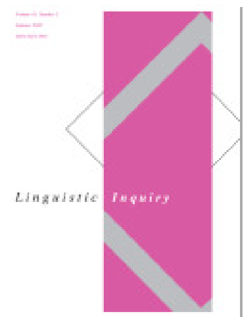
On Feature Inheritance: An Argument from the Phase Impenetrability Condition

Richards, Marc D.

Linguistic Inquiry, Volume 38, Number 3, Summer 2007, pp. 563-572
(Article)

Published by The MIT Press



Adriana Belletti and Luigi Rizzi, 11–62. Oxford: Oxford University Press.
Travis, Lisa. 1984. Parameters and effects of word order variation. Doctoral dissertation, MIT, Cambridge, Mass.
Zwart, C. Jan-Wouter. 1997. *Morphosyntax of verb movement: A minimalist approach to the syntax of Dutch*. Dordrecht: Kluwer.

On Feature Inheritance: An Argument from the Phase Impenetrability Condition

Marc D. Richards
University of Cambridge

This squib is organized as follows: the first part (section 1) identifies a potential conceptual flaw in the feature inheritance model of phase theory proposed in Chomsky 2005; the second part (sections 2–4) suggests a revision that removes this flaw.

1 Feature Inheritance, (Non)phase Heads, and the Strong Minimalist Thesis

Chomsky (2005; henceforth OP)[1] proposes a reinterpretation of the relation between the functional heads C and T: the Agree (ϕ-) and Tense features associated with the inflectional system are not an inherent property of T; instead, they belong to the phase head C. Traditional subject agreement and EPP (Extended Projection Principle) effects associated with T (A-movement of the formal subject to Spec,T, expletives, etc.) then arise via a mechanism of *feature inheritance*, whereby uninterpretable features are passed down from the phase head to its complement. It follows that T lacks uninterpretable features unless it is selected by C. That is, T is no longer a probe in its own right; it cannot initiate operations directly or independently of C.

Clearly, in this way, feature inheritance captures the long-standing observation that raising/ECM (exceptional Case marking)-infinitival T, which lacks C, also lacks ϕ-features (failing to value Case on DP) and independent tense (see MI:102, 105, BEA:13, OP:9). However, where the previous system had to stipulate this connection by means of a selectional restriction (C selects ϕ-complete T; V selects ϕ-defective T), the feature inheritance model offers an arguably more explanatory account of T's featural dependence on C: the features are simply C's, not T's. This, in turn, allows a uniform characterization of phase heads (C, v*) as the locus of uninterpretable features, as is desirable on computational grounds (see section 2).

Nevertheless, there is a major hurdle to be overcome before feature inheritance can be accepted as a principled property of the language system. Whereas T now clearly needs C in order to function at

My thanks to Noam Chomsky for discussion and encouragement, and to two anonymous reviewers for helpful comments. Remaining errors and omissions are my own.

[1] For convenience, the following abbreviations are used throughout: MI = Chomsky 2000, DbP = Chomsky 2001, BEA = Chomsky 2004 (pages cited from the ms. version, 2001), OP = Chomsky 2005.

all (e.g., in the Agree system), it is not at all obvious why C needs T in this feature inheritance model. In short, why should feature inheritance (and indeed T itself) exist at all? In order to conform to the Strong Minimalist Thesis (SMT), feature inheritance must be forced by considerations of good design. Unless it provides the optimal means of satisfying an interface condition or facilitates computation in some other way, feature inheritance is a redundant, extra device, a departure from optimal design that must be attributed to the unexplained residue of Universal Grammar. The challenge, then, is to find a principled reason why the Agree feature cannot simply remain on C.[2]

To this end, Chomsky (OP:9–10) offers an argument from the conceptual-intentional (C-I) interface: feature inheritance follows from the C-I-imposed requirement that the A/Ā distinction be structurally established. Spreading of Agree to T enables a structural distinction to be made between the A-position created by movement for C's Agree feature (Spec,T) and the Ā-position created by movement for C's edge feature EF (Spec,C, the phase edge). Further, once motivated in this way for the C-T relation, feature inheritance may reasonably be expected to hold of phase heads in general, on grounds of optimal design (OP:14). Bearing out this prediction, Chomsky notes that spread of Agree features from the v* phase head to its complement V, with A-movement of the object to Spec,V analogously with subject raising to Spec,T, immediately yields the famous ''raising to object'' paradigm that obtains with ECM, in which the embedded subject can bind into and take scope over matrix adverbials, indicating raising to a position in the matrix vP (see Postal 1974, Lasnik and Saito 1991, among many others, for examples, discussion, and alternative analyses). This movement has previously seemed unmotivated from the economy perspective, but now falls straightforwardly into place.

Although these arguments are suggestive, unfortunately neither has the force of necessity. The C-I-imposed A/Ā requirement would seem to imply only that we need two different types of features: the Agree type, yielding A-positions, and the EF type, yielding Ā-posi-

[2] Other possible objections to feature inheritance can be easily dismissed. The apparent countercyclicity of the operation is trivial once it is assumed that all operations (Agree, internal Merge, Transfer, as well as inheritance) are driven by the phase head, since only a single notion of cycle remains on this approach: the phase cycle. Thus, at the level of the phase, operations are unordered with respect to each other (there can only be ordering between phases themselves, not within them). The featural/locus/probe cyclicity of N. Richards (1999), Collins (2002), and others can then be formulated only in terms of phases (probe cyclicity and phase-level cyclicity having now become identical); see section 2. Feature inheritance also appears to violate Inclusiveness and/or the No Tampering Condition (NTC) of OP:5ff.: both C and T are ''tampered with,'' in that features are lost from C and added to T. However, the NTC in OP is a condition only on Merge (internal and external); that is, it constrains operations involving edge feature–type rather than Agree-type features (which have to acquire values, a kind of tampering that, like feature inheritance, does not violate Inclusiveness/the NTC).

tions. Given the possibility of multiple specifiers (itself due to EFs, as Chomsky has proposed[3]), there would seem to be a simpler and thus more optimal alternative to feature inheritance: both of C's features can be satisfied ''in situ'' on C, with the Agree feature raising its goal DP to form a first specifier of C (the head of the A-chain), and EF-driven movement creating a second, outer Spec,C (the Ā-/operator position).[4] Feature inheritance, and thus the very need for a separate T projection to exist at all, would still seem an extra stipulation on top of this. Further, although the generalization of feature inheritance to a property holding of all phase heads is indeed desirable for the reasons discussed above (OP:14), as yet it does not seem inevitable. The A/Ā rationale for feature spreading only holds for C-T; thus motivated, nothing actually forces it to extend to v*-V. Chomsky (OP:14) cites grounds of efficient computation, but these seem inconclusive to me: it would surely be at least equally efficient and optimal for feature inheritance *not* to obtain between v* and V, where the motivation (A/Ā) is absent.

What is lacking, then, is a general rationale for feature inheritance that holds by necessity and equally for all phase heads alike. In other words, the question raised above—why does C need T in a system where all features belong to C and all operations are triggered by C?—should be extended to the phase-nonphase relation in general. If phase heads do everything, then why do nonphase heads exist at all? Why does the computational system not consist exclusively of the phase heads? Without featural content of their own (beyond a categorial and/or edge feature, for selection and projection), nonphase heads like T become Agr-like dummy categories, in apparent violation of Full Interpretation. For such proxy heads to conform to the SMT in the OP system, then, there must be something that renders the existence and presence of these vacuous placeholders necessary. That is, nonphase heads follow from the SMT insofar as they enable the phase head to carry out its syntactic work. If the OP system of feature inheritance and all-powerful phase heads is correct, then it should be possible to show that the phase heads would actually be *unable* to do all the work on their own, without the mediation of proxy nonphase heads.

In the rest of this squib, I attempt to develop just such an argument. My claim is that feature inheritance follows by conceptual necessity from two basic assumptions about the phase-based derivational system: *Value-Transfer simultaneity* and the *Phase Impenetrability Condition.* These premises are briefly summarized in section 2, where I also expand and develop Chomsky's OP argument that Agree features belong to the phase head. The main argument for feature inheritance

[3] Class notes, LSA Summer Institute, MIT, July 2005.

[4] Assuming that the operator (head of Ā-chain) must c-command the A-chain over which it ranges, applying these two operations in the wrong order (i.e., EF preceding Agree, so that the outer Spec,C is the A-position containing the Ā-chain in the inner Spec,C) would simply yield a deviant interpretation (see OP:10–11 on overgeneration); possibly this could in any case be put right by ''tucking-in'' post-movement (see N. Richards 1999).

is then presented in section 3. Finally, some further, potentially far-reaching implications of the proposal are outlined in section 4.

2 Interpretability and Impenetrability

The OP model of the derivational system builds on the framework of multiple spell-out developed by Chomsky in MI, DbP, and BEA, extending and refining the notion of the phase. Phases are argued to be a necessary part of any well-designed language system, on two grounds: they eliminate redundant internal levels and cycles (yielding ''single-cycle generation''), and further reduce computational complexity (memory load) via the periodic ''forgetting'' (transfer to the interfaces) of derivational information, yielding a strict form of cyclicity. These general computational considerations lead to the expectation that phases should be as small as possible; the identity of the phase heads (i.e., the question of which categories are phases and which are not) then depends on supplying a principled definition of ''as small as possible,'' one that follows from interface conditions.[5] As argued in DbP:5, BEA:14, and OP:20, the relevant notion here is feature interpretability.

Given Full Interpretation (FI), uninterpretable features (uFs) must be deleted before they reach the semantic component (Sem); however, once valued by Agree, they are indistinguishable from interpretable (i.e., lexically valued) features, without reconstructing the derivation. It follows that uFs must be spelled out (transferred) as soon as they are valued if the system is to avoid lookback. If Value takes place before Transfer, the derivation will crash at Sem; if Value takes place after Transfer, then the derivation will crash at both interfaces (since unvalued features cannot be interpreted). Value must therefore be *part of* Transfer (these operations are, in this sense, ''simultaneous''); otherwise, no derivation can converge. We therefore arrive at the interface condition in (1).

(1) *Premise 1*
Value and Transfer of uFs must happen together.

As pointed out insightfully by Epstein and Seely (2002:71–72), the above logic implies that there cannot be any delay at all between Value and Transfer. Yet such a delay is unavoidable in the DbP/BEA system, in which uF (the ϕ-probe) belongs to T but is not valued until C (i.e., at the phase level, because of the simultaneity requirement). Thus, simultaneity of operations at the phase level is an unnatural, stipulated extension of the probe cycle in the DbP/BEA system, requiring a computationally dubious delay to operations initiated by the T probe (see also Epstein and Seely 2002:83).[6]

[5] For alternative views of what ''as small as possible'' should mean, see Svenonius 2001, 2004, Epstein and Seely 2002, and Müller 2004, among others.

[6] Additionally, Epstein and Seely (2002:83) criticize the apparent representationalism of this approach; Müller (2004) makes a similar point.

The OP system solves this problem via a neat twist in perspective: instead of belonging to T and being valued at C, the Agree probe belongs to C and is valued on T (because of feature inheritance). In this way, no delay is required in order to implement simultaneity; rather, operations become simultaneous at the phase level as a matter of course, since it is the phase head that initiates them all (thus, the probe and phase cycles, previously awkwardly distinct, collapse; see footnote 2). Since the OP system does without the inelegant, inefficient delay/lookahead of the previous approach, it better conforms to the SMT.

It thus follows from the SMT that uninterpretable (unvalued) features can only be a property of phase heads, that is, those heads that trigger Spell-Out/Transfer. If C is a phase head, then the uninterpretable ϕ-feature set previously attributed to T, a nonphase head, must actually belong to C (this being the simplest way to ensure that operations are simultaneous at the phase level, in conformance with (1), as argued in the previous paragraph).[7] When these features appear on T, they must therefore be derivative from C (OP:9), thus deriving the featural dependence of T on C discussed in section 1 (and once again confronting us with the question raised in section 1 of why T, and nonphase heads in general, should exist at all—the question to which this squib seeks an answer). (1) is thus fundamental to phase theory: not only does it provide a syntactic diagnostic for identifying phases (as uF sites: essentially, if X is a probe, then X must be a phase head), it also provides an argument from the SMT for why cyclic spell-out is a necessary component of any well-designed language system—namely, that without the immediate spell-out of valued probes, no expression could converge. Cyclic computation by phase thus follows from interface conditions (FI).

The second assumption underlying phase theory is best expressed in BEA:4–5: for cyclic computation to be meaningful, it cannot be the entire phasal category that is transferred, since this would exclude any continuation of the computation beyond the first phase level. For example, transfer of the entire phase would render the label (including EF) of the phase head inaccessible to further computation and therefore preclude selection of the phase by a higher head (compromising iterability of Merge). The edge (head/label and any higher, specifier material) of the phase head must therefore be carried over to the next phase, yielding an ''escape hatch'' for transphasal movement (movement out of a phase would be barred if the entire phase were transferred). The

[7] The alternative—that T is the phase head in the OP system (instead of, or as well as, C)—is argued against in BEA:21–22 and OP:18 on the basis of the intervention effect that is predicted to block Agree(T, Subject) across the copy of a raised *wh*-object in the v* phase edge (see also Müller 2004). Since no such intervention effect arises, T cannot be a phase or an independent probe (both now amounting to the same thing; see above and OP:18). A further, perhaps more compelling argument for why T cannot be a phase/probe is provided below in sections 3–4.

edge of a phase Ph_1 thus belongs to the *next* phase Ph_2 for the purposes of Spell-Out/Transfer, so that only the complement of Ph_1 is inaccessible to operations at the next phase level Ph_2. Meaningful cyclic computation, for Chomsky, thus implies a Phase Impenetrability Condition (PIC) with the following effect:[8]

(2) *Premise 2*
The edge and nonedge (complement) of a phase are transferred separately.

In sum, we now have a system in which phase heads do everything: they initiate Agree (via uF), internal Merge (via EF), and Transfer at the phase level.[9] All of these operations are driven by phase heads, and all uFs belong to phase heads. What, then, if anything, is the role of nonphase heads and feature inheritance in this system?

3 Inheritability

The observation I would like to make is that the assumptions in (1) and (2) are in fact incompatible as they stand: (2) renders the convergence condition in (1) impossible to meet. As a consequence, feature inheritance emerges from the SMT as the optimal, perhaps *only* device that can reconcile (1) and (2) and thus ensure convergence at the interfaces. The logic of the argument is simple, and runs as follows.

The assumption of Value-Transfer simultaneity in (1) requires that uF be transferred as soon as it is valued (by Agree). The syntactic life of valued uF is thus maximally short: it cannot survive into the next phase. Yet uF, as a property of phase heads (C, v*, etc.), is forced to do exactly this by the PIC/(2): the phase head is part of the edge and is thus carried over to the next phase/cycle. (2) thus renders (1) impossible to meet, with the result that all derivations should crash at Sem (see section 2)—a problem that has so far escaped notice in the literature. To take the example of phase head C, the Agree feature (uF) on C cannot be transferred at the phase level C, because the PIC ensures that C stays behind when its complement, TP, is spelled out. This should fatally violate the requirement that C's uF be transferred at the point when it is valued (here, by the subject contained in TP).

[8] The exact formulation of the PIC varies from MI to DbP, with different timing/accessibility statements leading to different patterns of search space for the nonphase head (see M. Richards 2004:chap. 3 for discussion). An interesting consequence of the OP system, in which the Agree probe belongs to C and not T, is that the PIC is no longer subject to such arbitrariness: the distinction between the MI and DbP versions of the PIC collapses as unformulable once uFs/probes are a property of phase heads, for the simple reason that T cannot do anything until C is merged and feature inheritance takes place. The PIC thus becomes a simpler, more natural, and necessary principle under the OP system—a further conceptual advantage over earlier versions of phase theory.

[9] And also, if it exists, the operation of feature inheritance (see section 3). Questions about the order in which these operations apply (as raised by a reviewer) thus do not arise, since there are no phase-internal subcycles or ordering of operations (see also footnote 2).

Clearly, this is an intolerable situation that must be resolved if the convergence condition in (1) is to be met. The solution is simple: feature inheritance, which now emerges as the only way to ensure that C's uF can indeed be valued at the same time that it is transferred. Since C will not be spelled out until the following phase level, its uF must descend onto the head that *is* spelled out by the PIC—namely, C's complement, T. That is, (1) (= interpretability) and (2) (= impenetrability) together entail (3) (= inheritability).

(3) *Conclusion*
uF must spread from edge to nonedge (i.e., from C to T, v* to V, etc.).

(3), the feature inheritance principle, has therefore been reduced to the level of principled explanation and perfect design—the system could not function without it.

4 Summary and Further Implications

The preceding sections have proposed a new rationale for Chomsky's (2005; OP) feature inheritance mechanism, one that follows by necessity and thus with greater force than the suggestive link with the A/Ā distinction put forward in OP. All that is required is that we accept the premises in (1) and (2). Although both have been challenged in the literature (see, e.g., the references in footnote 5), they arguably underlie any meaningful system of cyclic/phasal spell-out in conformance with the SMT (see section 2). With (1) and (2) in place as basic premises about the nature of phases, feature inheritance (3) is immediately deduced. By the PIC/(2), phase heads are not spelled out at the same time as their complements, and therefore uF on the phase head is not transferred until the phase following the phase in which it is valued, denying Value-Transfer simultaneity (1). Consequently, the derivation is doomed if valued uF remains on the phase head. The only way to overcome this fatal flaw and ensure that uF on C/v* is indeed valued as part of Transfer is for C/v*'s uF to be transmitted onto the category that *is* transferred, namely, the complement (T/V).

Since this edge-to-nonedge feature inheritance mechanism is directly entailed by (1) and (2), and thus ultimately by the SMT, it can be said to come for free. As a result, the connection with A/Ā can be reestablished: the now properly motivated feature inheritance mechanism provides an optimal, minimal mechanism for realizing the A/Ā distinction (which, in OP, was not the case; see section 1). Further, unlike the A/Ā rationale discussed in section 1, the argument from (1) + (2) necessarily holds for all phase heads alike: feature inheritance is an inevitable property of phase heads in general since the problem that it solves is inherent to the architecture of the phase system itself.

We now have an answer to the question raised in section 1. C needs T for the same reason that all phase heads need a nonphase complement: namely, to receive the phase head's uF. In so doing, the nonphase feature-receptacle enables valuation of the phase head's uF in a manner consistent with (1). Without nonphase heads, the Agree features on phase heads would be unable to carry out their work. In short, phase heads may well drive all operations, but they cannot do this on their own: successful Agree, and thus dependency formation, is crucially dependent on mediation by a nonphase head, enabling valuation and transfer to be synchronized.

The implications go yet further. Not only do we now have a reason for why proxy, featureless, nonphase heads should exist at all (i.e., why the system does not simply consist entirely and exclusively of phase heads), we can now also explain why T is not a phase head. Given the logic of the argument, the question of why T is not a phase head reduces to the question of why T exists at all. T, qua C's complement, exists precisely because it is *not* a phase head. Since phase heads require featureless, nonphase complements (for the reasons presented above), it follows that C cannot directly select v*.[10] Similarly, if T were a phase (i.e., the locus of uF), then it would be unable to be selected by the C phase or to take the v* phase as its complement and would itself have to select a nonphase to receive its uF.[11]

Generalizing still further, this logic implies that in any sequence of heads, phase heads cannot cooccur consecutively but must be separated by a nonphase head. Moreover, since functional nonphase heads exist *only* to keep phase heads apart in this way (i.e., to provide a receptacle for Agree-type features), it follows that only a single nonphase head should exist between any two phase heads; any additional nonphase heads would fail to be motivated by the SMT (i.e., by the logic of (1) + (2)), and so their existence would not be sanctioned. What emerges, then, is a very narrowly constrained picture of ph(r)ase structure, one in which neither phase heads nor nonphase heads may successively cooccur; rather, phases should consist, maximally and minimally, of one phase head (P) and one nonphase head (N).

(4) a. *P – P – P – P . . .
b. *P – P – N – P . . .
c. *P – N – N – P . . .
d. *N – N – N – N . . .
e. √P – N – P – N . . .

[10] Noam Chomsky (pers. comm.) points out that this possibility is further excluded since the inherited uF cannot probe and value the external argument from the v* position. Spec,v* is not within the probe range (c-command domain) of v*; thus, every (finite, transitive) clause would crash at the higher phase level because of unvalued uF.

[11] It also follows from this that T cannot be a ''phase by inheritance'' (i.e., T does not become a phase head by virtue of inheriting C's uF, a possibility raised by a reviewer), since its complement (v*) would be unable to accept its passed-down features. Any such creation of new phase heads from nonphase heads through inheritance would, furthermore, be genuinely countercyclic (i.e., in the sense of the phase cycle; see footnote 2), adding lower cycles within transferred material.

In sum, we conclude that the system implied by (1) and (2), in conformance with the SMT, is one in which phases are pairs of phase heads and nonphase heads, as in (4e) (hence the familiar core sequence C – T – v* – V).[12] If correct, the reasoning presented above might constrain the possible expansions of the core functional sequence into more richly articulated hierarchies (e.g., Rizzi 1997, Cinque 1999); it would also provide a heuristic for explorations into the structure of DP. I leave these issues for further research.

References

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. by Michael Kenstowicz, 1–52. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2004. Beyond explanatory adequacy. In *The cartography of syntactic structures*. Vol. 3, *Structures and beyond*, ed. by Adriana Belletti, 104–131. Oxford: Oxford University Press. Originally appeared as MIT Occasional Papers in Linguistics 20. Cambridge, Mass.: MIT, Department of Linguistics and Philosophy, MITWPL (2001).

Chomsky, Noam. 2005. On phases. Ms., MIT, Cambridge, Mass.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. Oxford: Oxford University Press.

Collins, Chris. 2002. Eliminating labels. In *Derivation and explanation in the Minimalist Program*, ed. by Samuel David Epstein and T. Daniel Seely, 42–64. Oxford: Blackwell.

Epstein, Samuel David, and T. Daniel Seely. 2002. Rule applications as cycles in a level-free syntax. In *Derivation and explanation in the Minimalist Program*, ed. by Samuel David Epstein and T. Daniel Seely, 65–89. Oxford: Blackwell.

Lasnik, Howard, and Mamoru Saito. 1991. On the subject of infinitives. In *CLS 27*, ed. by Lise M. Dobrin, Lynn Nichols, and Rosa M. Rodriguez, 324–343. Chicago: University of Chicago, Chicago Linguistic Society. Reprinted with minor revisions in Howard Lasnik. 1999. *Minimalist analysis*. Oxford: Blackwell.

Müller, Gereon. 2004. Phrase impenetrability and *wh*-intervention. In *Minimality effects in syntax*, ed. by Arthur Stepanov, Gisbert

[12] Interestingly, this same conclusion is reached on entirely independent grounds in M. Richards 2004, 2006, where I argue that the assumption of phase-nonphase pairs (as lexical subarrays) derives all the properties of the PIC that have to be stipulated in existing definitions (e.g., the timing, edge accessibility, domain, and search space statements of the MI and DbP definitions), as well as deriving why the phase heads are C and v (and not T). That is, phase-nonphase pairs yield the PIC, reinforcing from the opposite angle the conclusion reached in this squib (i.e., that the PIC yields phase-nonphase pairs).

Fanselow, and Ralf Vogel, 289–325. Berlin: Mouton de Gruyter.

Postal, Paul. 1974. *On raising: One rule of English grammar and its theoretical implications*. Cambridge, Mass.: MIT Press.

Richards, Marc. 2004. Object shift and scrambling in North and West Germanic: A case study in symmetrical syntax. Doctoral dissertation, University of Cambridge.

Richards, Marc. 2006. On phases, phase heads, and functional categories. Ms., University of Cambridge.

Richards, Norvin. 1999. Featural cyclicity and the ordering of multiple specifiers. In *Working minimalism*, ed. by Samuel David Epstein and Norbert Hornstein, 127–158. Cambridge, Mass.: MIT Press.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar*, ed. by Liliane Haegeman, 281–337. Dordrecht: Kluwer.

Svenonius, Peter. 2001. Impersonal passives: A phase-based analysis. In *Proceedings of the 18th Scandinavian Conference of Linguistics*, ed. by Arthur Holmer, Jan Olof Svantesson, and Åke Viberg, 109–125. Travaux de l'Institut de Linguistique de Lund 39.2. Lund: Lund University.

Svenonius, Peter. 2004. On the edge. In *Peripheries*, ed. by David Adger, Cécile de Cat, and George Tsoulas, 259–287. Dordrecht: Kluwer.

Rhetorical Questions and *Wh*-Movement

Jon Sprouse
University of Maryland

1 Observations

Rhetorical questions (RQs) have the syntactic form of a question but the semantic value of a declarative (Sadock 1971, Han 2002). In general, the meaning of RQs can be obtained by replacing the *wh*-word[1] with the appropriate negative quantifier (1a) or, in the case of polar questions, by adding negation (1b).

(1) a. After all, what does John know?
''John knows nothing.''
b. After all, does John ever help?
''John doesn't ever help.''

Despite the semantic difference, RQs appear to be identical to interrogative questions (IQs) with respect to syntactic behavior (at least

I am indebted to Norbert Hornstein, Howard Lasnik, and Ivano Caponigro for many useful comments and criticisms on an earlier draft. I am also grateful to all of my informants at the University of Maryland and elsewhere, in particular Hajime Ono, Tomohiro Fujii, Lydia Grebenyova, Julia Belopolsky, and Masaya Yoshida. As always, all mistakes are entirely my own.

[1] Here and throughout, for ease of exposition, I use the term *wh-word* to refer to both single *wh*-words such as *what* and *wh*-phrases such as *which book*.