. It has been worked out in a lot of detail concerning the syntactic, semantic and pragmatic restrictions on co-indexation (see for instance Büring (2006) for a useful survey.) However, it has also been criticized. One kind of criticism has focussed on the assumption that indices are present. This has been viewed as undesirable for the syntactic structures (Chomsky 1981, 1995, and others) and for semantic models (Landman and Moerdijk 1983, and others). I agree with this criticism, in particular, in the case of the semantic models: The semantic model is a language-independent part of our conceptual system, and it would be surprising to find a language specific element in this domain on the view that language itself evolved very recently (Hauser et al. 2002). But the use of assignment functions relies on the assumption that there are language dedicated subsystems of the semantic model that have the recursive structure of the natural numbers. Other criticisms of index-binding have focussed on certain empirical issues, many of which will become relevant below (cf. Fox 1999b; Jacobson 1999; Safir 2005b; Sauerland 2000a, among others).

The second existing approach to binding, the combinatory approach, has been applied to language first by Geach (1972).[10] It is based on the elegant system of combinatory logic that can express the same array of distinctions as predicate logic (Curry and Feys 1958) without recourse to a separate assignment function. Instead it makes use of the sequence of open argument

[10]See also Hepple (1990, 1992); Jacobson (1999); Szabolcsi (1987), and others.

positions of a function. The approach hence crucially relies on a recursive set of n-ary functions for any number n. If the approach is applied to language, these functions are assumed to be part of the semantic model. Therefore, the combinatory approach, though it does not use indices, nevertheless relies also on the idea of a language specific part of memory where memorized items are recalled solely by reference to their position in this structure. In addition to this short-coming, some of the phenomena I address below, specifically what looks like agreement between a binder and a bound pronoun (Heim 1994, to appear), have not yet been addressed in this system as far as I know. For this reason and since the flat-binding system I develop is closer to the index-binding system, I will not address the combinatory approach in any further detail in this paper.

2 Flat Binding

2.1 Intuitions

The presentation of the flat binding system is structured into three parts. First, I outline the guiding intuitions of the approach. Secondly, I present the entire system in a compact form. Finally, I discuss numerous examples that illustrate many aspects of the system.

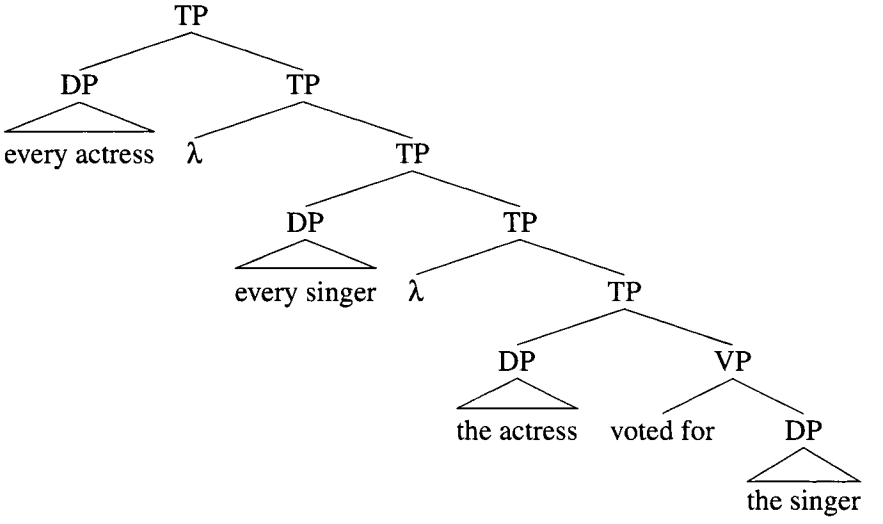
The central assumptions of my proposal are given in (6).

(6)
i. the semantic model contains a *memory set*
ii. bound elements are definite descriptions
iii. unindexed λ-operators can add elements to the memory set

The assumptions in (6) contrast with the three assumptions underlying index binding listed in (3) above. Specifically, (i) states that the role of the assignment sequence is taken on by the memory set—a flat structure where stored elements cannot be referred to by specific storage positions. The role of indices is taken on by definite descriptions. These pick out a stored element from the memory set by containing a property that uniquely describes an element. Finally the operation of adding an item to storage also does not require an index but just adds an element to a set.

For example (5), which I repeat in (7), a logical form representation in the flat binding system is shown below (7).

(7) Every actress voted for every singer.



This representation for (7) has important empirical support. As already mentioned in the introduction, the flat binding approach is based on two independent empirical observations about language. On the one hand, it has been argued that bound elements have in several cases been shown to have more descriptive content than evident from the phonological form. On the other hand, it has been argued that the referents of pronouns—a class of prototypical bound elements—are individual concepts (or guises). In the following, I present some partially novel motivation for these assumptions.

Consider first the evidence for the descriptive content of bound elements. There already is lot of evidence showing that elements that can be bound (i.e. traces and pronouns) can also contain silent lexical material (Evans 1977; Heim 1990; Sauerland 2000a, 2004, in print; Elbourne 2001, 2006). Furthermore, it is quite well-known that definite description can be bound in many cases, as illustrated by example (8) due to Schlenker (2005a).

(8) A linguist working on Binding Theory was so devoid of any moral sense that he forced a physicist working on particles to hire *the linguist*'s girlfriend in his lab.

The hypothesis that bound elements must always be definite descriptions at logical form is congruent with these two observations. However the hypothesis is not a necessary consequence of what we have seen so far: Even if some pronouns and traces have descriptive content and sometimes definite descriptions can be bound, bound pronoun that are not definite descriptions might also

exist. But at least two arguments show that bound pronouns must always have some descriptive content. I present one of these arguments based on focus in Sauerland (in print). The second argument for this conclusion is apparently novel, though it is straightforward. The argument comes from grammatical gender marking in languages like German and concerns pronouns: The gender of a noun phrase in German is determined by the head noun.[11] As we see in (9) the nouns for *spoon*, *fork*, and *knife* belong to different grammatical genders and the grammatical gender of the noun is morphologically marked on the indefinite and also on a pronoun that is anaphoric to the noun phrase.

(9) a. Tim hat einen Löffel gestohlen.
Tim has a.ACC.MASC.SG spoon stolen
Er war aus Gold.
pro.NOM.MASC.SG was of gold

b. Tim hat eine Gabel gestohlen. Sie
Tim has a.ACC.FEM.SG fork stolen *pro*.NOM.FEM.SG
war aus Gold.
was of gold

c. Tim hat ein Messer gestohlen.
Tim has a.ACC.NEUT.SG knife stolen
Es war aus Gold.
pro.NOM.NEUT.SG was of gold

The distribution of grammatical gender on pronouns must make reference to a noun. The most direct analysis is to assume that the pronoun contains a silent noun as shown in (10) (for (9a)). Then the same mechanism that transfers the gender of the noun to the entire noun phrase can apply in pronouns as well.

(10) $\overbrace{\text{er}_{\text{MASC}}\ [\text{der Löffel}_{\text{MASC}}]_{\Delta}}^{\text{elided}}$
$\underbrace{\phantom{\text{er}_{\text{MASC}}\ [\text{der Löffel}_{\text{MASC}}]}}_{\text{agreement}}$

The mechanism of transfer of gender marking from an elided noun also is at work in out-of-blue contexts, when a deictic pronoun or demonstrative is used. For example, a car can be either described as a *Auto* with neuter or as a *Wagen* with masculine gender in German, but there is no salient feminine noun for cars. Hence, either masculine or neuter can be used with deictic reference to

[11] I am simplifying and there is a lot of work on gender assignment (Enger et al. 2006; Yatsushiro and Sauerland 2006).

a car, but not feminine. In these cases, gender assignment as, for example, Corbett (2006) suggests could not be related to agreement since there is no antecedent.

(11) At the car-dealership, out of the blue.

a. Der/Das da gefällt mir.
the.MASC/the.NEUT there please me
'I like this one over there.'

b. *Die da gefällt mir.
The.FEM there pleases me

This argument for the presence of silent nouns in pronouns, though new for pronouns, is familiar from other cases of elided structure: For example in sluicing, the distribution of case marking on the remnant has been taken as evidence for silent, case assigning syntactic material (Ross 1969; Merchant 2001). The distribution of grammatical gender argues analogously for the presence of silent nominal material that assigns grammatical gender in pronouns.

The insight that all dependent elements are definite descriptions provides important support for the system I propose. In fact, representing definite descriptions as definite descriptions alone makes pronouns superfluous in simple examples like (12).

(12) Every boy said about every girl that he likes her.

Assume that the semantic evaluation of the sentences proceeds relative to a set of memorized items rather than an assignment sequence. The quantifiers *every boy* and *every girl* each add one item to this set; one is a boy and the other is girl. Of course, the quantifiers make us evaluate the complement clause for each boy and for each girl, but each of these evaluations is an independent process and in each process the memory will contain one boy and one girl. If the embedded clause is represented as (13) with the phrases *the boy* and *the girl* elided, no index positions are needed anymore: the description *the boy* selects the boy from the memory set in each evaluation process, and correspondingly *the girl* selects the girl.

(13) ... that he ~~the boy~~ likes her ~~the girl~~

But in general, something more than representing pronouns as definite descriptions is needed to capture binding: Example (12) has a special property,

which makes it particularly easy to account for it. The reason the definite descriptions *the boy* and *the girl* are sufficient to distinguish the two individuals, is that the two properties *boy* and *girl* are mutually exclusive: no individual is both a boy and girl. It is easy to come up with examples that do not have this property. Consider example (14) keeping in mind that there are people like Jennifer Lopez who are successful as both actress and singer.

(14) Every actress said about every singer that she likes her.

Sentence (14) has a reading where it entails that Jennifer Lopez said that she likes herself. This is shown by the fact that we can explicitly exempt this case by adding the exceptive *except Jennifer Lopez about herself* to (14). Example (14) also has other readings which we come back to in section 3.6. But for now, just consider the reading of (14) that entails self-liking of Jennifer Lopez.

If we apply the same strategy for (14) as in the analysis of (12), representation (15) is the representation of the embedded clause. Now, consider the semantic evaluation process initiated by the quantifiers with Nicole Kidman in memory as the actress under consideration and Jennifer Lopez as the singer under consideration. Because Jennifer Lopez is also an actress, the definite description *the actress* fails to uniquely refer—the sentence is predicted to result in a presupposition failure.

(15) ... that she ~~the actress~~ likes she ~~the singer~~

The presupposition failure arises in example (14) whenever there is a) a person like Jennifer Lopez who is both an actress and a singer, and b) there is at least one other actress or singer.

To get around the Jennifer-Lopez-problem, the system makes use of a second insight. This is the insight that memory always represents partial knowledge about the memorized items (Heim 1982, 1998)—in other words, what is stored are *Individual Concepts* rather than individuals. This is different from logical systems of binding where it is not necessary that knowledge be represented. Specifically, index binding is usually introduced in an extensional setting where knowledge is not represented. The need to represent partial knowledge about the memorized items has been recognized in the technical literature on binding independently of the Jennifer-Lopez-problem: In particular, Heim (1998) has argued this point (see also Thornton and Wexler 1999 for evidence from acquisition). Aloni (2001) and Percus and Sauerland (2003) provide further support from binding into modal and propositional attitude contexts. Heim's argument rests on an observation exemplified by the brief dialogue in (16).

(16) (Heim 1998, 213)[12]

a. Is this author Zelda?
b. How can you doubt it? She praises her to the sky. No competing candidate would do that.

Specifically, in the second sentence of (17b), the pronouns *she* and *her* occur in a syntactic configuration where the conditions of binding theory (for example, Condition B of Chomsky 1981) should block a binding and also a coreference relation of the two pronouns. Nevertheless, the information that the speaker of (16b) is trying to convey is that the referents of *she* and *her* are identical.

For the treatment of (16), Heim argues that it must be taken into account that the identity of Zelda and the speaker is under debate in the discourse above. Zelda and the author may be identical, but identity is not yet established with certainty. Such partial knowledge is commonly represented by recourse to possible worlds in model theoretic semantics. Following Heim, I adopt the possible worlds approach in the following, though, flat binding is in principle compatible with other approaches to knowledge representation as far as I can see. Within a possible worlds account, I make use of the notion of an *individual concept* (or *guise*). An individual concept is a function from a set of possible worlds onto individuals.

In example (16), two individual concepts are relevant. Both are defined for the context set in the sense of Stalnaker (1978): the set of possible worlds that are consistent with all the beliefs that are known to be shared by the two discourse participants in (16). In particular after the question (16a) the two discourse participants agree that there is a unique author and also a unique person named Zelda. But since it not yet known to be agreed that Zelda is actually the author, the context set contains worlds where Zelda and the author are two different individuals. Then, the following two concepts can be defined for the context set: the function that picks out the author in each world and set, and the function that picks out Zelda in each world. These two concepts are different functions since they yield different values for those worlds in the context set where Zelda is not the author. In the second sentence of (16b), Heim proposes that *she* refer to one of these two concepts—the one selecting the author—, and *her* to the other one. Since the two concepts are different, it is clear that binding theory is not violated in (16b).

Heim's argument shows that, at least in some cases, the items stored in memory are individual concepts rather than individuals; specifically, concepts

[12]Heim's example has *speaker* instead of *author*.

with the context set as their domain. Building on this assumption, I propose that at least for the sentence internal binding processes not only the context set is a possible domain for concepts, but also bigger sets of possible worlds. In fact, my solution to the Jennifer-Lopez-problem posed by (14) makes use of concepts the domain of which is as big as possible. (14) relies on a distinction between the concept of Jennifer Lopez as a singer and Jennifer Lopez as an actress even in a context where it is established that Jennifer Lopez is both. I introduce concepts that are have maximal domains for the properties *actress* and *singer* to this end. Observe first that global concepts—those defined for all possible worlds—will have only tautological properties since for any other property there are possible worlds where no entity will have that property. A maximal concept for a property has as its domain all the possible worlds that contain at least one object that has this property. Maximal concepts are useful since any concept maximal for property P has no other properties except for those that follow logically from P. If we view a concept as the lists of properties that the concept has wherever it is defined, a P-maximal concept is the list of just P and logical consequences of P. Relevant to the above example is the distinction between a concept x that assigns Jennifer Lopez to all worlds in the context set and is a maximal actress concept and a concept y that also assigns Jennifer Lopez to all worlds in the context set, but is a maximal singer concept. When (16) is evaluated relative to a memory set that contains y in addition to x or any other maximal singer concept, the definite description *the actress* will pick out x and not y, while the definite description *the singer* will pick out the other concept of the memory set. Hence, Jennifer Lopez can be part of the domain of quantification of both quantifiers, *every actress* and *every singer*, without leading to a presupposition failure.

At this point, the basic approach I am pursuing should be clear. Of course, many issues come up at this point. After all, binding is one of the most intensively studied and best understood areas of semantics. I try to address as many issues in this paper as possible in a formally explicit way. This way any potential problems can be seen clearly. To conclude this informal discussion, I briefly survey the major issues that are addressed below, and sketch the approach I develop in the formal part.

The first problem arising from the discussion of the Jennifer-Lopez-problem concerns sentences like (17), where two quantifiers with identical restrictors bind dependent elements in their scope. If both quantifiers range over maximal actress-concepts, there is no possible analysis of the two pronouns in the embedded clause as definite descriptions that could distinguish between the two maximal concepts so as to always select the one introduced

by the first or second quantifier. For example, if *she* or *her* is analyzed as the *the actress*, a presupposition failure will result.

(17) Every actress wrote about every actress that she likes her.

I solve his problem below by appealing to additional restrictors of the two quantifiers. The relevant interpretation has identical restrictions in the actual context, however, the restrictors may still differ as long as they are extensionally equivalent in the current context. Concretely the restrictor of the second occurrence of *every actress*, I will analyze as *actress possibly written about by the actress*. This is sufficient to bring about a difference of maximal concepts.[13] Furthermore, it explains that (17) could also be judged true if no actress wrote about herself, while it would be judged false when some but not all of the actresses wrote about themselves – it can be assumed only in the former situation that it is considered impossible that an actress writes about herself.

This leads directly to the second issue that needs to be addressed: Since I assume a lot of elided material in several place, I need a general account of ellipsis that predicts what material can be elided. In doing so, I have to turn away from the surface-anaphora analysis of Hankamer and Sag (1976) which assumes that elided material must always have a linguistic antecedent (see also Merchant (2005) for arguments against the surface-anaphora analysis). Instead I pursue an analysis of ellipsis stemming from the recoverability-up-to-deletion intuition (Chomsky and Lasnik 1993). In particular, I assume that the recoverability of reference licenses ellipsis. Recoverability of reference and recoverability of the lexical content differ clearly in the case of definite descriptions: if an actress with red hair and long legs is the most salient actress around, *the actress*, *the actress with red hair*, *the actress with long legs*, and *the actress with red hair and long legs* are lexically different, but referentially identical. In general, recoverability of reference for definite description is a less strict condition than recoverability of lexical content. I furthermore assume that the syntactic condition on the environment where ellipsis can take place is different from what is commonly assumed: I assume that ellipsis of adjuncts is always syntactically licensed. This predicts that in the example of the actress with red hair and long legs, all the definite descriptions listed before could be reduced to *the actress* by ellipsis. That adjunct ellipsis is

[13]To interpret *the actress* in the scope of both quantifiers, I introduce a preference for maximal concepts with a bigger domain – i.e. maximal actress-concepts rather than maximal actress-possibly-written-about-by-the-actress-concepts.

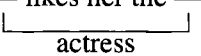
always possible follows naturally if the syntactic condition on ellipsis is not a licensing condition, but a blocking condition: For example, NP-ellipsis is blocked by phonologically weak determiners like *the*, *no*, and *a*.[14]

The necessary ellipsis in example (17) the shows us something about how (plural) individuals become salient in the course of processing a sentence – once we have processed *the actress wrote* the concept of individuals possibly written to by the actress must be salient. While I cannot discuss this assumption in further detail in this paper, it seems related to work in real-time processing such as Tanenhaus et al. (1995). Note also that work such as that of Stanley and Szabo (2000) and Marti (2003), which assumes silent NP-variables for the explanation of domain restriction, provides no account of which values these NP-variables may have in a particular context. Analyzing domain restriction as an instance of ellipsis is far more explicit on the licensing conditions.

A third issue, which I address briefly below and in more detail in other work (Sauerland 2007), are sloppy readings in ellipsis. The problem here is the following: If *her* is represented as *the actress* in the first conjunct and as *the singer* in the elided second conjunct, we need to understand why the analysis of VP-ellipsis licenses this switch of the definite description.

(18) The actress likes her agent and the singer does ~~like her agent~~ too.

I assume that, for this case, a syntactic representation that assumes structure sharing as in (19) (for the first conjunct of (18)): the NP *actress* is linked to two syntactic positions simultaneously.

(19) the — likes her the — agent


As I argue in the paper (Sauerland 2007), this analysis of sloppy interpretations compares favorably to the standard analysis assuming the binding/coreference distincting Reinhart (1983): All correct predictions are maintained, and one advantage of my analysis is that it is not restricted by c-command, which, for example, Tomioka (1999) argues to be correct for sloppy readings.

The fourth and final major issue I pursue below is the analysis of apparent agreement between a bound pronoun and its binder (Heim to appear). This issue arises in particular with indexical pronouns as in (20a), *de se*-pronouns

[14]Obviously, *every* is the one determiner in English that does not license NP-ellipsis, but is not phonologically weak.

as in (20b), and distributive plurals as in (20c).[15] In each of these cases, the bound pronoun bears a ϕ-feature which does not seem to be semantically licensed: In (20a), bound *my* is marked first person, but alternatives other than the first person are considered. In (20b), bound *they* is marked plural, but it is pragmatically clear that *they* always refers to a single boy, not a plural one. And in (20c), *he* must be a woman in all of John's dream worlds, but nevertheless the masculine pronoun must be used.

(20) a. Only I like my passport picture (nobody else likes his passport picture).
b. The boys each think that they are the only ones in the room.
c. John dreamt that he is woman.

In Heim's account of the data in (20), the bound item and the binder are co-indexed. Heim then states a syntactic agreement rule that ensures that binder and bound item have the same features. Since the flat binding account does not make use of indices, it is not possible to state such an agreement rule. The strategy I adopt below to the agreement facts in (20) is non-uniform. I assume that in (20b) plurality is checked not for the actual referent of *they* ~~*the boy*~~*, which is a number neutral definite description, but for the referent of *they* ~~*the boy*~~* under slightly different premises, changed so that *they* ~~*the boy*~~* refers to the group of boys that *the boys* also refers to. For (20c), I assume that agreement is checked not for the actual referent of the pronoun, but for its counterpart in another set of worlds: John. Finally, I assume a syntactic mechanism for (20a).

2.2 The Formal System

To make the theory verifiable and have predictive power, I now present a formal system implementing the intuitions spelled out in the previous section. Since the formal system depends on a number of interlocking assumptions, I first present the entire formal system in a compact form and in the next section go through several examples that show how these assumptions interact in the working system and how at least isolated changes within this system would

[15] Kratzer (2006) argues that in addition there is an optional agreement rule for predicatively used relative clauses. Probably even more such rules are needed than either Heim's or Kratzer's account assume. Consider for example (i), which allows an interpretation analogous to the bound reading of (20a).

(i) The picture that only I like is my passport picture.

not yield the same results. It might still be possible and is probably even likely that, based on similar assumptions, one could come up with a working system by changing several assumptions at once.

The system is situated within the framework of model-theoretic semantics as outlined in the introduction above: In a nutshell, it assumes that *Logical Forms*, phrase structure representations generated by the mind, are related by a recursive procedure to entities in the *Semantic Model*, a non-linguistic, scene representation also generated by the mind. To what extent the system I develop relies on any of the specific assumptions of this framework is up to further investigation, but not a question I will pursue here.

Logical Forms A logical form structure, I assume, is a directed graph where the relation 'x points to y' is a partial order of the set of nodes. *x immediately dominates y* is another way of saying that x points to y. The transitive closure of immediate domination is *domination*. The binary branching assumption, which I adopt, states that each node either immediately dominates two nodes or none. A node that does not immediately dominate any other nodes is a *terminal*. Terminals either contain a lexical item or a λ-operator. There must be one node that dominates all others: the *Root Node*. One important consequence of these assumptions that is particularly debated amongst syntacticians is the following (Gärtner 1999, 2002): *Structure Sharing* is specifically allowed. This means that two phrases may immediately dominate the same node. Structure sharing is important for my account of sloppy interpretations with traces and pronouns.

The Semantic Model For the semantic model, the flat binding system assumes a familiar intensional system of types with s the type of possible worlds, e the type of individuals and events, t the type of truth values, and $\langle a,b\rangle$ the type of functions from the domain of type a to the domain of type b.[16] In particular, one-place properties are functions of type $\langle e,\langle s,t\rangle\rangle$.[17] I use possible worlds to model partial knowledge. I assume that possible worlds are *centered*.

[16] At this point, it is simpler to adopt the standard recursive type system, though dissatisfying. I do it only for concreteness. I assume that, with improved understanding, a finite set of semantic types is seen to be sufficient for natural language. Particularly attractive from a conceptual point of view would be the position that all arguments of the same predicate must be of different basic types since then the arguments of a predicate would be distinguished by type rendering order unnecessary.

[17] These could be either partial or total functions. In the following, I often will not consider the case of partial properties for simplicity's sake.

I implement this in the following way: there is a function *self* defined for all possible worlds such that *self* applied to a possible worlds w yields the individual a which w is viewed from.

Some lexical entries are the following. Note that the world parameter w of the evaluation function does not directly determine the world of evaluation for a lexical predicate.

(21) a. $[\![\text{actress}]\!]^{m,s,C,w} = \lambda x \in D_e \,.\, \lambda w \in D_s \,.\, x$ is an actress in w
b. $[\![\text{sleeps}]\!]^{m,s,C,w} = \lambda x \in D_e \,.\, \lambda w \in D_s \,.\, x$ sleeps in w

At least for the verb *sleeps*, temporal information would still need to be added to the lexical entry in (21b). Currently there is no need to represent this here and standard treatments like Ogihara (1996) of tense should carry over as far as I can see. I assume that tensed clauses denote one-place properties of events, and therefore both NPs and TPs denote one-place properties. This is relevant for the section on ellipsis licensing.

Concepts concepts are partial functions of type $\langle s,e\rangle$. $\mathcal{K}$ is the set of all concepts. I will use the terminology that concept x *has property* P if and only if $P(x(w))(w)$ holds wherever it is defined. This is captured by the following definition of $P(x)$:

(22) For $P \in D_{\langle e,\langle s,t\rangle\rangle}$ and $x \in \mathcal{K}$: $P(x) = 1$ iff. $\forall w \in \text{domain}(x)$

$$w \in \text{domain}(P(x(w))) \rightarrow P(x(w))(w) = 1$$

Two kinds of concepts are of particular interest for the following: *maximal* and *contextual* concepts. For any property P the set of maximal concepts is defined as follows:

(23) For any $P \in D_{\langle e,\langle s,t\rangle\rangle}$: $\max(P) =$

$$\{x \in \mathcal{K} \mid P(x) = 1 \text{ and } \text{domain}(x) = \{w \mid \exists y \in \mathcal{K}: P(y(w))(w) = 1\}$$

The *Context Set C* is defined as in Stalnaker (1978) as the set of all possible worlds that are congruent with all publicly established joint beliefs.[18] A concept x is a *Contextual Concept* iff. domain(x) is the context set. For technical reasons, I assume that contextual concepts are never also maximal concepts for any of the properties under consideration. In other words, I

[18] Sometimes this set is also referred to as the *Common Ground*.

assume that the context set is always such that there is no expressible property P for which $C = \{w \mid \exists y \in \mathcal{K}: P(y(w))(w) = 1\}$. The assumption seems innocent to me for expressible properties: If I believe that a P might exist, there is also a world inconsistent with our common beliefs where a P exists – this just requires changing something that is part of our common belief which is entirely unrelated to Ps.

Evaluation Parameters/Salience The recursive semantic evaluation function takes four parameters: m, s, C, and w. The notation for the evaluation function with these parameters is $[\![—]\!]^{m,s,C,w}$. m, the *Memory Set*, is a set of concepts. s, the *salience function for m*, is a function from m to the rational numbers or some subset thereof.[19] C is the context set, i.e. a set of possible worlds. w, the current evaluation world, is an element of C.

The salience function plays a very minor role in the present account: it is only used for number agreement with distributive quantification. For the following, I assume that there are two salience values—medium and high—with high > medium. In addition to the salience captured by the salience function, a second concept of salience is defined by reference to the domain of a concept: to ensure that clause-internal binders are more salient than clause-external ones, I assume that the bigger the domain of a concept the more salient it will be. To distinguish the two notions of salience, I speak of *s-salience* determined by the salience function and *d-salience* determined by the domains of concepts. A concept x is *more s-salient* than y iff. $s(x) > s(y)$. A concept x is *more d-salient* than y iff. $\text{domain}(x) \supset \text{domain}(y)$.

Definite Descriptions I distinguish between normal occurrences and distributively binding occurrences of definite descriptions. Under both kinds, I include occurrences that are pronounced as pronouns or remain completely unpronounced as traces. The distributively binding occurrences are marked by the feature DIST. Normal occurrences of definite descriptions are evaluated relative to the memory set m and select the most-salient concept in m that satisfies the description. For the account of plural agreement, I assume here that d-salience outranks s-salience. The denotation of a definite description is defined as follows:[20]

[19] It would be possible to define m as the domain of s. For expository reasons, I chose not to do this.

[20] Recall that the application of the property $[\![\text{NP}]\!]^{m,s,C,w}$ to the concept x in (24) is defined in (22) above. The $\oplus$-Operator in (24c) refers to the mereological sum operation as used in the semantics of plurality (see e.g. Schwarzschild 1996).

(24) The denotation $[\![\text{the NP}]\!]^{m,s,C,w}$ is the concept x that satisfies one of the following conditions:

a. $x \in m$ and $[\![\text{NP}]\!]^{m,s,C,w}(x) = 1$ and $\forall y \in m: [\![\text{NP}]\!]^{m,s,C,w}(y) = 1 \Rightarrow \text{domain}(y) \subsetneq \text{domain}(x)$

b. $x \in m$ and $[\![\text{NP}]\!]^{m,s,C,w}(x) = 1$ and $\forall y \in m: [\![\text{NP}]\!]^{m,s,C,w}(y) = 1 \Rightarrow s(y) < s(x)$ and no $x \in m$ satisfies (24a)

c. $\text{domain}(x) = C$ and $\forall w \in C: x(w) = \bigoplus_{a \in D_e \wedge [\![\text{NP}]\!]^{m,s,C,w}(a)(w)=1} a$ and no $x \in m$ satisfies (24a) or (24b)

For distributive quantification over a plural definite description, the logical form is *[the*$_{\text{DIST}}$ *NP] XP*. In the scope of distributive plural, both the plurality it refers to and one singular part of this plurality are available as referents. To make the plurality available, *the*$_{\text{DIST}}$ *NP* must add the plural concept $[\![\text{the NP}]\!]^{m,s,C,w}$ to m. For my account of plural agreement, this plural concept must receive high salience: I.e. the salience function is extended to $s' = s[[\![\text{the NP}]\!]^{m,s,C,w} \mapsto \text{high}]$. Distributive quantification has to range over a set of maximal singular concepts that covers the plural concept.[21] Since NP might entail plurality—e.g. *three pigs* does—I assume that the distributive quantification ranges over maximal singular concepts for the property denoted by some subconstituent of NP. I use the term *singular* for the property of being one atomic part. This leads to the following definition:

(25) *[the*$_{\text{DIST}}$ *NP] XP* is true iff. there is a set of concepts B that each are maximal for $[\![\text{NP}']\!]^{m,s,C,w}$ where NP′ is a subconstituent of NP and $\text{singular}(b) = 1$ for all $b \in B$ such that for all $w \in C$ $[\![\text{the NP}]\!]^{m,s,C,w} = \bigoplus_{b \in B} b(w)$ and the following holds:

$$\forall b \in B: [\![\text{XP}]\!]^{m \cup \{[\![\text{the NP}]\!]^{m,s,C,w}\},s',w,C}(b) = 1$$

Lambda Operator The unindexed λ-operator adds a concept to the memory set and creates functions that take concepts as their arguments. The salience of the concept λ adds to m is medium, unless it was already an element of m. The interpretation rule for λ is: [22]

$$(26) \quad [\![\lambda\ \text{XP}]\!]^{m,s,C,w} = \begin{cases} \lambda x^{\langle s,e\rangle}.[\![\text{XP}]\!]^{m,s,C,w} & \text{if } x \in m \\ \lambda x^{\langle s,e\rangle}.[\![\text{XP}]\!]^{m \cup \{x\},s[x \mapsto \text{medium}],C,w} & \text{otherwise} \end{cases}$$

[21] Cf. Schwarzschild's (1999) notion of a cover.

[22] Recall from footnote 8 that I use the notation $f[a \mapsto b]$ for the function derived from f by modification in point a.

Universal Quantification The universal quantifier ranges over maximal concepts of single individuals. The sister constituent of a universally quantified DP will always be initiated by a l-operator, which will add each concept the universal ranges over to the memory set. I assume that *every NP* like distributive plurals also adds the concept $[\![\text{the NP}]\!]^{m,s,C,w}$, which will usually be plural, to the memory set m. For the added plural concept, *every* also extends the salience function, but in difference to the distributive plural with medium salience: $s' = s[[\![\text{the NP}]\!]^{m,s,C,w} \mapsto \text{medium}]$. In sum, the lexical entry for *every* is the following:

(27) $[\![\text{every NP XP}]\!]^{m,s,C,w} = 1$ holds iff. the following holds:

$$\forall x \in \mathcal{K}: \quad (x \in \max([\![\text{NP}]\!]^{m,s,C,w}) \wedge \text{singular}(x)) \rightarrow [\![\text{XP}]\!]^{m \cup \{[\![\text{the NP}]\!]^{m,s,C,w}\},s',w,C}(x)(w) = 1$$

Function Application Rule The composition rule for the complex phrase [X Y] depends on the types of $[\![\text{X}]\!]^{m,s,C,w}$ and $[\![\text{Y}]\!]^{m,s,C,w}$. If possible, X and Y are combined by standard functional application: $[\![\text{X Y}]\!]^{m,s,C,w}$ is $[\![\text{X}]\!]^{m,s,C,w}([\![\text{Y}]\!]^{m,s,C,w})$ or $[\![\text{Y}]\!]^{m,s,C,w}([\![\text{X}]\!]^{m,s,C,w})$, depending on which is defined. There are two new composition rules, though, for cases where functional application is inapplicable: the relative clause rule and the concept application rule.

Relative Clause Rule A relative clause RC of type $\langle\langle s,e\rangle,t\rangle$ needs to be combined with an NP of type $\langle e,\langle s,t\rangle\rangle$. They are to combined the following rule to yield a modified NP meaning of type $\langle e,\langle s,t\rangle\rangle$:

(28) $[\![\text{NP RC}]\!](a)(w) = 1$ iff. $\text{NP}(a)(w)$ and there is a concept x maximal for NP such that $x(w) = a$ and $\text{RC}(x)$.

Concept Application Rule Since a definite description refer to a concept x and verbs denote properties, functional application cannot combine a verb with its arguments when they are definite. Rather, the concept x must be used to determine an individual in the current world of evaluation. Depending on whether x is defined for the current context set or not a *de dicto-* or *de re-*interpretation results, which I implement in a way inspired by Kaplan's (1968) proposal: For the *de re*-interpretation, the concept x must be *acquaintance-related* to another concept y that is defined for the context set. This notion is

defined as follows (For a function x and $S \subset \text{domain}(x)$, I use the notation $x \restriction_S$ for the function x' with $\text{domain}(x') = S$ and $x'(w) = x(w)$ for all $w \in S$.):[23]

(29) Two concepts x and y are *acquaintance-related* in evaluation state m,s,C,w iff. there is a definite description DP such that the following two conditions hold:

$$\exists D(x \restriction_D \in m \wedge \forall w \in D[\![\text{DP}]\!]^{m,s,C,w}(w) = x(w))$$
$$\exists D'(y \restriction_D \in m \wedge \forall v \in D'[\![\text{DP}]\!]^{m,s,C,w}(v) = y(v))$$

Concept application can now be defined as follows:

(30) If the type of $[\![\text{X}]\!]^{m,s,C,w}$ is $\langle \text{s,e} \rangle$ and that of $[\![\text{Y}]\!]^{m,s,C,w}$ is $\langle \text{e},\alpha \rangle$ with α any type, then $[\![\text{X Y}]\!]^{m,s,C,w}$ is defined as

a. $[\![\text{Y}]\!]^{m,s,C,w}([\![\text{X}]\!]^{m,s,C,w}(w))$, if $\text{domain}([\![\text{X}]\!]^{m,s,C,w}) \supset C$
b. $[\![\text{Y}]\!]^{m,s,C,w}(y(w))$, if $\text{domain}([\![\text{X}]\!]^{m,s,C,w}) \not\supset C$ and $y \in CS$ with $\text{domain}(y) = C$ and $[\![\text{X}]\!]^{m,s,C,w}$ and y are acquaintance-related

Propositional Attitudes The only propositional attitude verb I consider in this paper is *believe*. However, the analysis can carry over to other attitude verbs as far as I can see. $\text{Dox}(a,w)$ is the set of centered worlds compatible with the beliefs of individual a in w. For *de se*-reference, the complement of *believe* is evaluated relative to a changed memory set m' where the self-concept of the believer is added. Furthermore, In addition to introducing a referent for *de se*-pronouns, there is a condition on m' that makes sure that concepts introduced by a quantifier (i.e. maximal concepts) are restricted so that they must be interpreted *de re* in the scope of *believe*.[24] In sum, m' is specified as follows:[25]

$$\begin{aligned} m' \;=\; & \{\text{self} \restriction_{\text{Dox}(a,w)}\} \cup \{x \mid x \in m \wedge \text{Dox}(a,w) \cup C \not\subset \text{domain}(x)\} \\ & \cup \{x \restriction_{\text{domain}(x) \setminus \text{Dox}(a,w)} \mid x \in m \wedge \text{Dox}(a,w) \cup C \subset \text{domain}(x)\} \end{aligned}$$

[23] The definition of *acquaintance related* in (i) is intended to capture what Kaplan (1968) calls *vivid* acquaintance.

[24] Without this restriction mechanism, the interpretation of (i) would entail that Mary believes that every boy (de dicto) sleeps, if Mary believes that there is at least one boy.

(i) Mary believes about every boy that he sleeps.

[25] Recall that I assume that all possible worlds come with a perspective and that self(w) selects the individual whose perspective a world w is from.

The salience function s' needs to be adjusted for domain m' on the basis of s, but at present it does not matter for any of the examples I consider below what values s' assigns. Hence, I leave this open for the time being. The lexical entry of *believe* is the following:

(31) $[\![\text{believe CP}]\!]^{m,s,C,w}(a) = 1$ iff. the following holds:

$$\forall w' \in \text{Dox}(a,w): [\![\text{CP}]\!]^{m',s',\text{Dox}(a,w),w'} = 1$$

Pronominalization/Ellipsis The rules for pronominalization and ellipsis licensing assume the syntactic structure of nominals by Sauerland (2003): a ϕ-head takes as its complement DP, which in turn takes an NP complement. Using the notation XP[YP] for an XP containing head X and complement YP, the nominal structure is ϕP[DP[NP]]. I analyze pronouns as a ϕP with DP elided. When DP is at least in part pronounced, the ϕ-head is not pronounced. I assume that unpronounced constituents are marked by the feature Δ and focused constituents by the feature F. No occurrence of F may be dominated by an occurrence of Δ. The phonological matrix of a phrase XP, PF(XP), is the pronounced form of XP. Generally this is the string of lexical terminals dominated by XP and not dominated by any Δ unless special pronunciation rules apply. The one relevant pronunciation rule is that PF(ϕP[DP$_\Delta$]) is the pronoun with the agreement features ϕ. XP_1 and XP_2 are *ellipsis alternatives* if $PF(XP_1) = PF(XP_2)$. XP_1 and XP_2 are focus alternatives iff. all material in XP_1 and XP_2 not dominated by an occurrence of F is identical. Deletion must be licensed by a *Focus Domain*, which is marked by $\sim$ (Rooth 1992b). I provide a semantics for $\sim$ applying to NP, ϕP and TP. In addition, deletion is also subject to a syntactic condition: While most researchers assume that there is a syntactic licensing condition (e.g. Lobeck 1995), I assume there is only a syntactic blocking condition: any phrase can be deleted except when it is selected by a head that blocks deletion of its complement. In particular, the English determiners *a, the, no* and *every* block deletion of their complement.[26] Deletion of adjuncts, however, is never syntactically blocked in my analysis and therefore only subject to semantic/pragmatic licensing. The semantic licensing conditions for focus domains capture the intuition that deletion

[26]The property of blocking deletion of the full complement can probably be derived for *a, the* and *no*: it might well follow from the fact that these determiners are phonologically weak in English. The German translations *ein, kein*, und *der* all do not block full NP-deletion of their complement. That *every* blocks deletion in English, however, does not follow from the phonology, and must be stipulated in the lexical entry.

must be recoverable depending on the salience of stored discourse entities. There are two licensing conditions, one based on s-salience derived from the salience function s as stated above, the other based on d-salience, which makes reference to the size of the domain of a concept (see page 216). I first define the notions of a maximally s- or d-salient focus alternative as follows:

(32) For any XP denoting a one-place property and for x either s or d, I define the *maximally x-salient focus alternative* in evaluation state m,s,C,w as the XP′ of the focus alternatives of XP such that $[\![\text{the XP}']\!]^{m,s,C,w}$ is defined and for any other focus alternative XP″ of XP either $[\![\text{theXP}'']\!]^{m,s,C,w}$ is undefined or $[\![\text{the XP}'']\!]^{m,s,C,w}$ is less x-salient than $[\![\text{the XP}']\!]^{m,s,C,w}$.

Now the licensing condition of the focus domain is captured as follows for NP, DP and TP:

(33) a. The focus domain $[\![\sim \text{NP}]\!]^{m,s,C,w}$ presupposes that for $x = s$ or $x = \text{d}$ the following holds: there is a maximally x-salient focus alternative NP′ of NP and for any ellipsis alternative $\widetilde{\text{NP}}$ of NP, if $\widetilde{\text{NP}}$ has a maximally x-salient focus alternative $\widetilde{\text{NP}}'$, then $[\![\widetilde{\text{NP}}']\!]^{m,s,C,w}$ is less x-salient than $[\![\text{NP}']\!]^{m,s,C,w}$.

b. The focus domain $[\![\sim \phi\text{P}[\text{DP}[\text{NP}]]]\!]^{m,s,C,w}$ presupposes that D is the definite determiner and that for $x = s$ or $x = \text{d}$ the following holds: there is a maximally x-salient focus alternative NP′ of NP and for any ellipsis alternative $\widetilde{\phi\text{P}[\text{DP}[\text{NP}]]}$ of $\phi\text{P}[\text{DP}[\text{NP}]]$, if $\widetilde{\text{NP}}$ has a maximally x-salient focus alternative $\widetilde{\text{NP}}'$, then $[\![\widetilde{\text{NP}}']\!]^{m,s,C,w}$ is less x-salient than $[\![\text{NP}']\!]^{m,s,C,w}$.

c. The focus domain $[\![\sim \text{TP}]\!]^{m,s,C,w}$ presupposes that for x being either s or d the following holds: there is a maximally salient focus alternative TP′ of TP and for any ellipsis alternative $\widetilde{\text{TP}}$ of TP, if $\widetilde{\text{TP}}$ has a maximally salient focus alternative $\widetilde{\text{TP}}'$, then $[\![\widetilde{\text{TP}}']\!]^{m,s,C,w}$ is less salient than $[\![\text{TP}']\!]^{m,s,C,w}$.

Traces Traces, I analyze as ϕP marked with Δ'. Δ' has the same phonological effect as Δ and must also not dominate any F, but it is not subject to the semantic licensing condition. Rather, I assume that Δ' is licensed by some structural mechanism of the type investigated in syntax. I assume that as a

consequence of this syntactic licensing, the NP in ϕP must always be structure shared with another occurrence of NP.[27]

Phi-Features For agreement licensing I assume that non-definite DPs must generally undergo movement from the complement of ϕP-position leaving behind a definite trace. I assume that NP has a number-neutral interpretation and that the choice between the singular and plural form is determined by agreement with ϕP (Sauerland 2003). In the following, I write *NP** to indicate the number-neutral interpretation whenever it is relevant. For definites, I assume the licensing presupposition in (34).

(34) $[\![\phi P[\text{the NP}]]\!]^{m,s,C,w}$ is licensed if x has the ϕ-properties, where x is a concept that satisfies one of the following conditions:

a. $x \in m$ and $[\![\text{NP}]\!]^{m,s,C,w}(x) = 1$ and $\forall y \in m : [\![\text{NP}]\!]^{m,s,C,w}(y) = 1 \Rightarrow s(y) < s(x)$

b. $x \in m$ and $[\![\text{NP}]\!]^{m,s,C,w}(x) = 1$ and $\forall y \in m : [\![\text{NP}]\!]^{m,s,C,w}(y) = 1 \Rightarrow \text{domain}(y) \subsetneqq \text{domain}(x)$ and no $x \in m$ satisfies (34a)

c. $\text{domain}(x) = C$ and $\forall w \in Cx(w) = \bigoplus_{a \in D_e \wedge [\![\text{NP}]\!]^{m,s,C,w}(a)(w)=1} a$ and no $x \in m$ satisfies (34a) or (34b)

Note that the order of s- and d-salience is the reverse for agreement licensing and for the reference of a definite description. This will predict the plural number agreement under distributive definite descriptions. The number-properties, specifically, are defined as follows:[28]

(35) a. $[\![\text{SING}]\!]^{m,s,C,w}(a)(w)$ is licensed iff. a is an atomic object in w

b. $[\![\text{PLUR}]\!]^{m,s,C,w}(a)(w)$ is licensed iff. a is not an atomic object in w

Indexicality A discourse state is a triplet m, s, C of a memory set, salience function, and context set. For reference of the first person in English, the self-concept of the root context must refer to the current speaker: I.e. I assume that the initial memory set must always contain the concept x which for any

[27] I assume that logical form structures can only contain movement structures of categories that denote concepts. These are DP and TP in the present theory (cf. Heycock 1995). Of these, only DP occurs in the examples discussed below.

[28] For my present purposes it is not important whether the interpretation of the plural in (35b) is primitive or derived by some pragmatic mechanism as Sauerland et al. (2005) argue. Both approaches are consistent with the analysis developed here.

w in C selects self(w), and C must be such that for all $w \in C$, self(w) is the unique a in w with $[\![\text{current speaker}]\!]^{m,s,C,w}$(a)(w). Then, the presupposition of the first person feature in English is the following:

(36) $[\![1\text{ST}]\!]^{m,s,C,w}(x)(w)$ is licensed iff. the following two conditions hold:

a. domain(x) = domain($[\![\text{the current speaker}]\!]^{m,s,C,w}$) and
b. $[\![\text{the current speaker}]\!]^{m,s,C,w}(v)$ is part of $x(v)$ for all $v \in \text{domain}(x)$

***De se*-Occurrences** For *de se*-occurrences in English, I assume the general pronunciation rule (37). Current work on other languages indicates that this pronunciation rule is English specific (cf. Schlenker 1999; Anand and Nevins 2004)

(37) PF(*the self related by acquaintance to* ϕP) = PF(ϕP)

This concludes the specification of the formal system. In the following section, we look at some examples that simultaneously show how the system works and provide a justification for some specific assumptions of the above.

3 Examples

3.1 The Basic Case

The basic case of a bound interpretation I considered in the introduction is repeated in (38).

(38) Every boy called his mother.

The logical form representation of (38) must contain a full DP in place of the pronoun since pronouns can only be derived by DP-ellipsis. One representation in the present system that yields the bound interpretation in many contexts is (39).[29]

(39) Every boy λ ϕP[the boy]$_{\Delta'}$ likes ϕP[the boy]$_{\Delta}$'s mother

[29]Here and in the following, I do not represent structure sharing unless it is important for the point under consideration. Since I indicate traces with $[]_{\Delta'}$, structure sharing relationships can be reconstructed. In (39), the first two occurrences of *boy* would be structure shared and the ϕP in subject position would be marked by Δ'.

Assume for example that (39) is interpreted relative to some context set C and that the memory set is empty except for the self-concept. The universal quantifier *every boy* following (27) is interpreted as the requirement that, for every maximal singular boy-concept, the predicate λ *likes [the boy]'s mother* be true. For this evaluation, the contextual concept x_c of the boys is also added to the memory set by *every* with medium salience. So, the definite description *the boy* is evaluated relative to the memory set m containing the maximal single boy concept x_b and the contextual concept of the boys x_c. Applying (24), the reference of both occurrences of *the boy* is x_b since x_b has a bigger domain than x_c. Furthermore, singular agreement is licensed on both occurrences of *the boy* according to (34) because x_b is equally s-salient, but more d-salient than x_c. Finally, deletion of the second occurrence of *the boy* is licensed at the DP-level in (33) because no other definite DP selects a more salient concept than x_b. I assume that $[\![\text{mother}]\!]^{m,s,C,w}$ is evaluated as the predicate m of type $\langle e,\langle e,\langle s,t\rangle\rangle\rangle$. The concept application rule (30) does not apply m directly to x_b, but first applies x_b to the actual world w, which is possible because $C \subset \text{domain}(x_b)$, and then applies m to the result. The reference of *the boy's mother* is therefore the contextual concept x_m of the mother of $x_b(w)$ for each $w \in C$. The verb *likes* is then applied to the two concepts x_m and x_b again by the concept application rule (30), which leads to the computation of $[\![\text{likes}]\!]^{m,s,C,w}(x_m(w))(x_b(w))$. This amounts to the proposition that $x_b(w)$ likes $x_m(w)$. Since for each boy a in a world w there is at least one maximal boy concept x_b, we see now that representation (39) is true iff. for all worlds $w \in C$, every boy in w likes his mother.

In fact, there will be many maximal boy-concepts x_b for each boy a in a world w such that $x_b(w) = a$. The one-to-many correspondence among individuals in the actual world and relevant maximal concepts is of interest for the understanding of quantification in language. Aloni (2001) observes that this one-to-many correspondence makes it difficult to treat many determiners as generalized quantifiers. In the case of the universal quantifier, however, the one-to-many correspondence is without truth-conditional consequences: The truth-conditions predicted for (38) with the lexical entry for *every* in (27) are equivalent to the standard truth-conditions. Aloni proposes to restrict the set of concepts available for quantification to guarantee a one-to-one correspondence between individuals in the actual world and relevant maximal concepts. However, her sole motivation is to allow all generalized quantifiers to range over concepts. If I had faith in the generalized quantifier analysis, I could follow Aloni. However, I am more inclined to believe that there are no quantifiers in natural language other than the universal and perhaps the

existential (cf. Hackl (2000); Krifka (2000); Sauerland (2000b)). If this is correct, Aloni's restriction is not necessary. In fact, abandoning it may help to explain the scarcity of quantifiers in natural language (see below).

One other logical form representation of the bound interpretation of (38) in the empty context contains *the single boy* instead of the *the boy*. This has exactly the same truth conditions of (39). Since I assume that ellipsis licensing only requires the recoverability of reference, there are generally several possibilities to reconstruct ellipsis.

(40) Every boy λ ϕP[the boy]$_{\Delta'}$ likes ϕP[the single boy]$_{\Delta}$'s mother

Next, compare (38) to (41), which does not allow a bound variable interpretation of the pronoun *their*. Rather, *their* can only be interpreted as referring to the plurality of all the boys.

(41) Every boy called their mother.

One representation that derives this interpretation is (42).

(42) Every boy λ ϕP[the boy]$_{\Delta'}$ called ϕP[the numerous boys]$_{\Delta}$'s mother

The initial steps of the semantic evaluation of (42) are the same as for (39). A difference occurs only at the pronoun: The DP *the numerous boys* selects x_c rather than x_b because x_b does not have the property *numerous*. Ellipsis of the *the numerous boys* is licensed at the level of the ϕP, since, for the plural agreement marking *their*, no ellipsis alternative exists that selects a concept other than x_c.

Now, consider (38) in richer contexts where other representations for (41) also lead to a bound interpretation. Consider a context where C is a subset of the worlds where Jane's class contains 16 boys and $m = \{x_c\}$ is the singleton set of the contextual concept of the boys in Jane's class. In this situation both representations in (43) are possible for (38) but they have different truth conditions.

(43) a. Every boy [in Jane's class]$_{\Delta}$ λ ϕP[the boy in Jane's class]$_{\Delta'}$ likes ϕP[the boy in Jane's class]$_{\Delta}$'s mother
b. Every boy λ ϕP[the boy]$_{\Delta'}$ likes ϕP[the boy]$_{\Delta}$'s mother.

In (43a), ellipsis of the adjunct *in Jane's class* is licensed at the NP-level because m only contains the concept x_c, which *boy [in Jane's class]* picks

out, and all other NPs either also select x_c or no concept in m at all.[30] Since the contextual concept of the boys in Jane's class is already a member of m, m remains unchanged for the evaluation of the complement of *every boy [in Jane's class]*$_\Delta$, but the salience of x_c is changed to medium. The evaluation procedure of *every* then applies the term starting with λ to each maximal concept of a single boy in Jane's class. Consider how the evaluation proceeds for one such concept x_b. The λ-operator adds x_b to m resulting in $m = \{x_c, x_b\}$. The definite description *the boy in Jane's class* selects x_b from this set because it has a bigger domain than x_c.[31]

In (43b), on the other hand, the evaluation proceeds much like in the case of the empty memory set discussed above. Specifically, quantification ranges over all concepts that are maximal for boy-hood and the definite description selects these.

Finally consider the representation in (44). The representation (44) is also possible for (38) and is equivalent to (43a). Specifically, when *the boy* is evaluated, x_b is selected since it is the boy-concept with the biggest domain in the memory set.

(44) Every boy [in Jane's class]$_\Delta$ λ ϕP[the boy in Jane's class]$_{\Delta'}$ likes [the boy]$_\Delta$'s mother.

3.2 Distributive Plurals

While in (38), the pronoun must be singular to be bound, it must be plural in (45), an example of distributive binding by a plural definite description. In fact, (45) is ambiguous between an interpretation where *their* is bound, and one where it is not.

(45) The boys like their parents.

A representation for the interpretation of (45) where *their* is not bound is (46). As for (38), I assume here that the initial memory set is the singleton set of x_c, the contextual concept of the sixteen boys in Jane's class.[32] Recall that I

[30] As mentioned above, *every* blocks deletion of its complement. If *every* is replaced with *each* in (42), deletion of the entire NP is possible.

[31] At this point, it is important that contextual concepts such as x_c are never maximal concepts. If the context set C was the set of all worlds where there are boys in Jane's class, x_c would also be maximal for $[\![\text{boy in Jane's class}]\!]^{m,s,C,w}$.

[32] In these distributive representations, I assume *sixteen* to be an NP-modifier. Therefore, it is expected that it can be late-adjoined and need not be represented in the trace position of the subject (Lebeaux 1988; Fox and Nissenbaum 1999).

assume that NPs are interpreted number-neutrally, which I indicate by the *: *boy** is true of a single boy and also of a plurality of boys.

(46) the_{dist} [sixteen]$_{\Delta}$ boy* [in Jane's class]$_{\Delta}$ λ ϕP[the boys in Jane's class]$_{\Delta'}$ like ϕP[the sixteen boy* in Jane's class]$_{\Delta}$ parents

For the bound interpretation, on the other hand, (47) is one possible representation.

(47) the_{dist} [sixteen]$_{\Delta}$ boy* [in Jane's class]$_{\Delta}$ λ ϕP[the boys in Jane's class]$_{\Delta'}$ like ϕP[the boy* in Jane's class]$_{\Delta}$ parents

In both cases, the distributive plural first adds the concept x_c defined as $[\![$the sixteen boy* in Jane's class$]\!]^{m,s,C,w}$ with domain C to the memory set at high salience. Then it selects a subphrase NP$'$ of *sixteen boy* in Jane's class*. I assume that this is *boy* in Jane's class*. It then constructs a set B of concepts that are maximal for $[\![$boy* in Jane's class$]\!]^{m,s,C,w}$ where the mereological sum in each $w \in C$ covers $[\![$the sixteeen boy* in Jane's class$]\!]^{m,s,C,w}(w)$. Having selected B it applies the complement to each of $x_b \in B$. Therefore the memory set contains two elements when the definite description corresponding to the pronoun is evaluated: x_c at high salience and one x_b at medium salience. The definite description *the sixteen boy* in Jane's class* selects the former, and agreement is plural. In (47), the definite description *the boy* in Jane's class* selects x_b because the domain of x_b is bigger than that of x_c, which is C. For the licensing of ϕ-features, however, salience is more important than the size of the domain. Therefore, the pronoun is plural, even if all $x_b \in B$ are singular. The difference between (38) and (47) is that, in (47), the salience of x_c is high, while in the evaluation of (38) it is medium.

Now consider how ellipsis of the complement DP of *their* is licensed in both cases. Here the difference between *s*-salience and d-salience is important for the explanation why two different ellipsis completions are available. In (46), the referent of *the sixteen boy* in Jane's class* is x_c which is the most *s*-salient element of the reference set. Ellipsis is therefore licensed at the DP-level because the given ellipsis completion is maximally *s*-salient. In (47), on the other hand, ellipsis is licensed because the ellipsis completion is maximally *d*-salient.

3.3 Relative clauses

The example in (48) illustrates how a relative clause works.[33] Assume again that initially the memory set is empty except for the self-concept and that the context set has some value not inconsistent with (48).

(48) Every boy who Mary met likes his mother.

A full logical form representation for (48) in this scenario, that captures the bound interpretation, is (49).

(49) every [boy λ Mary met ϕP[the boy]$_{\Delta'}$] λ [the boy λ Mary met ϕP[the boy]]$_{\Delta'}$ likes ϕP[the boy λ Mary met ϕP[the boy]]$_{\Delta}$'s mother.

Consider first the denotation of [λ Mary met the boy]: a predicate of concepts such that when they are added to *m*, it is true that Mary met the boy. In the evaluation of (49), the relative clause is applied to maximal boy-concepts by the special relative clause composition rule introduced as part of the formal system. The definite description *the boy* in the relative clause necessarily refers to these maximal concepts because they have the biggest possible domain. Therefore, the NP *boy Mary met the boy* overall express the property of being a boy that Mary met, as expected. The composition then proceeds as in the case without a relative clause in the previous section.

3.4 Grammatical Gender

Consider first example (50), which is repeated from (11) in the introduction. (50) shows that if there is a salient deictic referent, grammatical gender in German can be determined by a deleted noun that has no linguistic antecedent.

(50) At the car-dealership, out of the blue.

a. Der/Das da gefällt mir.
the.MASC/the.NEUT there please me
'I like this one over there.'

b. #Die da gefällt mir.
The.FEM there pleases me

[33] I consider here only matching relative clauses in the sense of Hulsey and Sauerland (2006).

A logical form representation of (50a) is (51a) or (51b) depending on the gender marking of the definite article. To capture the salience of the car I am pointing at, I assume that the initial memory set contains a concept x_c that selects the car I am pointing at for every world in the context set C. Then ellipsis is licensed in both cases because x_c has both properties *Auto* and *Wagen* and is maximally salient.

(51) a. I like the *Auto*.
b. I like the *Wagen*.

In (52), however, the gender of the pronoun must match that of the quantifier for the bound interpretation to be available. This is predicted by the present account because a maximal car-concept with the property *Auto* will not have the property *Wagen* and vice versa because the two nouns are not synonyms. Therefore the elided noun following the pronoun must match the noun of the quantifier.

(52) a. Jedes Auto hat Tage, an denen es/*er nicht
every.NEUT car[NEUT] has days on which it/*he not
funktioniert.
works
b. Jeder Wagen hat Tage, an denen er/*es nicht
every.MASC car[MASC] has days on which he/*it not
funktioniert.
works

The account furthermore predicts that a change of gender does become available in examples like (53). In this case, the elided noun following the pronoun *es* is *girl*. Feminine gender on *sie*, I assume, is interpreted as feminine followed possibly by an empty NP. Since being a girl entails female gender, *sie* selects the maximal girl-concept the universal quantifier introduced.

(53) Jedes Mädchen hat Tage, an denen es/sie nicht gut drauf
every.NEUT girl[NEUT] has days on which it/she not well on
ist.
is

'Every girl has days on which she isn't in a good mood.'

3.5 Multiple Binders

Next consider (54) with two overlapping binding relationships in the empty context.

(54) Every boy danced with every girl.

The representation that yields the expected interpretation for (54) is (55).

(55) every boy λ every girl λ ϕP[the boy]$_{\Delta'}$ danced with ϕP[the girl]$_{\Delta'}$

Both *every boy* and *every girl* lead to the addition of a maximal concept to the memory set—a maximal boy and a maximal girl concept. The definite description *the boy* and *the girl* then select the respective concepts.

In this case, it would probably be also sufficient to assume that the memory set contains individuals, rather than concepts. However, the next example shows that concepts are needed as mentioned with example (14).

(56) Every actress voted for every singer.

Imagine sentence (56) uttered in the following scenario: The actors' and singers' association is holding a vote for best performer among its members. Each member can vote for as many of the members as he chooses to. (57) shows the logical form of (56).

(57) Every actress λ every singer λ ϕP[the actress]$_{\Delta'}$ voted for ϕP[the singer]$_{\Delta'}$

If no members of the association are both a singer and an actress, (56) does not cause any obvious difficulty for the flat binding account. But, the example is interesting to consider in a situation where there are people like Jennifer Lopez who are successful both as a singer and as an actress. In such a situation, one interpretation of (57) requires, among other things, that Jennifer Lopez voted for herself to be true.

In this case, it would not be sufficient if the elements stored in the memory set were extensional individuals: Assume that JL is both a singer and an actress. The sentence *the actress voted for the singer* would be evaluated for each memory set {A,S} where A is an actress and S a singer.[34] However, this evaluation would not be possible unless there are no singers or actresses other than JL: Assume that NK is another actress. Then the expression in (58) would need to be evaluated, but because both NK and JL are actresses the uniqueness presupposition of the singular definite is violated.

[34] And specifically, the evaluation would take place with the singleton memory set {JL}.

(58) $[\![\text{the actress}]\!]^{\{NK,JL\}}$

Because of the Jennifer-Lopez-problem, a flat binding system with extensional individuals would evidently be a non-starter. But, extensional individuals have been given up for independent reasons already by Heim (1998) and others as was summarized above in section 2.1. Yet, moving to concepts alone does not solve the Jennifer-Lopez-problem. We need to furthermore ensure that a Jennifer-Lopez as singer concept does not have the property *actress*. For concepts that are defined for the context set this is not ensured: In a context, where the discourse participants have agreed that Jennifer Lopez is both a singer and an actress, the function that picks out Jennifer Lopez in all worlds of the context set satisfies both the property *actress* and the property *singer*. This would lead to the same problem with uniqueness that extensional individuals gave rise to. For this reason, I introduced the notion of a concept maximal for property *P*.

A concept maximal for *P* reflects that nothing is known about an individual other than that it has property *P*. This becomes apparent from the following characteristic of maximal concepts: If *x* is maximal for *P* and *x* also has property *Q*, then *P* entails *Q*.[35] For example, consider the properties *girl* and *younger than 20*. If *x* is a concept maximal for *girl*, it cannot have the property *younger than 20* because we can imagine worlds like the following *w*: Humans live as larvae underground for 20 years, until they hatch and become boys and girls. Since the maximal girl-concept *x* must select some girl in *w* and this individual is older than 20 years, *x* does not have the property *younger than 20*.

The Jennifer-Lopez-Problem does not arise when we make use of maximal concepts. The evaluation of sentence (56) now proceeds as follows. For any actress in the actual world there is a maximal actress concept x_a that selects this particular actress in the actual world. Similarly for any singer in the actual world there is a maximal singer concept x_s selecting this singer in the actual world. And when we evaluate the expression in (59) for such x_a and x_s no problem arises, because x_a uniquely has the property *singer* and x_s uniquely has the property *actress*.

(59) $[\![\text{the actress voted for the singer}]\!]^{\{x_a,x_s\}}$

[35]The characteristic formally follows from another feature we intuitively assume possible worlds and properties defined on them to have: For any two properties *P* and *Q*, if there is a possible world *w* and an individual *x* in *w* such that *x* has property *P*, but not property *Q*, then there is also a world *w'* where no individual that has property *P* also has property *Q*.

Note in particular that during the evaluation of (56) at least two different concepts that select Jennifer Lopez in the actual world play a role: one concept maximal for *singer* and one maximal for *actress*.

3.6 Indistinguishables

One further interesting case of overlapping dependencies are those involving two occurrences of the same NP such as the following:

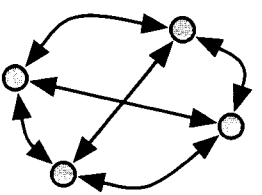
(60) Every dot is connected to every dot.

In these cases, the structure (61) is ruled out because the definite descriptions *the dot* would not uniquely refer even to a maximal concept.

(61) every dot λ every dot λ ϕP[the dot]$_{\Delta'}$ is connected to ϕP[the dot]$_{\Delta'}$

One natural interpretation of (60) is captured by the representation (62). This representation is true in the situation shown below.

(62) Every dot λ every dot [different from the dot]$_{\Delta}$ λ ϕP[the dot]$_{\Delta'}$ is connected to ϕP[the dot different from the dot]$_{\Delta'}$



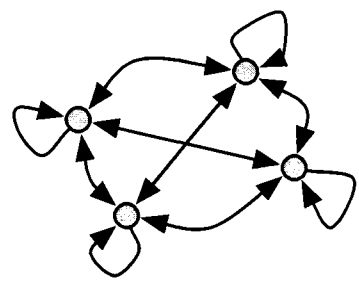
The interpretation procedure for (62) proceeds as follows. The higher quantifier *every dot* is treated in the same way as quantifiers in the previous examples. The lower quantifier ranges over maximal concepts for the property $[\![\text{dot different from the dot}]\!]^{m,s,C,w}$. Since *the dot* selects the concept introduced by the higher quantifier, this is then applied to w to yield an individual a. Recall here that I assume Kripkean individuals that exist in all possible worlds. So, the NP refers to the following property.[36]

$$\lambda b \in D_e \lambda w \in D_s \, \text{dot}(b)(w) \wedge a \neq b$$

[36]For concreteness I assume that *same* and *different* of natural language are directly captured by identity and non-identity on the model-theoretic level of individuals.

The maximal set of worlds where individuals satisfying this property exists is different from those where the maximal dot-concept is defined: While the former does not include worlds where some dot *a* is the only dot, while the latter does. Therefore, the two definite descriptions *the dot* and *the dot different from the dot* select different properties from *m*: *the dot* selects the maximal dot-concepts *every dot* introduced; *the dot different from the dot* selects the maximal concepts introduced by the second universal quantifier. The remaining question about representation (62) is why ellipsis of *different from the dot* is licensed – a question not usually considered. I will return to this question below, but first consider a different interpretation of (60).

While (62) captures one natural interpretation of (60), (60) also allows another interpretation: one that requires also connection of each dot to itself as in the following drawing:



Representation (62) is also true in this scenario, but the following two arguments show that (60) does have a stronger interpretation which requires the connections of dots to themselves. For one, a sentence like (63) seems to have a true reading.

(63) In the first drawing, not every dot is connected to every dot because no dot is connected to itself.

Furthermore, the fact that (64) is acceptable suggests that the self-exclusion in representation (62) is not the only interpretation of (60).

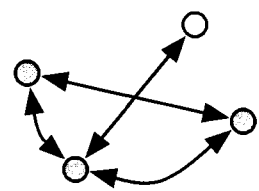
(64) Every dot is connected to every dot including/except itself.

There are many representations that would capture the interpretation that require self-connection within flat binding if ellipsis was not constrained: For example, representation (65) captures the relevant interpretation. Specifically, the concepts introduced by the two quantifiers have different domains because

dots are not necessarily grey. Therefore, the two definite descriptions correctly select the right concept. But since in all worlds in the context set all dots are grey, (65) does require for these worlds that each dot is connected to itself. And since there are many properties like *grey* that the dots in a concrete scenario all happen to have, but do not necessarily have, more representations equivalent to (65) are easy to come up with.

(65) Every dot λ every $[\text{grey}]_{\Delta}$ dot $\phi P[\text{the dot}]_{\Delta'}$ is connected to $\phi P[\text{the grey dot}]_{\Delta'}$

This shows that if silent domain restrictions are integrated in the account, a representation for (60) becomes available. However, the concrete representation (65) involving deletion of *grey* cannot be licensed in this scenario. This is shown by the fact that in a situation like the one depicted below, where some of the actual dots are not grey, (60) does not naturally exclude these. Rather (60) would be judged false in this situation. Therefore, (65) is not a possible representation for (60).



One representation for (60) that captures the self-connection requirement and where deletion is licensed is (66).

(66) Every dot λ every dot $[\text{connectable to by the dot}]_{\Delta}$ λ $\phi P[\text{the dot}]_{\Delta'}$ is connected to $\phi P[\text{the dot connectable to by the dot}]_{\Delta'}$

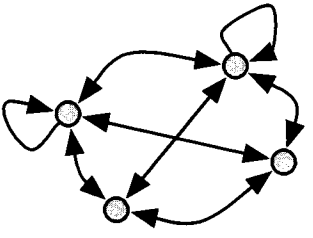
First consider the interpretation process for (66). The interpretation of the higher quantifier is as in the previous examples. For the lower quantifier, the property expressed by *dot connectable to by the dot* is computed. Assume that $[\![\text{the dot}]\!]^{m,s,C,w}(w)$ is some individual a. The dots connectable to by a are in some worlds a true subset of all dots, since for example dots on a different piece of paper are, in some worlds, not connectable to a. Especially, there are also worlds where there are no dots connectable to by a at all. Therefore, the concepts maximal for *dot connectable to by the dots* have a smaller domain than the maximal dot-concepts. As in the other examples, this is sufficient

for the two definite descriptions in (66) to correctly select different individual concepts from the memory set.[37]

An argument that quantificational statements in general are always implicitly restricted to instances where it is possible that the statement is true comes from the slight oddness of examples like (67).

(67) a. #Every dot is connected to every dot except for the ones on a different piece of paper.
b. #Everybody shook hands with everybody except for himself.

This now leaves the question of ellipsis licensing for both (62) and for (66). Actually though, representation (62) is no longer needed to explain a reading of (60): if in all worlds of the current context a dot is not connectable to itself, representation (66) is predicted to be true in exactly the situations where (62) is predicted to be true. Therefore, we restrict our attention to (66).[38] The reduction of (62) to (66) is supported by the fact that (60) is judged false in a scenario like the following because here the presence of some self-connection indicates that self-connection is possible.



[37]Philippe Schlenker (personal communication) points out that in mathematical examples like (i), the account requires that the properties of mathematical objects vary across possible worlds. However, it is well-known that mathematical examples raise problems for possible world accounts. My own feeling is that mathematical examples are conceptualized in ways not in accord with mathematical definitions when used in ordinary language.

(i) Every odd number is an even distance away from every odd number.

[38]Implicit restrictions of quantificational statements have been sporadically considered in the literature before, however, not within the context of ellipsis licensing theory (Westerståhl 1985; Stanley and Szabo 2000)—in fact, such a connection has been explicitly denied by Neale (2000). Somewhat related questions have been brought up in the literature on donkey anaphora, specifically by Heim (1990) and Elbourne (2006).

While I believe that there is evidence that the distinguishing content for (60) is actually present, the account of ellipsis licensing is still difficult.[39] Deletion must be licensed at the TP-level in the present system because otherwise it is impossible that two different NPs are reduced by deletion to the same PF as I assumed for (66) except when there is a mismatch between s- and d-salience. Furthermore, we need to determine the focus marking for (60). Since (60) could be preceded by any of the questions in (68), I conclude that (60) does actually not require a specific focus marking pattern for representation (66).

(68) a. What is connected?
b. What is up with dots?
c. What is every dot connected to?
d. What is connected to every dot?

In one case, deletion licensing is straightforwardly satisfied, namely, when there is focus on *every* and *dot* in the object position, which the question (68c) brings about. This representation is shown in (69):

(69) every dot λ $every_F$ dot_F [connectable to by the dot]$_\Delta$ λ ϕP[the dot]$_{\Delta'}$ is connected to ϕP[the <u>dot</u> connectable to by the dot]$_{\Delta'}$

I assume that focus marking interacts with structure sharing in the way discussed in the following section on sloppy readings. Therefore the focus alternatives of (69) are actually representations where the underlined occurrence of *dot* is replaced with the same lexical material as the F-marked occurrence of *dot*. If we replace the focused *every* with *the* and the focused *dot* and the underlined *dot* with *thing*, the focus alternative is a predicate of events paraphrased by (70).

(70) Every dot is connected to the thing connectable to by the dot

Since this predicate of events is true of the event introduced by the question (69c), which is plausibly more salient than any other event, the focus domain presupposition is satisfied by (69) in this case.

[39]Note that this a case of default domain restriction, a subclass of the general problem of domain restriction. The index-binding approach is not in a better position here. Consider for example the well-known problem of definites such as (i) (cf. Higginbotham 2006), where no tree has to be salient as long as there is a unique pair of house and tree such that the tree is in front of the house.

(i) The house with the tree in front is pretty.

It seems likely that a similar account carries over to all other contexts where (60) can occur, but I lack a general argument at this point.[40] I would like to point out though that in general licensing of DP ellipsis at the clausal level (TP on the present account) can be found in many other examples:

(71) a. Every sailor waved to every sailor. (Stanley and Williamson 1995)
b. If one bishop meets another bishop, he blesses him. (Hans Kamp p.c. Heim 1990)
c. John is talking to Bill. He likes him.

Example (71a) is considered by Stanley and Williamson (1995) in a special context; namely, one where we expect sailors on board a ship leaving the harbour to wave to the sailors remaining on land. In this context, the natural interpretation of (71a) is captured by (72a). (71b) and (71c) can be considered out of context, and (72b) and (72c) capture the natural interpretations of these examples.

(72) a. every sailor [on board] λ every sailor [on land] λ [the sailor on board] waved to [the sailor on land]
b. If a bishop meets a bishop [different from the bishop], [the bishop] blesses [the bishop different from the bishop]
c. John is talking to Bill. He [the John] likes him [the Bill]

Further evidence that structural parallelism is important for ellipsis resolution in these cases where there is no other disambiguating material comes from the following paradigms.

(73) a. *?If a bishop and a bishop meet, he plays chess against him. (Elbourne 2006)
b. ??If a bishop and a bishop meet, he and he play chess.

[40]Manfred Krifka (p.c.) suggests that ellipsis of expressions *former*, *latter* and the ordinal numbers could provide a different solution from the one developed in the text. While the suggestion is well worth exploring, note that a general solution would really require the ordinals. The ordinals, however, are in my judgment unacceptable unless there is a temporal or other order established among the referents. Hence, (i) is odd, while the example with pronouns instead in (ii) seems fine to me.

(i) #John made Bill meet Harry. The former knew the latter would like the third.

(ii) John made Bill meet Harry. He knew he would like him.

(74) a. *?John and Bill are talking. He likes him.
b. ?John and Bill are talking. He and he are getting together.

Finally, consider (75a) where structural parallelism is blocked. In this example, focus on both of the pronouns would not be licensed by the ellipsis completion in (75b). Therefore, the ellipsis completion in (75c) is accessible.

(75) a. John is talking to Bill. HE likes HIM.
b. *[HE [the John]]$_F$ likes [HIM [the Bill]]$_F$
c. [HE [the Bill]]$_F$ likes [HIM [the John]]$_F$

In sum, the section showed how 'indistinguishable' participants can be distinguished by a definite description. In particular, I showed that these do not raise any new problems. Rather they are closely related to the question how silent domain restrictions are licensed.

3.7 Strict and Sloppy

The ambiguity between strict and sloppy interpretations of VP-ellipsis is illustrated by the examples in (76).

(76) a. The boy called his mother and the teacher did too.
b. The child who dropped the spoon didn't pick it up, but the child who dropped the fork did.

Example (76a) is true if the teacher either called his own (the sloppy interpretation) or the boy's mother (the strict interpretation). Similarly (76b) is true if the fork-dropping child either picked up the spoon or the fork.

The ambiguity in examples with *only* like (77) has the same source.

(77) a. Only the boy called his mother.
b. Only the child who dropped the spoon found it again.

An index binding based analysis can provide an account of the ambiguity in examples like (77a) where the definite c-commands the pronoun with the same reference, but they generally do not extend to examples like (77b) where the definite does not c-command the pronoun. Tomioka (1999) first argued that the sloppy interpretation of examples like (77b) requires an E-type analysis of the pronoun *it* as a definite description. Tomioka proposes the content of the definite description amounts to the paraphrase *item he dropped*, where *he* is bound by the boy. Specifically, he proposes that it contain both a variable *R*

of type $\langle e, \langle e,t \rangle \rangle$ and a variable of type e that is bound by the c-commanding definite as shown in (78). How the value of R is restricted to only ever yield attested interpretations is left open by Tomioka's analysis.[41]

(78) The child who dropped the spoon λ_x didn't pick it[the $R(x)$] up, but the child who dropped the fork λ_y did [pick the $R(y)$ up]

One prediction of Tomioka's account is that the grammatical properties of the pronoun analyzed as $R(x)$ should be determined by R or the content of R rather than by the definite description in the preceding relative clause. However, this is not the case as (79) and (80) show. The German example (79) has the same ambiguity as the English example (77b), but German has grammatical gender marking. In (79), the grammatical gender of the pronoun *ihn* must correspond to that of the definite description *den Löffel*.

(79) Das Kind, das den Löffel fallen gelassen hat, hat
The child who the spoon.[MASC] drop let has hat
ihn nicht aufgehoben, aber das Kind, das die Gabel fallen
pro-MASC not picked up, but the child who the fork drop
gelassen hat, hat es getan.
let has has it done

The same point is made with grammatical number in the English example (80).

(80) The child who dropped the scissors didn't pick them up, but the child who dropped the fork did.

Therefore, I conclude that Tomioka's analysis, while correct about the link between E-type pronouns and sloppy interpretations, is not fully appropriate. Specifically, the restriction of the E-type pronoun cannot be a free property variable R, but must have specific lexical content.

The two interpretations of (76a) are captured within the present assumptions by the two different representations in (81a). For (76b), the two representations are given in (81b). These representations explain straightforwardly the match of formal properties, because the pronoun contains an occurrence of the same NP.

(81) a. (i) ... the teacher called the teacher's mother (*sloppy*)

[41] See also Safir (2005b) for discussion of Tomioka's analysis.

(ii) ... the teacher called the boy's mother (*strict*)

b. (i) ... the child who dropped the fork didn't pick up the fork (*sloppy*)

(ii) ... the child who dropped the fork didn't pick up the spoon (*strict*)

However, VP-deletion is not predicted to be licensed for the two representations of sloppy interpretations. This holds for both the case with c-command as well as the one without. Consider first (76a) with c-command since the solution will then carry over straightforwardly to the other case. My account of the sloppy interpretation of (76a) assumes a syntactic representation with structure sharing as in (82):[42]

(82) the $[\text{—}]_F$ called ϕP[the —$]_\Delta$'s mother and
boy
the $[\text{—}]_F$ called ϕP[the —$]_\Delta$'s mother
teacher

The correct interpretation is predicted for this representation. Furthermore ellipsis in the second conjunct is licensed because the TP that occurs in the first conjunct is a focus alternative of the TP of the second conjunct.

For the sloppy interpretation of (76b) the syntactic representation is (83).

(83) the child who ϕP[the child]$_{\Delta'}$
dropped the $[\text{—}]_F$ didn't pick up ϕP[the —$]_\Delta$
spoon
but the child who ϕP[the child]$_{\Delta'}$
dropped the $[\text{—}]_F$ did [pick up ϕP[the —]]$_\Delta$
fork

3.8 Propositional Attitudes

A basic example of a *de dicto* interpretation is (84), where we assume that John believes that there is a monster under his bed, while there really is none.

(84) John believes that the monster under his bed is French.

[42]The present analysis makes further novel predictions since it ties to structure sharing that I would want to explore in future work.

For the evaluation of (84), assume that $\mathrm{Dox}(j,w)$ is the set of John's belief worlds in w. Since it is known already that John believes that there is a monster under his bed, the initial memory set m contains the concept x that assigns to each world in $\mathrm{Dox}(j,w)$ the monster that is under John's bed. The interpretation rule for *believe* shifts the evaluation to $\mathrm{Dox}(j,w)$. It furthermore changes the memory set, but in this case it only adds the self-concept of $\mathrm{Dox}(j,w)$ to the memory set. The definite description *the monster under his bed* selects x_m. Since x_m is defined for $\mathrm{Dox}(j,w)$, the evaluation procedure then checks whether x_m actually selects a French individual for each world in $\mathrm{Dox}(j,w)$.

For a *de re* interpretation consider (85) relative to some C, $w \in C$, m and s. Here we assume that there is an actual American who was already introduced in the conversation, and therefore the memory set contains the concept x_a with domain C that selects this American for all worlds in C. Furthermore we know in C that John is looking at the American and thinks "That guy is French". Therefore, we assume that there also is concept x_f which selects in all of John's belief worlds the guy John is looking at.

(85) John believes that the American is French

Assume again that $\mathrm{Dox}(j,w)$ is the set of John's belief worlds in w. The interpretation rule for *believe* shifts the evaluation to $\mathrm{Dox}(j,w)$ and adds the self-concept of $\mathrm{Dox}(j,w)$ to the memory set. The definite description *the American* selects the concept x_a. Because this concept does not have a super-set of $\mathrm{Dox}(j,w)$ as its domain, the predicate expressed by *is French* is instead applied to x_f, a concept acquaintance-related to x_a because the definite *the person I am looking at* selects x_a in C and x_f in $\mathrm{Dox}(j,w)$.

A fuller version of the analysis would contain a proper dynamic system of adding individuals that were once referred to in the memory set of subsequent utterances. Then if (85) was followed by (86), (86) would be evaluated relative to a memory set containing a concept that is a French person in all of John's belief-worlds.

(86) The French person is now going to talk to John – John will be surprised by his American accent.

The determination of agreement explains the contrast in (87) which is adopted from Sauerland 2003. The initial memory set contains at least the following two concepts: x_s selecting Kai's mother for all worlds in C and x_p selecting the two monsters under his bed in all of Kai's belief worlds. While x_s

is singular, x_p is plural. The two concepts are acquaintance related because the definite description *the source of the noise* selects each one in their domain. In (87a), the singular agreement must be chosen because x_s is only mapped to x_p in the process of applying predication. In (87b), on the other hand, the definite description selects x_m directly.

(87) *Scenario:* Kai's mother hid underneath his bed and is making noise. Kai concludes that there must be two monsters under his bed making the noise.

a. Kai believes his mother/she is scary.
b. Kai believes the monsters/they are scary.

Next consider a *de re* interpretation with a bound variable pronoun in (88). A scenario bringing out the *de re* interpretation for (88) is the following: Every candidate saw a different ad on TV last night, which the candidate thought was impressive. He thought it was someone else's, but it actually was the one of his own campaign. So every candidate thinks the following thought: "I won't win. The guy I saw on TV last night will win."

(88) Every candidate believes that he will win.

The universal quantifier introduces a maximal candidate concept x_c. However, since both C and the belief set of every candidate are a subset of domain(x_c), the interpretation rule for *believe* changes this concept to x_c by narrowing the domain of x_c to not include Dox(a, w). Because of this change in the memory, the pronoun *he*, assumed to be the definite description *the candidate*, selects x'_c. This must be evaluated by the interpretation rule that applies, which works out to the concept of the person the candidate saw on TV resulting in the *de re*-interpretation. The change of the memory set by believes rules out the direct application of the x_c to worlds in Dox(a, w) in the evaluation of (88). If this was possible, the interpretation that would then be predicted could be paraphrased as *Every candidate believes that all the candidates will win.*, which is clearly not available for (88). Therefore, it must be ruled out.

Finally consider a *de se*-interpretation. Percus and Sauerland (2003) argue that examples like (89) possess a representation that must be interpreted *de se*.

(89) John believes that he is Brigitte Bardot.

The *de se*-interpretation is captured by the representation in (90), where the part corresponding to *he* is marked by an underbrace.[43]

(90) John believes that $\underbrace{\text{the self related by acquaintance to } \phi\text{P[John]}_{\Delta}}_{\text{he}}$ is Brigitte Bardot.

The operator *the self related by acquaintance to* remains unpronounced. *HeP[John]* refers to the concept of John defined for the context set. The self related by acquaintance to the concept is the self-function with domain Dox(j,w) introduced by *believe* into the memory set because the definite description *the person named John when I'm not Brigitte Bardot and Brigitte Bardot otherwise* selects these two concepts in the context set and Dox(j,w) respectively.

3.9 Personal Pronouns

I consider only the personal pronoun *I* of English here, but these remarks can be straightforwardly extended to *you* as the person *I* is addressing, and the third person and plural forms. For languages other than English, the relevant data are only now becoming available and it is not yet clear whether they are problematic for the general picture I am developing here. To my knowledge, the differences seem to only concern the pronunciation of *de se* pronouns.

Heim (1994) distinguishes three uses of English *I*: referential, bound, and *de se*. For my account, I distinguish furthermore two different bound uses: distributively bound and focus bound. The four uses are illustrated in (91), where the relevant occurrence of a first person form is marked by bold face.

(91) a. **I** am awake.
b. We each kissed **our** wife.
c. Only I did **my** homework; John didn't.
d. I dreamed **I** was Brigitte Bardot.

Third person pronouns also allow these four uses. Except for the referential use, it is not clear right away under what semantic conditions use of the first person form is licensed and therefore obligatory.

(92) a. John is awake.

[43]The part of (i) written out as *the self related by acquaintance to* could be abbreviated by a * in the spirit of Castañeda (1966).

b. Everyone of us kissed his wife.
c. Only John did his homework; I didn't.
d. John dreamed he was Brigitte Bardot.

Heim therefore proposes a syntactic account for the distribution of first person forms other than the referential case.[44]. In her system, the presence of first person is tied to a set of indexing conventions.

The present analysis predicts first person marking in all four cases in (91a) on purely semantic grounds. In the deictic case, the important assumption is that the initial memory set always not only contains the self-concept for the initial context set and that this must have the property *current speaker.* Therefore the representation (93) captures the referential interpretation.[45]

(93) I[the current speaker] is awake.

In the distributive binding case the analysis given for plural marking above carries over straightforwardly to the case of person marking. A logical form representation of the bound interpretation of (91b) is (94)—assuming (91b) is spoken by one of a group of explorers. When the definite description *the explorers* is evaluated, the memory set contains two relevant concepts: the contextual plural concept of the group of explorers a high salience and a maximal explorer concept at medium salience. The maximal concept will be selected for the reference of *the explorers*, but for the determination of the agreement properties of this ϕP the contextual concept is selected because of its higher salience.

(94) We[the$_{\text{DIST}}$ explorers] λ [the explorers] each kissed our [the explorers] wives.

The focus binding case requires an extension of structure sharing to the person features of the pronoun. The logical form representation for (91c) is shown in (95):

[44]See also (Kratzer 1998; Schlenker 1999; von Stechow 2003; Heim to appear), and Schlenker (2003); Safir (2005a) for partial alternatives.

[45]To the extent that (i) is grammatical, it requires third person agreement. I assume that ellipsis is obligatory in (93) therefore the subject in (i) cannot refer to the self-concept.

(i) The current speaker is awake.

(95) Only [[—] the —] did [[—] the —] homework.
current speaker
1st person

Representation (96) is a focus alternative of the scope of *only* in (95).

(96) [[—]$_F$ the —] did [[—] the —] homework.
John
3rd person

Finally the *de se*-case is accounted for by the pronunciation rule for *de se*-descriptions. One possible representation for (91d) is the following:

(97) I[the current speaker]$_\Delta$ dreamed that the self associated with I[the current speaker]$_\Delta$ is Brigitte Bardot.

4 Conclusion

In this paper, I have shown that it is possible to give a semantics for a rich fragment of natural language without use of storage sequences. All serious prior accounts of sentence interpretation include such storage sequences. Since the sequences would represent a language specific part of memory, a system without storage sequences would be clearly preferable. The account offered in this paper is clearly just a fragment. Important aspects of natural language that I did not consider in this fragment are indefinites, discourse level aspects of pronominal reference, reflexives and reciprocals, pied-piping as well as comparatives. These and probably other areas of sentence interpretation would need to be rethought to make present treatments in index based semantics compatible with the flat semantics I proposed. Furthermore, some of the analysis I offer above—for example the treatment of *de se*—can probably be improved, though the analysis above is the best I was able to do at this point.

My analysis as it stands is supported by a number of arguments in addition to not requiring sequences: most of the underlying assumptions have already been argued for, and it yields four interesting new predictions. Review first the claim that most of the assumptions the account builds on have been independently argued for. Specifically, the use of concepts as pronoun referents and the use of definite descriptions as pronoun and trace representations are both established in the respective specialized literature. The two main new features of my analysis are a novel view of ellipsis and the use of maximal

concepts with quantifiers, but both of these have been very limited departures from other work in the field. Consider first ellipsis. As is standard, I assume that there are syntactic and semantic conditions on ellipsis, and the semantic condition I proposed is similar to the proposals of (Rooth 1992a; Fox 1999a). Where I depart is the syntactic condition: I assume that deletion is always available except for a set of specific environments where it is blocked, rather than a specific need for syntactic licensing. The assumptions are very similar empirically, but, on my view, deletion of adjoined modifiers is licensed whenever it is semantically licensed.

The other new feature of my analysis is the assumption that a quantifier that is restricted by property P must range over concepts that are maximal for property P. These are concepts about which nothing is known except that they have property P. While this implementation is novel, it has been recognized that the referents introduced by quantifiers do have special properties. For example, Schlenker (2005b) proposes that there is a second storage sequence that stores only the items introduced by quantifiers. My use of maximal properties is a different way of capturing this distinction, which is based on the semantic intuition that we lack any knowledge other than P about the referent the quantifier *every P* introduced.

Now consider the four new predictions my analysis makes. First, my analysis straightforwardly accounts for Landman's generalization (Landman 2005). Landman argues that all dependent elements in natural language are elements of type *e*. On my view, all dependent elements are definite descriptions because dependent elements must select a member of the memory set.

My analysis also predicts the key aspects of Percus's generalization that predicates cannot be interpreted *de re* (Percus 2000). Percus's generalization is illustrated by (98). (98) allows an interpretation according to which some actual Italian is French in John's belief worlds. However, it does not allow an interpretation according to which some actual French person is Italian in John's belief worlds.

(98) John believes that the Italian is French

Percus captures the generalization by ruling out representation (99b), while allowing (99a).

(99) a. λ_w John believes $\lambda_{w'}$ that the Italian(w) is French(w')
b. *λ_w John believes $\lambda_{w'}$ that the Italian(w') is French(w)

On my view, *de re* interpretations only arise with definite descriptions. The system predicts that only items referring to individual concepts can be bound because only these can be retrieved from the memory set by a definite description. Worlds cannot be stored. Therefore, *de re* reference must be the consequence of having stored a concept with a different domain from the context set in the current evaluation state. This triggers the search for a concept related by acquaintance that is defined for the context set of the current evaluation state.

The third prediction of my approach concerns the lexical content of dependents. Several previous analyses have found empirical evidence for some form of lexical content of dependent elements—both for pronouns and for traces (Fox 1999b; Jacobson 2000; Sauerland 1998, 2004, in print; Schlenker 2005a). However, the lexical content of dependent elements plays no role in a sequence based account of binding because the content of the bound element is determined fully by a sequence position already. For example, I make use of the indexed definite determiner in (100) in my own earlier work, which presupposes that index *i* have a certain property. However, the reference is determined entirely by the index with no contribution from the lexical material in P. Such an analysis provides no semantic motivation for the presence of the lexical material in P.[46]

(100) $[\![\text{the}_i\ P]\!]^g = g(i)$ with a presupposition that P(i) holds

The present analysis, however, provides this semantic motivation: Every referring expression must contain so much lexical material as to uniquely select an item from the memory set.

The fourth and final interesting implication of this work concerns the set of quantifiers possible in human language. Quantificational expressions in human language have been argued to take a restrictor-argument in addition to the scope—they are so-called *generalized quantifiers*. However, it has been observed that actually only very few of the possible generalized quantifiers are expressible by lexical items in natural language (Hackl 2000; Krifka 2000; Sauerland 2000b). While there has been some progress made at explaining the so-called conservativity restriction (Fox 1999b), the scarcity of generalized quantifiers remains largely unexplained to my knowledge (cf. Geurts 2005). In the present work, quantifiers range over concepts that are maximal for

[46]For bound occurrences of overt definite descriptions, there is the pragmatic motivation that the lexical material can help disambiguating the index of the definite description (Schlenker 2005a).

some property P. Since the numerosity of such concepts is plausibly beyond human knowledge for any non-contradictory property P as is the numerosity of possible worlds, it would seem to me that only the Aristotelian quantifiers are straightforwardly expected on the present approach. This at least comes close to explaining the scarcity of quantificational expressions in human language.

These four implications provide additional motivation for the approach taken in this paper. However, the central motivation remains the goal of eliminating sequences as a language specific part of human memory. The elimination of sequences is clearly desirable within the biolinguistic approach (Chomsky 1995; Hauser et al. 2002). The elimination of sequences as part of the mechanism of sentence interpretation allows us to tie the semantic model of sentence interpretation more closely to other areas of human cognition that access the semantic model and memory. One remaining obstacle for such a tie-in is probably my use of possible worlds for the representation of knowlege. However, I use possible worlds only to characterize knowledge, so in principle the approach I argued for should be compatible with other approaches to knowledge representation.

References

Aloni, Maria
2001 *Quantification under Conceptual Covers.* Ph.D. thesis, University of Amsterdam, Amsterdam.
2005 A formal treatment of the pragmatics of questions and attitudes. *Linguistics and Philosophy*, 28(5): 505–539.

Anand, Pranav and Andrew Nevins
2004 Shifty operators in changing contexts. Unpublished Ms., MIT.

Büring, Daniel
2006 *Binding Theory.* Cambridge University Press, Cambridge, UK.

Castañeda, Hector-Neri
1966 'He': A study in the logic of self-consciousness. *Ratio*, 7: 130–157.

Chomsky, Noam
1981 *Lectures on Government and Binding: The Pisa Lectures.* Mouton de Gruyter, Berlin.
1995 *The Minimalist Program.* MIT Press, Cambridge, Mass.

Chomsky, Noam and Howard Lasnik
1993 Principles and parameters theory. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann, (eds.), *Handbook Syntax*, pp. 506–569. Walter de Gruyter, Berlin.

Corbett, Greville
2006 *Agreement.* Cambridge University Press, Cambridge, UK.

Curry, Haskell B.
1930 Grundlagen der kombinatorischen Logik. *American Journal of Mathematics*, 52: 509–536.

Curry, Haskell B. and Robert Feys
1958 *Combinatory Logic, Volume I.* North-Holland, Amsterdam.

Davis, Steven, (ed.)
1991 *Pragmatics. A Reader.* Oxford University Press, Oxford, UK.

Elbourne, Paul
2001 E-type anaphora as NP-deletion. *Natural Language Semantics*, 9: 241–288.
2006 *Situations and Individuals.* MIT Press, Cambridge, Mass.

Enger, Hans-Olav, Tore Nesset, and Curt Rice, (eds.)
2006 Special issue on the grammar of gender. *Lingua*, 116(9).

Evans, Gareth
1977 Pronouns, quantifiers, and relative clauses, I & II. *The Canadian Journal of Philosophy*, 7: 467–536, 777–797.

Fox, Danny
1999a Focus, parallelism, and accommodation. In Tanya Matthews and Devon Strolovitch, (eds.), *Proceedings of SALT 9*, pp. 70–90. Cornell University, CLC Publications, Ithaca, N.Y.

1999b Reconstruction, variable binding and the interpretation of chains. *Linguistic Inquiry*, 30: 157–196.

2000 *Economy and Semantic Interpretation.* MIT Press, Cambridge, Mass.

Fox, Danny and Jon Nissenbaum

1999 Extraposition and scope: A case for overt QR. In S. Bird et al., (eds.), *Proceedings of WCCFL 18*, pp. 132–144. Cascadilla Press, Somerville, Mass.

Frege, Gottlob

1879 *Begriffsschrift. Eine der arithmetischen nachgebildete Formelsprache des reinen Denkens.* Neubert, Halle, Germany.

Gärtner, Hans-Martin

1999 Phrase-linking meets minimalist syntax. In Sonya F. Bird, Andrew Carnie, Jason D. Haugen, and Peter Norquest, (eds.), *Proceedings of WCCFL 18*, pp. 159–169. Cascadilla Press, Somerville, Mass.

2002 *Generalized Transformations and Beyond: Reflections on Minimalist Syntax.* Akademie Verlag, Berlin, Germany.

Geach, P. T.

1972 A program for syntax. In Donald Davidson and Gilbert Harman, (eds.), *Semantics for natural language*, pp. 483–497. Reidel, Dordrecht, Netherlands.

Geurts, Bart

2005 Unary quantification revisited. University of Nijmegen, Ms.

Hackl, Martin

2000 *Comparative Quantifiers.* Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, Mass.

Hankamer, Jorge and Ivan Sag

1976 Deep and surface anaphora. *Linguistic Inquiry*, 7: 391–428.

Hauser, Marc D., Noam Chomsky, and Tecumseh Fitch

2002 The faculty of language: what is it, who has it, and how did it evolve? *Science*, 298: 1569–1579.

Heim, Irene

1982 *The semantics of definite and indefinite noun phrases.* Ph.D. thesis, University of Massachusetts, Amherst.

1990 E-type pronouns and donkey-anaphora. *Linguistics and Philosophy*, 13: 137–177.

1992 Presupposition projection and the semantics of attitude verbs. *Journal of Semantics*, 9: 183–221.

1994 Puzzling reflexive pronouns in *de se* reports. Unpublished Handout (Presented at Bielefeld), MIT.

1998 Anaphora and semantic interpretation: A reinterpretation of Reinhart's approach. In Uli Sauerland and Orin Percus, (eds.), *The Interpretive Tract, MITWPL 25*, pp. 205–246. MIT Working Papers in Linguistics, Cambridge, Mass.

to appear Features on bound pronouns. In David Adger et al., (eds.), *Φ-features*. Oxford University Press, Oxford, UK.

Heim, Irene and Angelika Kratzer

1998 *Semantics in Generative Grammar*. Blackwell, Oxford, UK.

Hepple, Mark

1990 *The grammar and processing of order and dependency: A categorial approach*. Ph.D. thesis, University of Edinburgh, Scotland.

1992 Command and domain constraints in categorial theory of binding. In *Proceedings of the Amsterdam Colloquium*, University of Amsterdam, Netherlands.

Heycock, Caroline

1995 Asymmetries in reconstruction. *Linguistic Inquiry*, 26: 547–570.

Higginbotham, James

1983 Logical form, binding, and nominals. *Linguistic Inquiry*, 14: 679–708.

2006 The simplest hard problem I know. In *Between 40 and 60 Puzzles for Krifka*. Zentrum für allgemeine Sprachwissenschaft, Online publication. *http://www.zas.gwz-berlin.de/publications/40-60-puzzles-for-krifka/*

Hulsey, Sarah and Uli Sauerland

2006 Sorting out relative clauses. *Natural Language Semantics*, 10: 111–137.

Jacobson, Pauline

1999 Towards a variable-free semantics. *Linguistics and Philosophy*, 22: 117–184.

2000 Paychecks, stress, and variable free semantics. In *Proceedings of SALT 10*, CLC Publications, Cornell University, Ithaca, N.Y.

Kaplan, David

1968 Quantifying in. *Synthese*, 19: 178–214.

Kratzer, Angelika

1998 Scope or pseudoscope? Are there wide-scope indefinites? In Susan Rothstein, (ed.), *Events in Grammar*, pp. 163–196. Kluwer, Dordrecht, Netherlands.

2006 Minimal pronouns. Unpublished Manuscript, University of Massachusetts, Amherst.

Krifka, Manfred

2000 At least some determiners aren't determiners. Unpublished ms., University of Texas, Austin.

Landman, Fred and Ieke Moerdijk

1983 Compositionality and the analysis of anaphora. *Linguistics and Philosophy*, 6: 89–114.

Landman, Meredith

2005 *Variables in Natural Language*. Ph.D. thesis, University of Massachachusetts, Amherst.

Lasersohn, Peter
2005 The temperature paradox as evidence for a presuppositional analysis of definite descriptions. *Linguistic Inquiry*, 36: 127–134.
Lebeaux, David
1988 *Language acquisition and the form of grammar*. Ph.D. thesis, University of Massachusetts, Amherst.
Lewis, David
1970 General semantics. *Synthese*, 22: 18–67.
Lobeck, Anne
1995 *Ellipsis*. Oxford University Press, Oxford.
Marti, Luisa
2003 *Contextual variables*. Ph.D. thesis, University of Connecticut, Storrs.
McBurney, Susan L.
2002 Pronominal reference in signed and spoken language: Are grammatical categories modality-dependent? In Richard P. Meier, Kearsy Cormier, and David Quinto-Pozos, (eds.), *Modality and structure in signed and spoken languages*. Cambridge University Press, Cambridge, UK.
Merchant, Jason
2001 *The Syntax of Silence: Sluicing, Islands, and the Theory of Ellipsis*. Oxford University Press, Oxford, UK.
2005 Fragments and ellipsis. *Linguistics and Philosophy*, 27: 661–738.
Miyake, Akira and Priti Shat, (eds.)
1999 *Models of Working Memory: Mechanisms of Active Maintenance and Executive Command*. Cambridge University Press, Cambridge, UK.
Montague, Richard
1974[1970]The proper treatment of quantification in ordinary English. In R. Thomason, (ed.), *Richard Montague: Selected Papers*, pp. 247–270. Yale University Press, New Haven, Conn.
Neale, Stephen
2000 On being explicit: Comments on Stanley and Szabo, and on Bach. *Mind and Language*, 15: 284–294.
Ogihara, Toshiyuki
1996 *Tense, attitudes, and scope*. Kluwer, Dordrecht.
Percus, Orin
2000 Constraints on some other variables in syntax. *Natural Language Semantics*, 8: 173–229.
Percus, Orin and Uli Sauerland
2003 On the LFs of attitude reports. In Matthias Weisgerber, (ed.), *Proceedings of Sinn und Bedeutung 2002*. Universität Konstanz, Konstanz, Germany.
Quine, Willard van Orman
1960 *Word and Object*. MIT Press, Cambridge, Mass.

Reinhart, Tanya
1983 Coreference and bound anaphora: A restatement of the anaphora questions. *Linguistics and Philosophy*, 6: 47–88.

Rooth, Mats
1992a Ellipsis redundancy and reduction redundancy. In Steve Berman and Arild Hestvik, (eds.), *Proceedings of the Stuttgart Ellipsis Workshop*. Arbeitspapiere des Sonderforschungsbereichs 340, Bericht Nr. 29, IBM Germany, Heidelberg.
1992b A theory of focus interpretation. *Natural Language Semantics*, 1(1): 75–116.

Ross, John Robert
1969 Guess who? In Robert I. Binnick, Alice Davison, Georgia M. Green, and Jerry L. Morgan, (eds.), *CLS 5*, pp. 252–278. Chicago Linguistics Society, Chicago.

Safir, Kenneth
2005a Person, context and perspective. *Rivista di Linguistica*, 16.
2005b *The Syntax of (In)dependence*. MIT Press, Cambridge, Mass.

Sauerland, Uli
1998 *The meaning of chains*. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, Mass.
2000a The content of pronouns: Evidence from focus. In Tanya Matthews and Brendan Jackson, (eds.), *The Proceedings of SALT 10*, pp. 167–184. Cornell University, CLC Publications, Ithaca, N.Y.
2000b No 'no': On the crosslinguistic absence of a determiner 'no'. In Hidekazu Suzuki, (ed.), *Proceedings of the Tsukuba Workshop on Determiners and Quantification*, pp. 415–444. Tsukuba University, Tsukuba, Japan.
2003 A new semantics for number. In *The Proceedings of SALT 13*, pp. 258–275. Cornell University, CLC-Publications, Ithaca, N.Y.
2004 The interpretation of traces. *Natural Language Semantics*, 12: 63–127.
2007 Focus dependency as structure sharing. In *Proceedings of SuB 12*. Barcelona, Spain. (to appear).
in print The silent content of bound variable pronouns. In Kyle Johnson, (ed.), *Topics in Ellipsis*. Oxford University Press, Oxford, UK.

Sauerland, Uli, Jan Andersen, and Kazuko Yatsushiro
2005 The plural is semantically unmarked. In Stephan Kepser and Marga Reis, (eds.), *Linguistic Evidence – Empirical, Theoretical, and Computational Perspectives*. Mouton d'Gruyter, Berlin, Germany.

Schlenker, Philippe
1999 *Propositional attitudes and indexicality*. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, Mass.
2003 A plea for monsters. *Linguistics and Philosophy*, 26: 29–120.

2005a Minimize restrictors! (notes on definite descriptions, condition C and epithets). In *Proceedings of Sinn und Bedeutung 9*. Nijmegen, Netherlands.

2005b Non-redundancy: Towards a semantic reinterpretation of binding theory. *Natural Language Semantics*, 13(1): 1–92.

Schönfinkel, Moses

1924 Über die Bausteine der mathematischen Logik. *Mathematische Annalen*, 92: 305–316.

Schwarzschild, Roger

1996 *Pluralities*. Kluwer, Dordrecht, Netherlands.

1999 GIVENness, AVOIDF and other constraints on the placement of accents. *Natural Language Semantics*, 7: 141–177.

Stalnaker, Robert

1978 Assertion. In Peter Cole, (ed.), *Pragmatics*, pp. 315–332. Academic Press, New York, N.Y. (reprinted in Davis 1991).

Stanley, Jason and Zoltan Szabo

2000 On quantifier domain restriction. *Mind & Language*, 15: 219–261.

Stanley, Jason and Timothy Williamson

1995 Quantifiers and context-dependence. *Analysis*, 55: 291–295.

Szabolcsi, Anna

1987 Bound variables in syntax (are there any?). In *Proceedings of the Amsterdam Colloquium*, pp. 331–351. University of Amsterdam, Netherlands.

Tanenhaus, Michael K., Michael J. Spivey-Knowlton, Kathleen M. Eberhard, and Julie C. Sedivy

1995 Intergration of visual and linguistic information in spoken language comprehension. *Science*, 268: 1632–1634.

Tarski, Alfred

1936 Der Wahrheitsbegriff in den formalisierten Sprachen. *Studia Philosophica*, 1: 261–405.

Thornton, Rosalind and Kenneth Wexler

1999 *Principle B, VP Ellipsis, and Interpretation in Child Grammar*. MIT Press, Cambridge, Mass.

Tomioka, Satoshi

1999 A sloppy identity puzzle. *Natural Language Semantics*, 7: 217–248.

von Stechow, Arnim

2003 Binding by verbs. In Makoto Kadowaki and Shigeto Kawahara, (eds.), *Proceedings of NELS 33*, GLSA, University of Massachusetts, Amherst.

Westerståhl, Dag

1985 Determiners and context sets. In Johan van Benthem and Alice ter Meulen, (eds.), *Generalized Quantifiers in Natural Language*, pp. 45–71. Foris, Dordrecht, The Netherlands.

Yatsushiro, Kazuko
2001 The distribution of *mo* and *ka* and its implications. In *Proceedings of the Third Formal Approaches to Japanese Linguistics*, pp. 181–198. MIT Working Papers in Linguistics, Cambridge, Mass.
Yatsushiro, Kazuko and Uli Sauerland
2006 [Feminine] in a high position. *Snippets*, 13.

The Grammar of Focus Interpretation*

Sigrid Beck

1 Introduction

This paper investigates the use of alternatives in natural language. Our starting point is the observation that many constructions in natural language have been argued to involve an alternative semantics. The most prominent example is focus. A standard analysis of interpretation (1b) of sentence (1a), for instance, goes something like this (Rooth (1985, 1992)): (1a) without *only* makes available two semantic objects, the proposition in (2a) and the set of alternative propositions in (2b). The adverb *only* says that out of all the alternatives, the single true one is the normal semantics, cf. (3).

(1) a. Rachel only invited ROSS.

b. Nobody other than Ross is such that Rachel invited them.

(2) a. that Rachel invited Ross

b. {that Rachel invited Ross, that Rachel invited Joey, that Rachel invited Monica}

(3) Out of all the propositions in (2b), the single true one is (2a).

Besides focus, alternatives have been argued to be used in the semantics of questions, negative polarity, disjunction and others.

I would like to raise the question how we decide whether a particular phenomenon involves an alternative semantics or not. This is not predictable from the interpretation we want to assign to the construction considered. Take questions, for example: one can reasonably claim that the classical Hamblin/Karttunen semantics for wh-questions is derived by means of an

*Acknowledgements: I would like to thank Daniel Hole, Uli Sauerland, Junko Shimoyama, Kazuko Yatsushiro and the audience at the 2005 Symposium "Interfaces + Recursion = Language".

alternative semantics in Hamblin (1973), but without recourse to alternatives in Karttunen (1977). I propose an empirical criterion to help us decide, namely minimality or intervention effects. Following Rooth (1992), I adopt a focus semantics from which the constraint in (4) follows.

(4) Evaluation of alternatives cannot skip an intervening focus sensitive operator.

The constraint shows up in intervention effects like (5) (from Pesetsky (2000)), where a *which*-phrase in situ cannot be interpreted when it occurs in the scope of a focus sensitive operator like *only*. This is an argument in favour of using an alternative semantics in the compositional interpretation of wh-questions (Beck (2006)).

(5) ?* Which boy did only Mary introduce which girl to __?

I use the minimality constraint on alternative evaluation to argue that negative polarity, alternative questions, and *either... or*-constructions, in particular, do indeed use an alternative semantics. A further outcome will be that there is crosslinguistic variation with respect to where an alternative semantics is employed. The use of alternatives is a choice offered in the interpretation component.

Section 2 introduces linguistic phenomena which have been analysed on the basis of an alternative semantics. This will give us a set of expressions we can assume introduce alternatives into the semantics. In section 3 I address the issue of how these alternatives are used, i.e. where and how they are evaluated. We are then equipped with a grammar that derives the minimality condition mentioned in (4). Section 4 argues that neither formal nor interpretive criteria tell us conclusively what constructions are compositionally interpreted via an alternative semantics. The minimality constraint on alternative evaluation is offered as an empirical criterion instead. Its consequences in terms of intervention effects are observed and compared to other constraints that might show up if a different compositional interpretation without alternatives were at work. Conclusions are drawn in section 5.

2 Alternatives in Natural Language

In this section I collect a set of expressions that can plausibly be assumed to trigger introduction of alternatives into the semantic calculation. The collection is not intended to be exhaustive.

2.1 Focus (Rooth 1985, 1992)

We have already said that focus has an alternative semantics (Rooth (1985, 1992, 1996)).[1]

Focus on *Ross* in the example below has the effect that besides the ordinary semantics of that DP (reference to the individual Ross), the DP introduces a set of alternative individuals. I call the focused DP the alternative trigger or Alt-trigger. The Alt-trigger is the expression responsible for introducing alternatives into the calculation. The alternatives introduced by $Ross_F$ become relevant when we want to evaluate the effect of focus semantically or pragmatically – in the example presumably at the level of the sentence as a whole. At this point, we have alternative propositions (calculated compositionally for example in the manner developed in Rooth (1985)). I use XP-Alt to refer to the phrase at the level of which alternatives become semantically operative.

(6) Rachel invited ROSS.

(6') a. Alt-Trigger: $\text{Ross}_F \rightarrow \{\text{Ross, Joey, } \ldots\}$

b. XP-Alt: [IP Rachel invited ROSS] $\rightarrow$
{that Rachel invited Ross, that Rachel invited Joey, ... }

2.2 Questions (Hamblin 1973, Stechow 1991)

One way in which the focus alternatives of an utterance can become relevant is when we want to decide whether the utterance is an appropriate answer to a question. Sentence (6) with the accent indicated would be an appropriate answer to the question in (7). The semantics of Hamblin (1973) assigns to (7) the meaning in (7'b) – a set of alternative possible answers to the question. The wh-phrase is responsible for introducing the alternatives, (7'a).

(7) Who did Rachel invite? [wh-question]

(7') a. Alt-Trigger: who $\rightarrow$ {Ross, Joey, ... }

b. XP-Alt: [CP who did Rachel invite] $\rightarrow$
{that Rachel invited Ross, that Rachel invited Joey, ... }

[1] I would like the term 'alternative semantics' to be understood loosely enough to also include an implementation in terms of structured meanings (Krifka (1992)). Genuinely excluded would be a focus semantics on the basis ofpresupposition (Geurts & van der Sandt (2005)); see that issue of Theoretical Linguistics and Sauerland (2005) for discussion.

The same idea – that the semantics of a question is the set of possible answers to the question – would associate (8) with the interpretation in (8'b). (8) is an alternative question, AltQ for short. Stechow (1991) proposes that the disjunctive phrase is the alternative trigger and acts in a way parallel to the wh-phrase in (7).

(8) Did Rachel invite Ross or Joey? [alternative question]

(8') a. Alt-Trigger: Ross or Joey → {Ross, Joey}

b. XP-Alt: [CP Did Rachel invite Ross or Joey] →
{that Rachel invited Ross, that Rachel invited Joey}

Rooth (1992) observes that there is a simple way of formulating question-answer congruence on the basis of this semantics for questions: the meaning of the question has to be the same as the alternative semantics of the answer.

2.3 Japanese *-mo* and *-ka* Constructions (Shimoyama 2001)

Our next example is a little different in that alternatives are used as the input to another operator and affect the truth conditions of the whole. The example in (9) is an instance of a Japanese *mo*-construction taken from Shimoyama (2001). There is a DP 'which student's mother' containing what looks like a wh-phrase (an 'indeterminate phrase' in Shimoyama's terms), attached to which is the morpheme *-mo*. The whole DP semantically contributes 'every student's mother' (Shimoyama 2001:12).

(9) [Dono gakusei-no okaasan]-mo odotta.
which student-GEN mother-MO danced
'Every student's mother danced.'

It seems that the morpheme *-mo* contributes universal quantification over the indeterminate phrase. (10) is another example; here we have a DP corresponding to 'the teacher that which student invited', to which *-mo* attaches.

(10) [[Dono gakusei-ga _ syootaisita] sensei]-mo odotta.
which student-NOM invited teacher-MO danced
'For every student x: a/the teacher(s) that x had invited danced.'

Shimoyama argues that quantification is not actually over the indeterminate phrase, but over alternatives, as indicated in (9'). The indeterminate phrase

introduces alternatives, which become relevant as we encounter the operator *-mo*. *-Mo* expresses universal quantification over the alternatives provided by its sister, cf. (11). The example is interpreted in (12), resulting in the desired semantics.

(9') a. Alt-Trigger: which student → {Rachel, Ross, Joey}

b. XP-Alt: [DP which student's mother] →
{Rachel's mother, Ross's mother, Joey's mother}

(11) ⟦ XP -MO ⟧ = λP. for all x in XP-Alt: P(x)=1

(12) a. ⟦[which student's mother] -MO⟧ =
λP. for all x in {Rachel's mother, Ross's mother, Joey's mother}:
P(x)=1

b. ⟦ [which student's mother] -MO danced ⟧ =
for all x in {Rachel's mother, Ross's mother, Joey's mother}:
x danced

Shimoyama groups *-mo* with another morpheme *-ka*, which also quantifies over alternatives but quantification is existential. See also Hagstrom (1998) on *-ka*. An example taken from Yatsushiro (2005) is given in (13) (the '%' indicates that not every speaker accepts this use of *-ka* as existential quantification at a distance).

(13) a. [[dare-no hahaoya]-ka-ga tukue-no ue-ni iru.
who-GEN mother-KA-NOM desk-GEN above-LOC is
'Someone (or other)'s mother is on the desk.'

b. % [[dare-o hihansita] hito]-ka-ga John-o hometa.
[who-ACC criticized] person-KA-NOM John-ACC praised
'Someone or other who had criticized someone praised John.'

2.4 NPIs (Krifka 1995, Lahiri 1998)

It has been argued (in particular by Krifka (1995), Lahiri (1998)) that alternatives play a role in NPI licensing. According to these theories, example (14) means something like (14').

(14) I didn't give him a red cent.

(14') I didn't give him the minimal amount of money (a red cent), & I didn't give him an alternative amount of money, & the most likely amount of money for me to give him is the minimal amount.

According to these theories, the NPI itself has as its ordinary meaning a minimum, in the example the minimum amount of money. In addition, it triggers the introduction of alternatives (15a). Those alternatives become relevant when we encounter the licenser of the NPI. I simplify things somewhat in (16), taking the licenser to be the expression 'not even' with the meaning indicated. (16) derives the interpretation in (14') as desired.

(15) a. Alt-Trigger: a red cent →
{the minimal amount of money, ..., the minimum +n, ...,
a lot of money, ...}

b. XP-Alt: [IP I give him a red cent] →
{that I gave him the minimal amount of money,...,
that I gave him the minimum +n, ...,
that I gave him a lot of money, ...}

(16) ⟦ not even XP ⟧ =1 iff ⟦XP⟧ =0 &
for all p in XP-Alt: p=0 &
⟦XP⟧ is more likely than the other p in XP-Alt

I should add that the purpose of this semantics for NPIs is to derive predictions about licensing environments for NPIs from the semantics. I will not go into the details here, see the works cited.

2.5 Disjunctions (Zimmermann 2000, Beck & Kim to appear)

A final example of an alternative trigger I want to consider are disjunctions, which already came up in the context of AltQs. Zimmermann (2000) proposes that (17) means (17').

(17) Either John ate rice or beans.

(17') It is possible that John ate rice and it is possible that he ate beans and nothing else is possible.

Beck & Kim (to appear) propose to derive (17') through the use of alternatives. As indicated in (18), the disjunction introduces alternatives (just

like it did in the case of AltQs). Those alternatives become relevant as we encounter *either*, which is analysed in (19) as an alternative sensitive operator. The reader can see that (19) will use (18) to derive the desired (17').

(18) a. Alt-Trigger: rice or beans → {rice, beans}

b. XP-Alt: [IP John ate rice or beans] →
{that John ate rice, that John ate beans}

(19) ⟦ either XP ⟧ = 1 iff [for all q in XP-Alt: may q] &
¬∃p [for all q in XP-Alt: p ∩ q={} & may p]

2.6 Perspective

From this collection of phenomena the following general perspective emerges. Certain expressions in natural language have the purpose of introducing alternatives:

(20) *Alt-Triggers:*
focused phrases
indeterminate phrases (e.g. in wh-questions, many NPI-constructions, *-mo/-ka* constructions)
disjunctions (e.g. in AltQs, *either... or* constructions)

Furthermore, the semantics of certain operators uses the alternative semantics of their arguments:

(21) *Alt-Evaluation:*
focus evaluation (i.e. semantic evaluation of the contribution of focus)
question formation
-mo and *-ka*, *either*

In (20) and (21) I have only listed the phenomena we have considered. It is likely that there are further instances – plausible candidates being free choice contexts with both disjunctions (Aloni (2003), Simons (2004)) and indefinites (Kratzer & Shimoyama (2002)), and perhaps further operators similar to *-mo* and *-ka*. I have limited myself to the discussion of the cases that I feel reasonably confident about.

3 Evaluating Alternatives

The purpose of this section is to discuss evaluation of alternatives in more detail. This is necessary in order to provide the empirical criterion for the use of an alternative semantics promised in the introduction. In order to proceed I need the notation introduced in (22) for the ordinary and the alternative semantic value of an expression. What I called XP-Alt in the examples above is the alternative semantic value of the phrase at the level of which alternatives are evaluated.

(22) a. $[\![\alpha]\!]^o$: the ordinary semantic value of α;
$[\![XP]\!]$ above is now $[\![XP]\!]^o$

b. $[\![\alpha]\!]^{Alt}$: the alternative semantic value of α;
XP-Alt above is $[\![XP]\!]^{Alt}$

3.1 Focus Evaluation: Rooth (1992)

The most important case of alternative evaluation for present concerns is focus evaluation. I adopt Rooth's (1992) theory, according to which example (23a) has the structure in (23b) when the focus on *Ross* is evaluated at sentence level. An operator $\sim$ is adjoined to the sentence, which is responsible for focus evaluation. The $\sim$ operator is accompanied by a variable C, the focus anaphor. Rooth's semantics for the $\sim$ operator is given in (24).

(23) a. Rachel invited ROSS.

b. [$\sim$ C [Rachel invited Ross_F]]

(24) a. $[\![[\sim C\ XP]]\!]^o$ is only defined if $[\![C]\!]^o \subseteq [\![XP]\!]^{Alt}$.
If defined, $[\![[\sim C\ XP]]\!]^o = [\![XP]\!]^o$

b. $[\![[\sim C\ XP]]\!]^{Alt} = \{[\![[\sim C\ XP]]\!]^o\}$ [Rooth (1992)]

c. C is the focus anaphor; context provides a value for C

The impact of (24) is best explained on the basis of an example. In the case of (23b), it amounts to (25). One effect is a requirement on the context, which has to provide an appropriate value for the focus anaphor C (focus alternatives to the expression the $\sim$ attaches to). The $\sim$ does not affect the ordinary semantics of the structure it is contained in. The second effect of the $\sim$ is that focus is taken into account, and the alternatives introduced by a

focus in the scope of the $\sim$ are no longer preserved in the focus alternatives of larger structures.

(25) (23b) is only appropriate in a context that provides a value for C such that $[\![C]\!]^o \subseteq$ {that Rachel invited Ross, that Rachel invited Joey, ...}. In such a context, (23b) denotes the proposition
that Rachel invited Ross.
The alternative semantic value of (23b) is reset to the singleton
{that Rachel invited Ross}.

The first aspect of the semantics of the $\sim$ operator accounts for the contrast (26) vs. (26') – question-answer congruence. (26b)=(23) is predicted to be an appropriate answer to (26a) because the focus anaphor finds an antecedent in the context that corresponds to focus alternatives of the answer – namely the question in (26a). The same structure is not an appropriate answer to the question in (26'a), and this is predicted because (26'a) is not a set of focus alternatives to (26'b), hence not an appropriate antecedent for the focus anaphor in (26'b)=(23).

(26) a. Who did Rachel invite?

b. Rachel invited ROSS.

(26') a. Who invited Ross?

b. # Rachel invited ROSS.

The other semantic effect of the $\sim$ can be illustrated with embedding. The contrast (27) vs. (27') intuitively mirrors the one in (26) vs. (26').

(27) a. Who did Rachel invite?

b. I think that Rachel invited ROSS.

(27') a. Who invited Ross?

b. # I think that Rachel invited ROSS.

The analysis in the theory of Rooth (1992) goes as follows: The focus evaluating operator is attached to the embedded clause, (28a). The focus anaphor thus looks for an antecedent that corresponds to focus alternatives of the embedded structure. The question in (27a) is such an antecedent, but

the question in (27'a) is not. Note that the alternatives that are introduced by focus on *Ross* will not be passed on further, beyond the embedded clause. This means that we will not end up looking for the context to provide focus alternatives to the structure as a whole – something amounting to the question in (28b). And this is right, because there is no discourse antecedent like (28b) in the perfectly appropriate (27). This is the motivation for 'forgetting' the alternatives introduced by the focus on *Ross* for the purposes of further alternative calculation. (The focus on *Ross* is not forgotten for the purposes of discourse coherence, note, because it has affected the value of the focus anaphor attached to the embedded clause – cf. the way we account for the contrast (27) vs. (27')).

(28) a. I think that [∼C [Rachel invited ROSS]]

b. Who do you think that Rachel invited?

According to Rooth (1992), the contribution of focus (that is, intonational focus in English and presumably in related languages) is uniformly evaluated by the ∼ operator. Thus whenever we detect a semantic or pragmatic effect of focus, it must come about through the work of a ∼ operator. A case that will concern us later is that of focus sensitive particles like *only*. An example is given in (29a). The sensitivity to focus that *only* descriptively exhibits is mediated by the ∼ in the proposed structure for the example (29b). In this structure, the resource domain variable of *only* and the focus anaphor are identical. Thus *only* will in effect quantify over focus alternatives of the sister of the ∼ operator. The semantics of *only* is given in (30), and we will derive the desired interpretation for (29) in (31).

(29) a. Only ROSS left.

b. [only$_C$ [∼C [Ross$_F$ left]]]

(30) $[\![\text{only}_C\ \text{XP}]\!]^o = 1$ iff out of all propositions in $[\![\text{C}]\!]^o$, the single true one is $[\![\text{XP}]\!]^o$.
$[\![\text{C}]\!]^o = [\![\text{XP}]\!]^{Alt}$.

(31) out of all propositions in $[\![\text{XP}]\!]^{Alt}$ = {that Ross left, that Joey left, ...}, the single true one is the ordinary meaning of $[\![\text{XP}]\!]^o$ = that Ross left.
= The only one who left is Ross.

3.2 Question Formation: Hamblin (1973), Stechow (1991), Beck (2006)

The next case of alternative evaluation we will consider is question formation. I assume (like Hamblin (1973), Stechow (1991), Beck (2006)) that a question operator Q is responsible for using alternatives introduced by a wh-phrase. The structure of a simple wh-question like (32a) is (32b). As proposed in section 2, the wh-phrase is the Alt-trigger, introducing a set of individuals (33a) (according to Beck (2006), the wh-phrase has no ordinary semantic value). At the point where we encounter the Q operator, we have the alternative propositions in (33b).

(32) a. Who left?

b. [Q [$_\phi$ who left]]

(33) a. $[\![\text{who}]\!]^{Alt} = D = \{\text{Ross, Joey, } \ldots\}$

b. $[\![\phi]\!]^{Alt} = \{\text{that x left} \mid x \in D\} = \{\text{that Ross left, that Joey left, } \ldots\}$

This is, of course, already the desired Hamblin semantics for the question. All that remains to be done is to elevate this semantic object from the level of alternative semantics to the level of the ordinary semantics. This is the task of the question operator, defined in (34). Application to the example (35) yields the desired result.[2]

(34) $[\![Q\ \phi]\!]^o = [\![\phi]\!]^{Alt}$

(35) $[\![\ [Q\ [_\phi \text{ who left}]]\ ]\!]^o = [\![[_\phi \text{ who left}]]\!]^{Alt} = \{\text{that x left} \mid x \in D\}$

The same operator can derive AltQ meanings (Stechow (1991), Beck & Kim (to appear)). The structure of example (36a) is (36b). The Alt-trigger disjunction will provide us with the relevant alternatives (37a), and we derive the desired interpretation (37b).

(36) a. Did Ross or Rachel leave?

b. [Q [$_\phi$ Ross or Rachel left]]

[2] This is a simplified version. See Beck (2006) for the actual proposal. The version in the text is simplified in particular with regard to selectivity of the Q operator. (34) above makes Q evaluate all alternatives triggered in its scope. This is not accurate, as Beck (2006) discusses. Baker ambiguities and focus inside a question show that Q (in contrast to the $\sim$) is a selective operator.

(37) a. $[\![\phi]\!]^{Alt}$ = {that Ross left, that Rachel left}

b. $[\![$ Q [$_\phi$ Ross or Rachel left]$]\!]^o$ = $[\![$ [$_\phi$ Ross or Rachel left] $]\!]^{Alt}$ = {that Ross left, that Rachel left}

3.3 Others: Japanese *-mo*, English *either*

We have already seen a few more operators that use alternatives. I will work with the assumption that they have direct access to those alternatives and use the interpretations in (38), (39).

(38) a. $[\![$XP -mo$]\!]^o$ = λP. for all x in $[\![$XP$]\!]^{Alt}$: P(x)=1

b. $[\![$XP -mo$]\!]^{Alt}$ = { λP. for all x in $[\![$XP$]\!]^{Alt}$: P(x)=1 }

(39) $[\![$ either XP $]\!]^o$ = 1 iff for all q in $[\![$XP$]\!]^{Alt}$:
may q & $\neg\exists$p [for all q in $[\![$XP$]\!]^{Alt}$: p $\cap$ q = { } & may p]

Note that *-mo* as an alternative evaluating operator is similar to the $\sim$ in that it uses all alternatives in its scope and resets the alternative semantic value of the whole to the singleton set containing the ordinary semantics. Note also that I have not specified the effect of *either* on alternative semantic values. See below for discussion.

4 Where Is an Alternative Semantics Used?

We are now ready to address the central question posed in this paper: how do we know when an alternative semantics is used? I first observe that certain potential indicators do not provide conclusive evidence, and then I suggest an empirical criterion based on the semantics of focus evaluation we have just sketched.

4.1 Inconclusive Indicators: Meaning and Morphology

I would like to point out that the meaning of a construction does not tell us whether it will be derived from an alternative semantics. We saw that Japanese *-mo* amounts to universal quantification using the alternative semantic value of its complement. Other universal quantifiers do not plausibly operate on alternative semantic values, for instance English *every*. It would be perfectly possible to give a semantics to (40) in the manner indicated in (40b,c). But this is not a plausible analysis, for instance because there is no independent

indication that the argument of *every*, *them* in the example, makes available the necessary alternatives. A plausible analysis of *every* has it quantify over a range provided by the ordinary semantic value of its sister (although this may then be further constrained by alternatives evaluated by the $\sim$ – see below). Hence some but not all universal quantifiers are alternative evaluating operators.

(40) a. every one of them

b. $[\![\text{them}]\!]^{Alt} = \{\text{Rachel, Ross, Joey}\}$

c. $[\![\text{every one of XP}]\!]^{o} = \lambda P.\ \text{all x in } [\![\text{XP}]\!]^{Alt}: P(x)=1$

Phillippe Schlenker (p.c.) points out to me that there are several expressions in French that can express disjunction, as seen in (41). But not all of them are suitable for the formation of an alternative question (42). This suggests (although a more thorough investigation would be needed) that the various 'or's in French do not give rise to the same alternative semantic values. That would mean that some but not all disjunctions are Alt-triggers.

(41) a. Hans a bu du thé ou (bien) du café.
Hans has drunk of tea or (well) of coffee

b. Hans a bu soit du thé, soit du café.
Hans has drunk be.it of tea be.it of coffee
'Hans drank tea or coffee.'

(42) a. Est-ce que Hans a bu du thé ou (bien) du café? (ok AltQ)
is.it that Hans has drunk of tea or (well) of coffee

b. Est-ce que Hans a bu soit du thé, soit du café. (* AltQ)
is-it that Hans has drunk be.it of tea be.it of coffee
'Did Hans drink tea or coffee?'

Thus I think that it is unclear from an item's basic truth conditional contribution whether this semantics arises through the use of alternatives.

Kratzer & Shimoyama (2002) following Haspelmath (1997) observe that there is an interesting morphological connection between natural language expressions that are likely to function as Alt-triggers. Indeterminate phrases crosslinguistically tend to occur in several environments: as wh-phrases, NPIs and certain kinds of indefinites, possibly with some extra morphology involved. Some examples are given in (43).

(43) a. Malayalam: *aar-um* (who-also) = anyone (NPI)

b. Mandarin: *shei ye* (who also) = anyone (NPI)

c. Japanese: *dare-mo* (who-also/even) = anyone (NPI)

This is good confirmation of the general perspective pursued here, that certain expressions – namely indeterminate phrases – have the purpose of introducing alternatives. The fact that the ties show up crosslinguistically certainly support an analysis that identifies a common semantic core. But I think that as a criterion for the use of alternatives in semantics, morphology must be incomplete. For one thing, there will be historical accidents for individual morphemes (like an item changing from being an NPI to not being one or vice versa). For another, morphology will tell us that if we find morphological property X, we should look for an alternative semantics; but we may well have an alternative semantics without particular formal properties.

4.2 Minimality Condition on Deriving the Interpretation

I suggest instead that the circumstances under which the relevant interpretation arises tell us whether an alternative semantics is at work. The system of compositional interpretation sketched in section 3 makes the prediction in (44) (compare Beck (2006), Beck & Kim (to appear)):

(44) General Minimality Effect MIN:
The evaluation of alternatives introduced by an XP
cannot skip an intervening $\sim$ operator.
$*[\ Op\ [\ \sim C\ [_{\phi} \ldots\ XP \ldots\ ]]]$

This is predicted because the $\sim$ operator according to the semantics in (24) will use alternatives introduced in its scope, hence will evaluate the alternatives introduced by XP in the schema above. Furthermore, the alternative semantic value of the structure containing the $\sim$ operator will be reset to the ordinary semantics of the sister of the $\sim$ (see (45) below). Hence the alternatives introduced by XP will no longer be available for alternative set calculation above that $\sim$. As a result, Op does not have access to alternatives triggered by XP. According to this theory of focus evaluation, we expect a constraint on the evaluation of alternatives that amounts to the restriction that there cannot be an intervening $\sim$ between an Alt-Trigger and its evaluating operator.

(45) a. [∼C φ] uses $[\![\phi]\!]^{Alt}$, i.e. evaluates the alternatives triggered by XP

b. $[\![\ [\sim C\ \phi]]\!]^{Alt} = \{[\![\phi]\!]^o\}$, i.e. the alternatives triggered by XP are no longer passed on.

Now, the ∼ itself is not visible. Thus MIN only becomes an empirical constraint when we know where to expect a ∼. A visible intervener will be an element that needs a ∼, i.e. an element that is sensitive to focus:

(46) Problematic interveners are elements that trigger the introduction of a ∼ in their scope, i.e. elements that can give rise to a focus affected reading. In English and German, focusing and quantificational elements are problematic interveners (*only*, *even*, *not*, *nobody*, *every* N, *rarely*, ...). Focus sensitive particles should be problematic interveners cross-linguistically.

To this I add the generalization derived in Shimoyama (2001) and Kratzer & Shimoyama (2002), and by the semantics in (38) for *-mo*, that the evaluation of an indeterminate pronoun cannot skip an intervening evaluating operator. Thus *-mo* and the ∼ are problematic interveners for alternative evaluation (cf. Shimoyama 2001, Kratzer & Shimoyama 2002).

(47) * [... [[... indeterminate pronoun ...] *-mo/ka*] ...]*-mo/ka*

We are now ready to observe the effects of MIN in the constructions introduced in section 2 as plausibly involving an alternative semantics. In all of them, there are effects that can be analyzed as consequences of MIN, supporting the analysis in terms of alternatives.[3]

4.2.1 MIN in focus evaluation

Beck (2006) (see also Beck & Vasishth (in prep.)) observes that (48B) below cannot have the interpretation indicated. This means that *only* cannot associate with the focused *Bill* across the intervening quantifier, compare the schematic structure in (49). The quantifier is one of the elements that bring a ∼ with them. The ungrammaticality of (48B) thus follows from MIN, or in other words, focus evaluation itself shows the expected minimality effect.

[3]Note that not all alternative evaluating operators trigger this Minimality Effect; the Q operator in particular does not. Moreover, the informed reader will wonder about multiple focus data discussed for instance in Krifka (1991) and Rooth (1996). See Beck (2006) for a detailed discussion of the circumstances under which MIN effects do occur.

(48) A: You told nobody that Maria met Sally.
B: No – I only told nobody that Maria met BILL.
?* The unique x such that I told nobody that Maria met x is Bill.

(49) * [only [$\sim$ [... [nobody [$\sim$ [$_{\phi}$ ... XP ...]]]]]]

4.2.2 MIN in question formation

The data in (50) (from Pesetsky (2000)) and (51) (from Beck & Kim (to appear); the judgement given refers to the AltQ reading.) illustrate the intervention effect in questions.

(50) a. ?* Which book didn't which student read?

b. ?* Which boy did only Mary introduce which girl to _?

(51) a. ?* Didn't Sue read 'Pluralities' or 'Barriers'?

b. ?* Did only Mary introduce Sue or Molly to Bill?

As illustrated in (52), all of these data are further instances of MIN, where the wh-phrase or the disjunction introduces alternatives that want to be evaluated by the Q operator, but encounter an intervening $\sim$ first.

(52) * [Q ... [only/NOT [$\sim$ [$_{\phi}$ ... wh/DisjP ...]]]]

4.2.3 MIN in -ka/-mo constructions

The datum in (53) is an instance of MIN in the Japanese *-ka*-construction taken from Shimoyama (2001). The only interpretation acceptable is the one in which both indeterminate phrases are evaluated by the most local operator. Thus the configuration in (54) is impossible. This is in fact an important motivation for the semantic system developed by Shimoyama.

(53) Taro-wa [Yamada-ga dare-ni nani-o okutta ka] tazunemasita ka?
Taro-TOP Yamada-NOM who-DAT what-ACC sent Q asked Q
a. 'Did Taro ask what Yamada sent to whom?'
b. ?* 'For which person x did Taro ask what Yamada sent x?'
c. * 'For which thing y did Taro ask to whom Yamada sent y?'
d. ?* 'For which x,y did Taro ask whether Yamada sent y to x?'

(54) * [... [[... indeterminate pronoun ...] *-mo/ka*] ...]-*mo/ka*

The minimality effect for *-mo* can be illustrated with the following example from Yatsushiro (2005). Interpretation (55'a) in which the indeterminate phrase would associate with the higher *-mo* is impossible. It must be evaluated by the lower *-mo*, with the higher *-mo* being interpreted as an additive particle.

(55) [[dare-no kaita] hon-mo syookai-sita] hito-mo paatii-ni kita.
who-GEN wrote book-MO introduction-did person-MO party-to came

(55') a. * For every x: the person who also introduced the book that x wrote came to the party.

b. The person who introduced the book that everyone wrote also came to the party.

4.2.4 MIN in NPI licensing

Intervention effects in NPI licensing have been known since Linebarger's work. An example is (56) which shows that a quantifier cannot intervene between an NPI and its licenser. Linebarger (1987) accordingly proposes the Immediate Scope Constraint.

(56) I didn't give Joe/*most people a red cent.

(57) Immediate Scope Constraint (Linebarger (1987)):
A negative polarity item is acceptable in a sentence S if in the LF of S the subformula representing the NPI is in the immediate scope of the negation operator. An operator is in the immediate scope of NOT only if (i) it occurs in a proposition that is the entire scope of NOT, and (ii) within this proposition there are no logical elements intervening between it and NOT.

In our terms, what is excluded is a structure such as (58) in which the licenser 'not even' wants to use the alternatives introduced by the NPI, but those are caught by the intervening $\sim$ that accompanies the quantifier 'most people'.

(58) * [not even [$\sim$ [... [most people [$\sim$ [$_{\phi}$ ... NPI ...]]]]]]

4.2.5 MIN in either... or *constructions*

Finally, Beck & Kim (to appear) detect MIN in *either... or* constructions. An example and the relevant structure is given below.

(59) ?* Either only John ate rice or beans.

(60) * [either [only [~C [... [$_{DisjP}$ rice or beans] ...]]]]

In conclusion, an alternative semantics for the constructions discussed receives additional support from the fact that MIN is observed, as expected under such an analysis.

4.3 Scope vs. Minimality

Suppose we find that a construction obeys the constraint MIN. This is an argument in favor of an alternative based interpretation of that construction in so far as MIN is different from the constraints we would expect otherwise. That is, MIN is empirically interesting to the extent that methods of compositional interpretation not based on alternatives would lead to different restrictions. Thankfully, this is indeed the case, as I illustrate below. I compare standard scope taking of regular (non-alternative based) quantifiers with quantification over alternatives.

4.3.1 Scope taking by standard mechanisms possible where alternative evaluation is not

(61) is an example of inverse scope (taken from Heim & Kratzer (1998)). The relevant interpretation paraphrased in (61b) is one in which a quantified DP takes scope over another DP that it is syntactically contained in. I assume that that interpretation is derived from the Logical Form in (61c) in which the inner DP has been raised out of its containing DP. This is not particularly important though; there could be another way to compositionally derive the inversely linked reading. The point is that English does in fact allow the relevant interpretation, and hence we have to have a method of scope assignment which derives it. Let me call this method 'standard scope taking'. Standard scope taking is implemented via Quantifier Raising here.

(61) a. One apple in every basket is rotten.

b. For every basket x: one apple in x is rotten

c. [every basket [1 [[DP one apple in t_1] is rotten]]]

In (62) I try to construct the same example in the form of a Japanese *-mo/-ka* construction. There are two versions, a genitive DP, (62a), and a DP with a relative clause, (62b). In both cases, there are two indeterminate phrases and two evaluating operators, one universal and one existential. The reading we are after is one in which the universal would take wide scope (hence *-mo* is the outermost morpheme on the DP). Contained inside the universal quantification is an existential quantifier, here a *-ka* morpheme. Neither version of the example has the relevant interpretation, which would of course violate MIN.

(62) a. [[[dono kago-no dono ringo]-ka]-mo] kusatteiru.
which basket-GEN which apple has.rotten
* 'For every basket x: one apple in x is rotten'

b. [[[dono kago-ni aru dono ringo]-ka]-mo] kusatteiru.
which basket-in is which apple has.rotten
* 'For every basket x: one apple that is in x is rotten'

The comparison between (61) and (62) shows us that standard scope taking is possible in circumstances in which an alternative based interpretation is not. Standard scope taking does not obey a minimality constraint like MIN (confirming the point of view expressed in section 4.1. that English *every* does not work with an alternative semantics).

4.3.2 Alternative evaluation possible where standard scoping is not

There are also instances of the reverse situation, alternative interpretation being possible while standard scoping is impossible. The data in (63) and (64) are from Shimoyama (2001) and show that alternative evaluation is not constrained by syntactic islands. In (63) universal quantification takes scope outside a relative clause in which the Alt-trigger is contained. In (64) an adjunct clause separates the Alt-trigger and the alternative evaluating operator. These examples, notice, mirror the well-known fact that focus evaluation can cross island boundaries in English (compare e.g. Rooth (1985)).

(63) [[[[Dono T.A.-ga _ osieta] gakusei]-ga _ syootaisita] sensei]-mo
which T.A.-NOM taught student-NOM invited teacher-MO
kita.
came

'For every T.A. x: the teacher(s) that the student(s) that x had taught invited came.'

(64) [[_ [Taro-ga nani-o katta-kara] okotta] hito]-mo
Taro-NOM what-ACC bought-because got.angry person-MO
heya-o deteitta.
room-ACC left

'For every thing x, the people who got angry because Taro had bought x left the room.'

(64') John only got angry because Taro had bought ICE AGE II.
The only x such that John got angry because Taro had bought x is Ice Age II.

In (65) and (66) I construct parallel English examples. They do not have the relevant interpretations. Standard scope taking cannot take the quantifier outside the relative clause island or the adjunct island. This is expected under a movement analysis of scope assignment.

(65) The teacher that the student that every TA had taught invited came.
* For every T.A. x: the teacher that the student that x had taught invited came.

(66) The people who got angry because Taro bought everything left the room.
* For every thing x, the people who got angry because Taro had bought x left the room.

Thus once more the empirical conditions under which the relevant reading of the universal quantifier is available are different. Hence Shimoyama argues for employing two different strategies of compositional interpretation (contra competing analyses of Japanese that treat the Japanese construction parallel to quantification in English).

4.4 Focus Sensitive Particles in Mandarin

I will end this paper with a short case study of Mandarin that confirms important aspects of the view of alternative evaluation we have arrived at largely on the basis of English. In addition, it provides an opportunity to raise some questions regarding cross-linguistic aspects of alternative evaluation. My discussion is based on Hole (2004), who offers an extensive and detailed overview

of focus particles in Mandarin. Mandarin has a set of particles including *cai*, *jiu*, *dou* and *ye*, all of which seem to be able to act as focus sensitive particles. I concentrate on *ye* ('also'/'even') for this exposition. All data are taken from Hole (2004) ((67a): p. 189 (109); (67b): p. 38 (64a)).[4]

(67) a. Lao Li lian XINGQITIAN ye gongzuo
old Li even Sunday YE work
'Old Li works even/also on SUNDAYS.'

b. (Lian) NUWANG ye hui lai
(even) queen YE will come
'Even the QUEEN will come / the QUEEN will come, too.'

Hole notes that in addition to interacting with a focus, the particle *ye* can associate with an NPI, an indeterminate pronoun and a disjunction, as illustrated below. Hence we have an overt element here that can evaluate alternatives from all the various expressions we have identified as Alt-triggers. In Mandarin there is thus obvious motivation for the idea that the constructions considered have a common semantics, from the use of the same morpheme. Notice that this morpheme is the alternative evaluating operator, not the Alt-trigger (unlike in the morphological evidence we considered in section 4.1.). ((68a)[NPI *yi-di*]: p.198 (118b); (68b/b')[NPI *zai*]: p.201 (123); (68c)[indeterminate pronoun *shenme*]: p.204 (127b); (68d)[disjunction *xia bu xia*]: p.218 (138))

(68) a. Ta (lian) YI-DI jiu ye mei he
(s)he (even) 1-CL:drop wine YE not.have drink
'(S)he hasn't even had a drop of wine.'

b. Wo bu zai qu le
I not once.more go PRT
'I won't go there anymore.'

b'. Wo ZAI ye bu qu le
I once.more YE not go PRT
'I'll never ever go there again.'

[4]I omit diacritics. Hole indicates focus in the Mandarin example through capitalization, which I have followed. Sometimes I have adapted the English translation according to what Hole says in the text surrounding the example. Hole argues that while the most appropriate translation of a sentence with *ye* is often with *even*, the actual contribution of the *ye*-construction is closer to *also*. I follow him there.

c. Lao Li SHENME ye mei mai
old Li what YE not.have buy
'Old Li hasn't bought anything at all.'

d. Ta buguan xia bu xia yu ye hui lai
(s)he no.matter fall not fall rain YE will come
'(S)he'll come no matter whether it rains or not.'

Hole observes that while it is possible to have two such particles in one sentence, the constraint in (69a) holds. This is once more the now familiar MIN condition. Some relevant data are given below. There are other interveners besides focus particles, in particular quantifiers, motivating part (69b)(cf. p.276f. (44)-(47)).

(69) a. * $F_i \ldots F_j \ldots PART_{i/j} \ldots PART_{i/j} \ldots$

b. * $F_i \ldots$ Quantifier $\ldots PART_i \ldots$

(70) Zhiyou TIANQI HAO, women cai lian BINGQILIN ye hui chi
only.if weather good we CAI even ice-cream YE will eat
'Only if the weather is good will we even eat ICE-CREAM.'

(71) * Zhiyou TIANQI HAO, women lian BINGQILIN {cai ye / ye cai}
only.if weather good we even ice-cream CAI YE / YE CAI
hui chi
will eat
'Only if the weather is good will we even eat ICE-CREAM.'

(72) Zhiyou TIANQI HAO, (*pingchang) women (*pingchang) cai
only.if weather good (usually) we (usually) CAI
(*pingchang) lian BINGQILIN (*pingchang) ye hui chi
(usually) even ice-cream (usually) YE will eat
'Only if the weather is good will we usually even eat ICE-CREAM.'

Within the present framework, I propose (as a first approximation, following Hole's suggestions) the semantics in (73) for the particle *ye* (this is for the association with focus use; under a Krifka/Lahiri analysis it should extend fairly straightforwardly to NPIs and indeterminate pronouns, but the analysis of the disjunction example is not completely clear to me).[5]

[5]I remain silent on the details of the syntactic structure of this construction. Hole (2004) argues that the particle is only a morphological reflex of an *also*-focus, but does not carry the

For our purposes, the important properties are that *ye* evaluates the alternatives contributed by its complement, and it resets the alternative semantic value. These are the same properties that in the case of the $\sim$ operator are responsible for MIN, and they account for the analogous effect (69).

(73) *The semantics of* YE:

a. $[\![\text{YE XP}]\!]^o$ is only defined if there is a p in $[\![\text{XP}]\!]^{Alt}$ such that $p \neq [\![\text{XP}]\!]^o$ and p is true. Then, $[\![\text{YE XP}]\!]^o = [\![\text{XP}]\!]^o$

b. $[\![\text{YE XP}]\!]^{Alt} = \{[\![\text{YE XP}]\!]^o\}$

The case of Mandarin focus particles raises some interesting questions regarding the cross-linguistic side of focus evaluation.

Generally, there is the question in how far there is variation in the semantics of focus evaluating operators. Rooth (1996) suggests that different languages or possibly even different constructions in the same language might use different alternative evaluating operators, for instance ones with a stronger semantics than the $\sim$. Still, one might speculate (so Rooth) that something like the $\sim$ is present as a common core whenever we evaluate alternatives. The Mandarin operator called YE above would not be identical to the English focus evaluating operator, having a stronger semantics incorporating *also*.

A related issue, though, is the question of what precisely the alternative evaluating operator is. Rooth's (1992) theory for English is fairly abstract in postulating just one focus evaluation operator, which indirectly accounts for all the apparently focus sensitive elements in the language like *only*, *even*, *also*, quantifiers and so on. I have for the moment opted for a more direct semantics for *ye*, and also Japanese *-mo* and *-ka*, in which the particles themselves are the alternative evaluating operators. It seems though that if a language has one operator that appears to have direct access to alternatives (like *-mo* and *ye* on the semantics actually given here, without mediating operator), then that language has a few more (like *cai*, *jiu* etc. in Mandarin, and at least *-ka* in Japanese). One might try to unify these particles under one evaluation operator, which would then have a semantics weak enough to combine with all of them. I do not know what the right strategy is.

Another issue is selectivity. The question operator Q in English and German, for instance, can only evaluate *wh*-phrases. Mandarin YE on the other

semantics. He does not provide an interpreted Logical Form. His proposal is compatible with what I say in the text if YE is not the morpheme itself, but the locus of the *also*-semantics of which *ye* is the reflex. Mandarin particles would then be similar to the abstract ONLY proposed for Korean in Lee (2004).

hand can access various Alt-triggers. See Kratzer & Shimoyama (2002) for some discussion of the issue of particular operators evaluating alternatives introduced by particular Alt-triggers (although it should be pointed out that Kratzer & Shimoyama do not address focus, and therefore do not have a two-tiered semantic system but only one tier corresponding to alternative semantic values. Including focus would complicate the discussion somewhat).

A final question is which alternative evaluating operators trigger a MIN effect. We have seen that the $\sim$ and *-mo* and *ye* do. Beck (2006) shows that the Q operator does not. Is this predictable somehow, or a lexical accident?

5 Consequences

The use of alternatives in natural language semantics seems to be a wide-spread phenomenon. At the same time, there is variation wrt. where a language uses an alternative semantics. Universal quantification can be expressed in this way by Japanese *-mo*, but not English *every*. The different 'or's in French suggest that items with the same ordinary meaning can have different properties regarding their alternative semantics. This means that use of alternatives is not a definitive demand made by any cognitive system outside of the grammar – at least not in every particular case – hence the variation. Rather, it is part of grammar to determine where such a semantics is used. The variation is in the interpretation component of the grammar.

My personal view is that all parts of the grammar leave room for variation. The aspects of interpretation that are external to grammar do not dictate a unique syntax-semantics interface. The use of alternatives is one example of an area where choices are left open. We need to take the syntax-semantics interface seriously as a component of grammar. Once this is established, it seems that the role of recursion in the equation "Interfaces + Recursion = Language" is overstated. Grammar is not limited to recursive structure building. The way in which structure is associated with meaning – the syntax-semantics interface – is an equally important part of the grammar. Regarding that interface, i.e. interpretation component of the grammar: It must certainly be recursive, but that doesn't seem to be the issue regarding the questions that I am interested in in that domain, concerning language variation, interpretation principles available, systematic lexical variation etc.

References

Abraham, Werner (ed.)
1991 *Discourse Particles.* John Benjamins, Amsterdam.

Aloni, Maria
2003 Free Choice in Modal Contexts *Proceedings of Sinn und Bedeutung*, 7: 25–37. available at: http://ling.uni-konstanz.de/pages/conferences/sub7

Beck, Sigrid
1996 *Wh-Constructions and Transparent Logical Form.* Ph.D. thesis, University of Tübingen. available at: http://www2.sfs.nphil.uni-tuebingen.de/Alumni/Dissertationen/beck.pdf
2006 Intervention Effects Follow from Focus Interpretation. *Natural Language Semantics*, 14: 1–56.

Beck, Sigrid, and Shin-Sook Kim
to appear Intervention Effects in Alternative Questions. *Journal of Comparative Germanic Linguistics.*

Beck, Sigrid, and Shravan Vasishth
in prep. *Multiple Focus.* Manuscript, University of Tübingen & University of Potsdam.

Geurts, Bart and Rob van der Sandt
2004 Interpreting Focus. *Theoretical Linguistics*, 30: 1–44.

Guerzoni, Elena
to appear *Intervention Effects on NPIs and Feature Movement: Towards a Unified Account of Intervention.* Manuscript, MIT.

Hagstrom, Paul
1998 *Decomposing Questions.* Ph.D. thesis, MIT.

Hamblin, Carl Lee
1973 Questions in Montague English. *Foundations of Language*, 10: 41–53.

Haspelmath, Martin
1997 *Indefinite Pronouns.* Oxford University Press, Oxford

Heim, Irene and Angelika Kratzer
1998 *Semantics in Generative Grammar.* Blackwell, Oxford

Hole, Daniel
2004 *Focus and Background Marking in Mandarin Chinese.* Routledge Curzon Asian Linguistics Series, London and New York

Karttunen, Lauri
1977 Syntax and Semantics of Questions. *Linguistics and Philosophy*, 1: 3–44.

Kratzer, Angelika and Junko Shimoyama
2002 *Indeterminate Pronouns: The View from Japanese.* available at http://semanticsarchive.net

Krifka, Manfred
1991 A Compositional Semantics for Multiple Focus Constructions. *Proceedings of SALT*, 1: 127–158.
1995 The Semantics and Pragmatics of Polarity Items. *Linguistic Analysis*, 25: 209–257.
Lahiri, Utpal
1998 Focus and Negative Polarity in Hindi. *Natural Language Semantics*, 6: 57–123.
Lappin, Shalom (ed.)
1996 *The Handbook of Contemporary Semantic Theory*. Blackwell, Oxford.
Lee, Youngjoo
2004 Scope of Focus Particles: Abstract ONLY in Korean. *Proceedings of Sinn und Bedeutung*, 8: 167–180. available at: http://www.ub.uni-konstanz.de/kops/volltexte/2004/1383
Pesetsky, David
2000 *Phrasal Movement and its Kin*. MIT Press
Rooth, Mats
1992 *Association with Focus*. Ph.D. thesis, University of Massachusetts at Amherst
1998 A Theory of Focus Interpretation. *Natural Language Semantics*, 1: 75–116.
1996 Focus. In Lappin (1996), pp. 272–297.
Sauerland, Uli
2005 Don't Interpret Focus. *Proceedings of Sinn und Bedeutung*, 9: 370–384.
Shimoyama, Junko
2001 *Wh-Constructions in Japanese*. Ph.D. thesis, University of Massachusetts at Amherst
Simons, Mandy
2004 *Dividing Things Up: The Semantics of 'or' and the Modal/'or' Interaction*. Paper presented at NELS 35.
von Stechow, Arnim
1991 Focusing and Backgrounding Operators. In Abraham (1991), pp. 37–84.
Yatsushiro, Kazuko
2005 *The Distribution of Suffixes of Quantifiers in Japanese*. Manuscript, ZAS Berlin.
Zimmermann, Thomas Ede
2000 Free Choice Disjunction and Epistemic Possibility. *Natural Language Semantics*, 8: 225–290.

List of Contributors

Prof. Dr. Artemis Alexiadou
University of Stuttgart
Institute for Linguistics
Keplerstraße 17 / 4b (KII)
70174 Stuttgart
Germany

Prof. Dr. Elena Anagnostopoulou
University of Crete
Department of Philology
74 100 – Rethymno
Greece

Prof. Dr. Sigrid Beck
University of Tübingen
Department of English Philology
Wilhelmstraße 50
72074 Tübingen
Germany

Prof. Dr. Noam Chomsky
Massachusetts Institute of Technology
Department of Linguistics and Philosophy
77 Massachusetts Avenue, 32-D808
Cambridge, MA 02139
USA

PD Dr. Hans Martin Gärtner
Center for General Linguistics
Schützenstraße 18
10117 Berlin
Germany

Dr. Jens Michaelis
University of Osnabrück
Institute for Cognitive Science
49069 Osnabrück
Germany

Prof. Luigi Rizzi
University of Siena
Interdepartmental Center of Cognitive Studies on Language
Via Roma, 56
53100 Siena
Italy

PD Dr. Uli Sauerland
Center for General Linguistics
Schützenstraße 18
10117 Berlin
Germany

Prof. Dr. Ur Shlonsky
University of Geneva
Departement of Linguistics
Rue de Candolle 2
1205 Genève
Switzerland

Index of Languages

≠Hoan, 41, 46

Arabic, 35, 37

Basque, 124
Brazilian Portuguese, 129, 141
Bulgarian, 70, **72**, 77, **102–103**, 189

Catalan, 35, 38, 50
Celtic, 35, 37, 49, 50, 147

Danish, 65, 67, 69, 71, 72, 91, 94–97, 104

English, 22, 32, 36, 37, 49, 50, 52, 63, 65, 69, 70, 73, 76, 83, 87, 88, 116, 126, 143–146, 148–151, 219, 222, 264, 266, 269, 272–275, 277, 278
  West-Ulster, 144

Flemish, 132
French, 32–34, 36–38, 46, 50, 115, 116, 130, 132, 133, 135, 138, 140–144, 146, 148, 149, 151, 153, 267, 278
  Québec, 131, 142

German, **39**, 54, **70–72**, 74–77, 83, 94, **97–102**, 106, 127, 152, 205, 227, 238, 269, 277
Germanic, 129, 138, 146
Greek, **38**, 49, 50, 54
Gungbe, 141

Hebrew, **120–123**, 140, 152

Icelandic, 35–37, 46, 49, 50, 65, 66, 68, 73, 78, 81, 91, 94–96
Imbabura Quechua, **123–125**, 141
Italian, 35, 38, 50, 117, 119, 122, 123, 127, 129, 130, 153
  Old, 129

Japanese, 10, 124, **258–259**, 266, 268, 270, 273, 274, 277, 278
Ju'hoansi, 41–43, 45, 46

Khoisan, 31, 41, 53, 55
Kilega, **39–40**, 54
Kinande, **41**, 55
Korean, 70, 83, 277

Malayalam, 268
Mandarin, 268, **274–278**

Romance, 37, 129
Romanian, **38**, 54

Scandinavian, 65, 67, 69–71, 75, 76, 95, 97, 98, 100, 153
Sign Languages, 13
Sinhala, 124
Spanish, **38**, 54, 129, 153
Swedish, 66

Index of Subjects

A′-movement, 17, 18, 24, 25, 69, 96, 97
A-movement, 18, 19, 21–25, 67, **72–73**, 76, 77, 97, 98, **103**, 128, 137, 145
adverb
  intervening, 93
Agree
  counter-cyclic, 52
  cyclic, **51–53**, 53
  multiple, 9
agreement
  number, **133**, 138
  past participle, 138
  probe-goal, 22, 138
  redundant, 18
alternative
  ellipsis —, 220, 224
  focus —, 219, 235, 244
alternative semantics, 255–258, 261–263, **265–269**, 272, 273, 277, 278
anti-symmetry, 74

binding
  semantics of, 197–247
binding theory, 9, 67

Case
  quirky, 22
  structural, 17–19, 26, 32, 41, 46, 55, *see also* feature, structural Case
Case checking, 31, **35–37**, 38, 43, 46, 49, 51–53
clitic doubling, 38, 39
complement, 9, **11**, 17, 24–26, 70, 82, 83, 85, 90, 93, 94, 116, 146, 147
concept
  individual, 15, 199, 204, 208, 214–215
  maximal, 214
counter-cyclic, 19, 53
cyclic linearization, *see* linearization, cyclic
cyclicity, 1, 10, 24, 51, 63
  counter, *see* counter-cyclic
  strict, 5, 16, 17, 24, 53

D-linking, 118
definite description, 203, 204, 206, 209, 210, 212, **215–216**, 217, 218, 221, 223, 225, 229, 231, 233, 237, 238, 245, 246
definiteness, 25, **133–134**
derivational organization of grammar
  strictly, 62, 107
derivational syntax, **77–78**
disjunction, 255, **260–261**, 261, 265, 267, 270, 275, 276

edge, 11, 12, 17, 21, 62–64, 67, 69, 73, 75, 76, 78, 80, **86**, 88, 89, 93, 102, 103, 105, 126
edge feature, **11**, 14, 24, 25
EF, *see* edge feature
ellipsis, 210, 219–220, 235, 237, 245
EPP, 24, 31, **48–153**

evolution of language, 3–5, 7, 14
expletive *il*, 133, 134
expletive *pro*, 119, 127, 129, 130, 141
expletive *there*, 126
Expletive Drop, 145
externalization, 13–15, 31, 36, 48, 49

feature, **6**, 20, 23, 62, 65, 78, 79, 103, 118, 137–139, 147, 152
  and intermediate movement, 79
  availability, 80
  criterial, 118
  D, 88, 103
  EPP, 31, 37, 49–51, 53, 62, 64, 78, 79, 98, 105
  finiteness, 149
  force, 149
  goal, 79
  inflectional, 20–22, 24
  inheritance, 23, 25, 26
  interpretable, 18
  non-local, 85, 88, 105
  number, 6, 7, 11, 18–20, 135–137
  phi-, 20, 22, 38, 44, 51, 52, 132, 143, 146, 147, 149, 152, 212, 226
  π, 91
  probe, 79, 88, 92–94, 99, 104
  ψ, 85
  q-, 152
  selectional, 24
  stem, 18
  structural Case, 31, 35, 36, 38, 41, 47, 48, 55
  subcategorization, 85, 86, 91, 93
  top, 105
  uninterpretable, 11, 18–24
  wh, 64, 80, 81, 125
  wh-, 144
focus, 18, 116, 117, 128, 129, 143, 150, **220**, 235, 243, 255, 256, **257**, 257, 261–266, 268, 269, 273–278
for–trace effect, **146**, 147
force, 31, 41, 46, 148–151

government
  proper, 115, 116, 127, 129, 132, 147
Grice, 13

Holmberg's Generalization, **65–67**, 68, 75, 78, 90, 95, 97

indefinite, 25, 26, 133, 261, 267
indexical, 211, 221
infinitive
  coherent, 101
interface conditions, 5, 7, 11, 17, 18, 21
intervention effect, 22, 256, 270, 271
Inversion
  Quotative, **32**, **34–35**, 49
  Stylistic, **32–34**, 46, 49

landing site, 31, 44, 45, 49–51, 64, 67, 69, 71, **76**, 77, 98, 99, 103, 106, 146
Landman's generalization, 245
linearization, 9, 54, 63, 64, 67, 69, 72, 75, 83, 86, 88, 90, 93–96, 99, 101, 105
  cyclic,**61–65**, 67, 69, 71, 76–78, 86, 88, 90, 94, 104, 106, 107

relativized, 62, **82**,**84**, 92
VP, 64, 67–69, 75, 88
locative, 40–47, 49
logic, 129, 271
logic, combinatory, 12, 197, 198, 202
logical form, 200, 201, 203, 204, 213, 216, 221, 222, 224, 227–229, 243

Merge, 5, **5**, 6–11, 14, 16, 17, 62, 63, 65, 83–96, 98–104, 106
pair-Merge, 8
parallel, 8
set-Merge, 8
unbounded, 8, 11, 14
minimality, 61, 73, 105, 256, 268, 269, 271, **272–274**
relativized, 61, 132, 137
movement, 31, 33, 35, 37, 39, 40, 49, 52, 62, 64, 65, 67–71, 73, 75–82, 85, 86, 88–94, 96–101, 103–107, 116, 117, 119, 123–126, 129, 130, 135–138, 140–146, 152, 153, 221, 274
covert, 12, 69
improper, 25, 67, 76
intermediate, 62–64, 67, 69, 79, 107
overt, 12, 38, 69
partial, 127, 128
successive-cyclic, 17, 25, 62–64, 73, 76, 78, **78**, 79–82, 84, 86
multiple-agree, *see* Agree,multiple

negation polarity item, **259–261**, 267, 268, 271, 275, 276
negative polarity, 255, 256, 271
NPI, *see* negation polarity item
Null Subject Language, 119, **126**, 127, 129, 132, 133, 141, 144, 153

object
direct, 33, 41, 46, 72, 95, 96, 100–102
indirect, 34, 72, 94–96, 100, 102
raising to, 20, 99
secondary, 41
vP-internal, 31, 35, 37–39, 46, 47, 50, 51, 81, 100
object movement, 31, 32, 34, 41–50, 75, 91, 95, 96, 106, 116, 124, 132, 137
object shift, 49, **65–67**, 71, 73, 76, **90–94**, 94–98, 100, 107
multiple, **67–69**, 70, **94–97**, 107
only, 255, 256, 264, 269, 277
ordering, 1, 5, 6, 9, 10, 22, 53, 54, 62–64, 67–70, 72–74, 77, 78, 82–84, 87–96, 98–107, 140, *see also* linearization

parasitic gap, 76, 87, 97
participial, 22, 50
particle
focus sensitive, 264, 269, 274–276
passive, 17, 22, 72, **72**, 73, 76, 77, 89, 103, **103**, 145
Percus's generalization, 245
phase, 5, 9, 10, **16**, 17–26, 54, 61–65, 67, **69–70**, 71–83, 86, 88–90, 93, 103–105, 107, 125
strong, 54

Phase Balance, **79**, 80–82, 85, 88–92, 99, 100, 102–106
phase-impenetrability condition, 16, 17, 19, 24, 62, 65, 69, 77, **78**, 81–83, 86
PIC, *see* phase-impenetrability condition,
pied-piping, 37, 38, 117, 119, 123, **123**, 124–126, 153
  clausal, 124, 126
plurality, 144, 212, 215, 216, 224, 226
predication, 14, 197, 241
presupposition, 257
pronoun, 67, 68, 70, 71, 74, 76, 77, 91, 93–95, **97**, 98–102, 104, 106, 107, 122, 123, 140, 197–247, 269, 271, 275, 276
  de re, 241
  de se, 222, 241, 244
  resumptive, 13, 119, 120, **120**, 121–123, 140

quantification, 11, 24, 201, 202, 206, 258, 259, 266, 267, 269, 271–274, 278
quantifier
  floated, 33, 34, 130, 144, 145
  universal, 217, 223, 266
quantifier raising, 76, 272
quasi-expletive, 142, 148–151
que-qui, 130, **130**, 132, 136–139, 141, 142, 153
question, 69, 71, 81, 119, 131, 141, 142, 144, 150–152, 255, 256, **257–258**, 261, 270, 277, 278
  multiple, 72
  subject, **140**, 141, 143, 144, 152
  wh, *see* wh-movement

recursion, 5–8, 124, 198–200, 202, 203, 213, 215, 278
reflexivization, 9, 18
reflexives, 23
relative clause, 32, 62, 65, 78, 121, 122, 130–133, 273, 274
repetitions
  independent, 10

salience, **215**
scope, 12, 21, 23, 24, 116, 118, 124, 127, 128, 138, 141, 143, 153, 256, 263, 265, 266, 268, 269, 271, **272–274**
scrambling, 70, **106**, 144
semantics
  duality of, 10–12, 15, 17, 22
sluicing, 206
specifier, **11**, 17, 26, 62, 80, 81, 86–96, 98–102, 104–106, 125, 138
spell-out domain, *see* phase
SSG, *see* subject-in-situ generalization
subject
  vP-internal, 31–35, 37–39, 41, 46, 47
Subject Criterion, **116**, 118, 119, 121, 124–128, 130, 132, 135–139, 141, 143, 147, 149, 152, 153
subject extraction, 115, 116, 119, 123, 126, 127, 129–131, 136, 141, **146**, 147–149, 151–153

subject focalization, 130
subject-in-situ generalization, **31–53**, 54, 55
syntactic islands, 273

TEC, *see* Transitive Expletive Construction
that-trace, 129, **148**
TP, 18–20, 23, 31, 35, 48–50, 53, 65, 67, 69, 70, 73, 74, 81, 82, 87, 89–91, 94, 97–99, 102–105
trace, 88, 96, 115, 142, 144, 201, 202, 204, 215, 221, 222, 244, 246
Transfer, 11, 16, 18, 19, 23, 24
Transitive Expletive Construction, 35, 37, 50, 96, 133
truncation, 150, 151

verb-second, 65, 66, 69, 71, **73–75**, 77, 94, **103–105**

weak crossover, 21
wh-movement, 33, 63, 64, 67, 69, 70, **72**, 76, 77, 81, **89–90**, 96, **102–103**, 107, 118, 123–127, 141, 145
  successive cyclic, **63–65**, 121, 125