Viable Syntax: Rethinking Minimalist Architecture

Ken Safir

Hauser *et al.* (2002) suggest that the human language faculty emerged as a genetic innovation in the form of what is called here a 'keystone factor'—a single, simple, formal mental capability that, interacting with the pre-existing faculties of hominid ancestors, caused a cascade of effects resulting in the language faculty in modern humans. They take Merge to be the keystone factor, but instead it is posited here that Merge is the pre-existing mechanism of thought made *viable* by a principle that permits relations interpretable at the interfaces to be mapped onto c-command. The simplified minimalist architecture proposed here respects the keystone factor as closely as possible, but is justified on the basis of linguistic analyses it makes available, including a relativized intervention theory applicable across Case, scope, agreement, selection and linearization, a derivation of the A/A'-distinction from Case theory, and predictions such as why *in situ wh*-interpretation is island-insensitive, but susceptible to intervention effects.

Keywords: A'-movement; Case theory; c-command; evolution; intervention; Merge; minimalism

1. Introduction

The goal of the Minimalist Program (MP) has been to reduce syntactic operations down to the simplest possible mechanism or mechanisms that are consistent with the full complexity and variation that is manifested in natural language. Insofar as this goal is achievable, it appears to open up new opportunities for addressing why it might be that the human language faculty (HLF) as we know it to function now should have emerged with relative suddenness in the evolutionary history

I would like to thank Jonathan Bobaljik, Martin Everaert, Kyle Johnson, Terje Lohndahl, Jim McCloskey, Jason Merchant, Eric Reuland, two anonymous referees, and particularly Mark Baker and Noam Chomsky for useful conversations and comments about various parts of the work presented here. I would also like to thank those participating in my Spring, 2008 seminar at Rutgers, where these ideas were presented for the first time and these include Carlo Linares, Paidraigh Houghton, Dafina Gasparin, Heeshin Koak, Nadya Vinokurova, and Hyunjoo Kim. Thanks also to audiences at the University of Iceland, the University of Tromsoe, the University of Amsterdam, the University of Groningen, and the University of Utrecht where earlier versions of these ideas were also presented. This research was partially supported by NSF BCS-0523102, NSF BCS-0919086, and NWO-040.11.065.

ISSN 1450–3417

of humans, as some propose. The strongest hypothesis consistent with the MP is that there is only one purely syntactic mechanism that interacts with the rest of human cognition to produce HLF, which suggests that the full complexity of grammar may have arisen by virtue of a minimal change in human cognition that had cascading and transformative effects.

From this perspective, the formal goals of the MP dovetail nicely with what may be the necessary ingredient for a satisfying explanation of the sudden emergence of HLF. Put another way, one could treat the solution to the 'sudden emergence problem' as an additional boundary condition on minimalist theorizing: Any candidate for 'the simplest mechanism' that emerges from the MP-informed normal operation of linguistic science *must also* have the right properties to serve as the minimal *keystone factor* that could suddenly coordinate disparate linguistic pre-adaptations into a functioning broad HLF. The only proposal so far that is consistent with such a requirement is that of Hauser *et al.* (2002), who suggest that the Merge operation is (what I am calling) the keystone factor that achieved this fundamental reorganization, i.e. the sudden emergence of the broad HLF (see also Chomsky 2007b).

My proposal springs from the same dovetailed concerns that inform the Hauser *et al.* (2002) proposal, but I argue that the keystone factor (KF) is not Merge, which may have been part of cognition of non-*homo sapiens*. Rather, I will argue that the KF is the *Mapping Principle* that makes Merge-generated syntax *viable*, that is, interpretable by the semantic and morpho-phonological interfaces. In other words, the ability to generate recursive embedding may have predated the HLF of modern *homo sapiens*, but that HLF only emerged with the advent of the ability to interpret structures generated by Merge.

I am assuming that many faculties of mind and body, each with their own evolutionary trajectory, turned out to be useful pre-adaptations for HLF in the broad sense, but the pace of incremental changes in these pre-adaptations appears insufficient to explain the cognitive leap to HLF that appears to distinguish modern *homo sapiens* from all predecessors. This is a broad statement which I expect any number of those expert on particular anatomical and cognitive abilities to take issue with, but on the conceptual level, it is perhaps underappreciated what must be assumed if syntactic complexity is taken to be the incremental result of natural selection.

If one accepts that humans are innately prepared to learn natural languages, then it is difficult to treat the ability to learn constructions of grammar as less than general. Otherwise, one would have to argue that the ability to learn specific complex constructions, especially those not found in every language, arose because some were genetically prepared to learn them when they had to, and they passed the ability to learn the specific structure down to their offspring. Those genetically unprepared to learn the specific construction must have perished or dwindled in the population. For example, one would have to argue that exceptional Case-marking or bare infinitives, or headless relatives with matching effects arose individually in evolutionary history and were proliferated because those who had the ability to generate some of these constructions, but not others, produced more surviving offspring, even if the offspring of these successful individuals happened never to be exposed to a language with one of

these constructions.[1] Similarly, one would have to argue that those hominids that could master both Ergative/Absolutive and Nominative/Accusative case systems would have out-reproduced those that could only master one or the other, even in parts of the world that seem devoid of one or the other construction for long periods of history. This is the sort of scenario that must be accepted if HLF, as manifested in the structural configurations linguists call constructions, grew by selected genetic accretion.

Rather, it seems much more plausible that the ability to master syntactic constructions, including many that are not in the language to which one is exposed, is general, in which case humans have the capacity for knowledge about syntactic constructions that they have never been exposed to, or that their ancestors may never have been exposed to, that is, the class of possible constructions must include many that have not been specifically selected for. If we grant this much reasoning, then most of what we experience as syntactic complexity and variety must be a consequence of more general factors, not of individual constructions of grammar added step by step by natural selection. If, indeed the emergence of HLF was sudden, then the strongest assumption is that the KF consists of a *single* change in cognitive capacity that can account for the (sudden) emergence of complexity, including complexity that is not selected for. This essay is an attempt to argue for the strongest assumption.

Returning now to the Hauser *et al.* (2002) version of the strongest assumption, no single device in the history of generative grammar has ever been adequate to achieve what must be expected of the KF, and so it is no surprise that Merge is not up to it. However to see this, it is necessary to be a bit more precise about Merge. Suppose Merge is as simple as possible, that is, it is essentially Chomsky's (2004) set-Merge (see also Seely 2006).

(1) *Merge*

If α and β are labels or the output of Merge, then Merge of α and β yields $\{\alpha,\beta\}$.

As formulated in (1), $\{\alpha,\beta\}$ can be a term in a Merge operation, for example, with γ, to create $\{\gamma, \{\alpha,\beta\}\}$. Nothing prevents a subpart of a tree from being a term in a Merge operation since any node in a tree is either a label or the output of Merge. Chomsky (2004: 110) describes cases where Merge applies to a term that is already part of a tree as 'internal Merge' (*iMerge* henceforth):

> [Narrow syntax] is based on the free operation Merge, [the Strong Minimalist Thesis] entails that Merge of α, β is unconstrained, therefore either *external* or *internal*. Under external Merge, α and β are separate objects; under internal Merge, one is part of the other, and Merge yields the property of 'displacement', which is ubiquitous in language and must be captured in some manner in any theory […]. Accordingly, displacement is not an 'imperfection' of language; its absence would be an imperfection.

[1] This is not a straw man position. See Christiansen & Chater (2008: 499) who make the following assertion: "Specifically, we adopt a Construction Grammar view of language [references omitted—KS], proposing that individual constructions consisting of words or combinations thereof are among the basic units of selection".

When nodes are Merged, neither of the nodes is changed as a result of Merge (the No Tampering Condition), thus the result of iMerge applying to a sub-constituent B of dominating X will be a copy of B in the position where B occurs before Merge. On these assumptions, as Chomsky points out, a system that did not permit iMerge would require a stipulation to prevent it from occurring (an 'imperfection'), since it comes for free with Merge as defined in (1). Let us assume for the time being that *Extension*, introduced in Chomsky (1995), is a condition on Merge.

(2) *Extension*

Always Merge to an undominated node.

I assume that non-terminals formed by Merge or individual morphemes bearing labels may be thought of as terms, and hence nodes, in syntax.

The formulation of Merge in (1) conditioned by Extension in (2) is very elegant, but it is not often what is assumed *in practice* in most current minimalist architectures. For example, the reduction of Move to an instance of Merge, has been obscured because residues of earlier ideas have not been reevaluated. In much current practice, iMerge and eMerge (Merge of α and β when α is not contained within β) are still distinguished by special features and triggers that formerly played the role of making more expensive iMerge possible, and by different outputs of Merge, such as pair-Merge (see Chomsky 2004: 117–122) as opposed to set-Merge (where only the latter is expressed in (1)). In many minimalist accounts, the architecture is enriched beyond Merge with Agree, *uninterpretable features*, *projection*, *numerations*, *percolation*, *Spec-Head feature-checking*, and certain combinations of these which amount to operational triggers distinguishing instances of iMerge from eMerge. If Merge is the KF, then every one of these additional linguistically specific operations or entities represent departures from KF reasoning.

Once the economic distinction between iMerge and eMerge is discarded, as it is in Chomsky (2004), and some common accretions to Merge have been set aside, only Merge and a slight revision of Extension are required to generate syntactic structure. I argue further, however, that the KF must be something other than Merge. As a result, the architecture of derivations will be quite different.

1.1. A New Direction

Suppose that HLF does not consist of a separate component of human cognition, but in a change of the interface relations that occurred between pre-HLF faculties (a possibility suggested in Hauser *et al.* 2002: 1573). If a single factor is to recruit the pre-adaptations that constitute HLF broad, then it must at minimum introduce or presuppose that the pre- adaptive domains become interpenetrable by virtue of a factor that permits a common interface for (at least some) pre-existing human cognitive capacities. Suppose that Merge already existed in human cognition, but was perhaps encapsulated in another cognitive component, or formally available but dormant. Accordingly, I propose that the KF was not Merge, but a change that has the effect of making Merge *viable* by permitting

varieties of cognition to be mapped onto a common medium, the forms generated by Merge.

The essential proposal of this article, roughly put, is that *the KF is the ability to map interface relations onto c-command relations defined on the output of Merge*. Since c-command will be argued to be, on the one hand, independently (empirically) necessary,[2] and on the other, inexpressible unless it applies to the output of Merge, I will proceed to explore the theoretical intuition that Merge is a largely non-viable pre-*homo sapiens* (pre-HS) capacity, and that only the introduction of a c-command-like notion makes complex syntax available for the integration and expression of cognitively significant relations. The goal of this line of analysis, respecting the burden of the KF, is to derive every aspect of syntactic architecture relevant to HLF from the factors listed in (3).

(3) a. Merge;
b. interpretatively significant relations mapped onto c-command (the KF);
c. lexical properties; and
d. a preference for relative proximity that may not be specifically linguistic.[3]

As a working hypothesis, then, any minimalist architecture that does not derive from (3a–d) must be jettisoned in favor of architecture that respects the burden of the KF.

A reason why c-command is an especially good candidate for the factor that produces a cascade of complexity is that it can be thought of as a structural way of defining *potential closeness* relations (by providing a downward vector) between nodes that may be (potentially) indefinitely far away. C-command also serves to define *relative* closeness (e.g., if X c-commands Y and Y c-commands Z then Y is more local to Z than X). Local domains may then be seen as emerging from the interaction of pre-existing conceptual relations that become expressible in viable syntax and that are then susceptible to proximity interventions. In other words, the existence of syntactic locality domains may be an emergent consequence of syntax made viable by the KF: Mapping of interpretively significant relations onto the potential closeness relation.[4]

---

[2] In Appendix A to this article, proposals made in the literature purporting to derive the effects of c-command from Agree are shown to be untenable, unless Agree is so expanded as to reduce to the essence of the Mapping Principle. Even if Agree were sufficient to replace c-command mapping as the KF, Agree is also an accretion to the theory after Merge, and would face the same KF burden that c-command does.

[3] Conditions on 'computational complexity' play a similar role of inducing locality for Chomsky's version of minimalism, where it is also hypothesized that the relevant constraint is not linguistically specific.

[4] Of course, one can ask why c-command should be the relation to define this relationship rather than, say, dominance. See Neeleman & van de Koot (2002) for an attempt to derive c-command from dominance and projection of labeling. Viable syntax architecture does not permit projection of labeling. If Neeleman & van de Koot's theory were disciplined according to the burden of the keystone factor, assuming labeling but not c-command, an alternative approach to some of the argumentation for c-command presented here might be possible, but a very different architecture would emerge, permitting relations in addition to

To grasp the substance of this proposal, as opposed to other imaginable ones, it is important to understand that viability is only partial in HLF. The structures generated by Merge can be quite complex, consisting of a great number of nodes that define subunits, depending on how large a sentence is. Syntacticians differ markedly on what exact structure they would assign to a sentence like (4), but the (node dominating the) word *him* and the (node dominating the) word *he* are not normally regarded to be in a direct relationship that is regulated by syntax, even if *he* and *him* in (4) are taken to pick out the identical object of thought.

(4) A woman who knew *him* thought that *he* was brave.

Yet there is a structural relation between *him* and *he* that one could calculate in terms of the number of nodes that each is dominated by up to the node that dominates them both. If such a relation were viable, semantic generalizations could be stated on such a correspondence. No such relationship is viable in UG, nor are any number of other such relationships that could be defined on structures generated by Merge. It is big news in linguistic theory when new structural relations are posited to be viable. Selection under sisterhood and c-command were among the first of these, to be followed by various locality relations, not all of them still assumed to be part of modern theory, including subjacency, binding domains, maximal projections, government, and phases, to name a few. The theory developed here attempts both to limit viable relations and to permit just a small set of local relations that arise as interventions determined by relative closeness in contexts where potential closeness holds.

Suppose, then, for the sake of argument, that Merge is a pre-HS capability (which may be viable for some other cognitive domain, such as predation planning or kinship calculation), but the ability to map relations in other cognitive domains to relations between nodes is not. As suggested above, a structural relation is *viable* if and only if an interface relation is permitted to exploit it. Suppose further that 'sister' and 'dominate' are not viable relations, that is, 'X dominates A' and 'X dominates B' are not interpretively visible on their own, except for the calculation of c-command, as stated in (5).[5]

(5) *C-Command*

B is *c-commanded* by A if B is dominated by a sister of A.

If the structural domination relation is not viable, then no interpretive relation can map onto it, and if so, then inheritance of a label, if posited as a contentful relation spreading from A or B to dominating X, would be invisible to interpre-

---

those allowed by unique appeal to c-command. On grounds of parsimony, a c-command primitive may be preferable to one based on labeling, but this depends on how the primitives are formulated.

[5] Seely (2006: 197) assumes that a node without a label would be uninterpretable at LF and therefore cannot exist, that is, non-terminal nodes cannot exist, hence a set-theoretic notion of containment must be preferred to dominance. However it is assumed here that LF only evaluates interpretable properties for well-formedness, not uninterpretable ones. See fn. 10.

tation. Now suppose that Merge creates [Y [$_X$ A, B]] and we call the whole structure Z, as in (6).[6]

(6) [$_Z$ Y [$_X$ A, B]]

Neither A nor B of the composite [$_X$ A, B] would be interpretable in relation to Y, unless c-command is a viable relation. The structure in (6) is generable, but pointlessly so, because it cannot be used to shape, for example, semantic interpretation. Only the availability of the *Mapping Principle* in (7) permits interpretation to exploit relations between Y and sub-constituents A and B embedded in X, where X is the sister to Y as illustrated above in (6).

(7) *The Mapping Principle*

If x and y are in relation R, then R is mapped onto syntax if and only if either x c-commands y or y c-commands x.

Thus Merge may have been in the cognitive arsenal of pre-HS hominids—and if so, they had the potential to generate complex trees, possibly useful for other purposes, but could not use them for semantic or phonological interpretation until the advent of The Mapping Principle.

C-command is a powerful template for interpretive relations. For example, if A or B is a non-terminal in (6), then Y also c-commands all daughters of A and B, so the recursive structures always generable by Merge are now viable and may be exploited by interpretive relations. C-command will permit a head to be in a selection relation to something inside its sister, it will permit antecedents to antecede anaphors and a copy left by internal merge to be related to a copy that is c-commanded by it. Discourses may then be thought of as concatenations of the final outputs of Merge operations. If A and B are concatenated nodes in a discourse, then neither A nor any node inside it c-commands B nor vice versa. If the Mapping Principle as stated is the KF, then antecedent relations that are not in a c-command relation are uninterpretable by any system or component that has access only to syntactic structure.

The burden of the KF is now carried by the Mapping Principle as it makes structures created by Merge interpretable and creates a common vehicle for the integration of relations expressed by interface components. Thus all of the complexity of grammar should now be expected to follow from the formulation of Merge, interpretive relations mapped onto c-command (as the means of expressing proximity of interpretive relations), *and* the general assumption that interpretive relations can be sensitive to relative proximity, that is, to interventions.

---

[6] Epstein *et al.* (1998) point out that if Extension ensures that each new constituent is a sister to the whole structure, then c-command relations could simply be recorded as moments in the history of a derivation (a point further developed in Seely 2006). This assumption is compatible with viable syntax architecture and I will be exploited here with respect to 'Firstness', but whether c-command is derivationally defined or not does not explain why c-command is a viable relation.

(8) *Intervention*

B *intervenes* between A and C if A c-commands B and C, and B c-commands C.

(9) *Intervener*

B is an *intervener* between A and C if
i. B intervenes between A and C, and
ii. for some relation *R*, *R* cannot relate A and C when B is present.

This account of local relations as bounded by intervention is essentially a restatement of relativized minimality defined on c-command, as in Rizzi (1990) (not just relative proximity of any sort, as in Koster 1976). On this account, however, *locality only emerges when the Mapping Principle makes viable the relations on which interventions can be defined*. Put another way, potential proximity in the form of c-command is logically prior to locality, just as Merge is logically prior to c-command.[7] Thus interventions do not block Merge operations, but they may block one interpretive relationship or another from being established on the output of Merge operations.

Consider, for example, how intervention works if *c-selection* is relation *R*. Since labels on nodes are only introduced into a derivation on the lexical items (heads) that bear them (since projection relations are not viable), the complement relation in viable syntax architecture (VSA) can only be formulated as a relation between a selecting head and a head bearing a label that it selects. A head H selects a head Y if H must seek Y, H c-commands Y, and there are no intervening selectors between H and Y. So if V selects D (e.g., the verb *kill* selects for a nominal), then V must c-command D and no other selector can intervene between V and D. Thus, *depend* could not select for *the* in *Don depended on the boat* because *on* is a selecting head that intervenes between *depend* and *the*. The reasoning here is in the spirit of Collins (2002), though I do not follow his proposal in detail. A head cannot select for a non-terminal Y as its complement (e.g., a verb cannot select for a PP) because projection is not viable, and, more specifically to VSA, H cannot select a sister to H because sisterhood is not c-command (i.e. there are no viable head-head sister relations). Thus a simple sentence like *John must leave* would require that if *must* selects *leave*, then *leave* is not a sister to *must*, but is embedded in a branching sister to *must*, as it is in all theories that assume 'little *v*'.[8]

---

[7] Frank & Shanker (2001) propose that c-command, as opposed to dominance, should be treated as a primitive, which could conceivably be regarded as an impetus to remove the logically prior status of Merge for c-command (i.e. if trees are created by the formation of c-command relations). As a justification for this move, they point out, as I have here, that many sorts of relations stated in terms of dominance and precedence might be treated as significant, but only c-command seems to play a significant role in syntax. Their axiomatization of syntactic tree structure in terms of c-command starts from the assumption that sisters are in a mutually c-commanding relation, which is not assumed here. It is unclear what the consequences for their account would be if sisters are not in a c-command relation (it is also assumed in the text that domination is not reflexive, or sisters would c-command each other, given (5)).

[8] Since c-command is always asymmetric in this system, no viable relation results when

Thus VSA replaces the boundary-forming function of projection with boundaries created by intervention. This approach to intervention begs what might be called the Natural Intervention Hypothesis, as stated in (10).

(10) *Natural Intervention Hypothesis*

a. Strong Form: If B intervenes between A and C for relation *R*, then B must be eligible to participate in *R* with A and C.

b. Weak Form: If B intervenes between A and C for relation *R*, then B must have properties in common with A and C.

Although they have not always been distinguished, both versions of the Natural Intervention Hypothesis have been explored in the literature on interventions of various kinds (e.g., Koster 1976) and even much earlier in phonology. Although (10a) makes the strongest predictions about possible interveners, it has well-known problems that will not be solved here. My proposals are consistent with at least (10b). Moreover, as I develop VSA, interventions will more typically arise from intervening heads rather than intervening 'specifiers', even though the apparent interpretive relation may be one between non-terminal nodes.[9]

The foregoing introduction is meant to introduce the following propositions:

(11) a. The sudden emergence problem, with its subparts (single factor change, unselected complexity, recruitment of existing capacities) requires positing a KF.

b. The KF is the advent of the Mapping Principle.

c. Relations mapped onto structure are sensitive to a pre-linguistic preference for relatively closer relations (once c-command provides the vector).

d. All the architecture of syntax and the restrictions that shape constructions of grammar derive from the interaction of Merge, interpretive relations stated uniquely on c-command, relative proximity, and the distinct properties of lexical items.

(11a–d) are empirical hypotheses, but (11a, c) are assumed here without argument, as they are in Hauser *et al.* (2002). My main goal is to show that an architecture for syntax consistent with (11b) is defensible, namely, the one in (11d). In defending (11d), appeal will be made to pre-existing cognitive capacities from which mappable relations arise, and those attributions are empirical hypotheses that will not be explored. An honest appraisal of other syntactic theories would reveal similarly rich assumptions about human capacities that are not defended,

---

terminals are merged. Only merger where one of the terms is non-terminal results in a c-command relation. In theories with mutual c-command (sisters), there would be unintended intervention effects (as Seely 2006: 202 points out for the theory of selection in Collins 2002).

[9] The viable syntax instantiation of relativized minimality is more restrictive than the original formulation, insofar as no minimality restriction can be placed on intervening maximal projections, for example, node-naming locality statements like subjacency list nodes that are in the class of interveners. See section 6 on cyclicity.

but the point here is not that my assumptions should escape scrutiny. Rather, my assumptions should be put into perspective with respect to the task at hand, which is defending (11d).

1.2. The Road Ahead

The program throughout is to show that VSA, as a particular instantiation of the MP, can be constructed and that it comes close to respecting (11d) and which dispenses with most of the ancillary mechanisms surrounding Merge, or else reduces those mechanisms to interpretive relations permitted by the Mapping Principle. Where the program fails to truly reduce a problem, as it occasionally does, I will try to show that VSA fares no worse than other existing minimalist accounts, but it is to be acknowledged that in the larger picture this is not good enough to meet the burden of the KF, and is just the best that I was able to do. In spite of this caveat, I will argue that VSA offers certain advantages over competing analyses based on assumptions common to most minimalist accounts. These include a relativized intervention theory applicable across Case, scope, agreement, selection and linearization, a revision of Extension consistent with late attachment and head-to-head movement, derivation of the A/A'-distinction from Case theory, and a derivation of why *in situ* wh-interpretation is island insensitive, yet susceptible to intervention effects. Whether or not one accepts the evolutionary reasoning that has inspired my proposal, my central contention is that VSA deserves to stand on its own merits as a parsimonious and insightful account of the relation of syntax to interpretation at the interfaces.

The remainder of this essay fleshes out a version of VSA that respects the KF as closely as possible. Section 2 clears the field of accretions to Merge in minimalist syntax that cannot be countenanced in VSA. Section 3 develops some major design features of VSA, further specifying of the principles for assigning interpretation to structure and the interaction between structure and interpretation in the course of a derivation, while keeping in mind new devices should not complicate the KF, or ascribe implausible properties to other components. Subsequent sections extend VSA design to Case-marking (section 4), to the derivation of the A/A'-distinction from Case Theory and the interpretive role of criterial positions (section 5), to cyclic linearization and the locality of extraction (section 6), and to contrasts between terminal and non-terminal node realization that derives the role of pied-piping on intervention effects (7). Section 8 concludes by briefly summarizing the main arguments for VSA.

2. What We Must Do Without

The Hauser *et al.* (2002) proposal that Merge is the KF does not succeed because it relies, in practice, on many additional devices to achieve a descriptively adequate account of phenomena considered central to natural language grammar. The list that follows includes a range of devices, not all of them found in every minimalist proposal, but many of them found in most minimalist proposals, that do not stand up to the logic of the KF if that factor is just Merge. Residues of

earlier accounts of iMerge as distinct from Merge must also be swept away, in particular, the notion that only iMerge must be operationally triggered (e.g., by feature-checking). In presenting a list of what we must do without, I clarify what VSA must achieve.

Consider first projection. A labeling relation between a node and a head it dominates is something that can be added to the Merge operation, but it is not a consequence of simply combining two terms. Projecting a label is not a viable relation in VSA since it is not a c-command relation. All that VSA can countenance is a label on a terminal node (head) and nodes that, because they are not terminal, have no label.

If 'the feature composition of a node" has no meaning for a non-terminal in a theory without projection, then there is no way to express Spec-Head checking, which is another accretion on Merge. There is no label with features on a non-terminal node for a head to check. Since heads do not c-command their 'Spec', no viable relation holds between head to Spec either. Even if a theory with Agree does not require Spec-Head relations (e.g., Chomsky 2004), VSA goes further by rendering Spec-Head relations ineffable.[10]

Removing Spec-Head checking removes a mechanical necessity for percolation, which has been used to account for how a node dominating K can be attracted because K is in it. Percolation permits a property of a maximal projection A (containing K) to inherit feature(s) of K, and then, potentially, for the feature of A to be inherited by B, a maximal projection that contains A, and so on. Percolation is not viable in any case, since it involves mapping interpretive properties onto dominance relations and not c-command relations.[11]

---

[10] It is usually assumed in minimalist accounts that ill-formedness in phonology or semantics might result because *an interpretable feature has been left uninterpreted*. Although I don't express this claim in terms of features, the idea that formal properties of interpretation (morpho-phonological or semantic) condition possible outputs is also a crucial part of VSA.

However, the notion 'uninterpretable feature' goes beyond these considerations and requires a retreat from the logic of the KF. It is not at all clear that any phonological feature, whether it is checked or not, should ever be visible to semantic interpretation by its very nature. Similarly, no semantic feature should necessarily be visible to phonology if unchecked. Being visible in the 'wrong' component is not a necessary assumption, and if not, then phonological features do not supply any information to the semantic component, *including information about ill-formedness*, nor can semantic features contribute to the ill-formedness of phonological representations. In other words, segregating features by component restricts their descriptive power (and thereby derives the proposal of Pesetsky & Torrego 2001, that every proposed feature should be interpretable in *some* component). Structural case features play a role in phonology, for example, but are not relevant to semantics, whereas inherent Case features, which have both morphological and semantic value, may be visible in both components. Why Case features should have an origin in some pre-linguistic component remains mysterious (see Bobaljik & Wurmbrand 2008 for a summary of the issues), and it is in the latter sense that they represent a challenge to KF reasoning.

[11] Percolation has been stipulated to apply whenever a non-terminal of arbitrary size is moved (e.g., containing *wh*-REL, as in *Bill, Al's pictures of whom, we will soon see*), but then percolation is just a notion that describes what is empirically possible—it is not an explanation (as pointed out by Heck 2004, 2007). Watanabe (2006) actually introduces a pied-piper feature that can be probed for by Agree in order to trigger movement. VSA countenances neither the role of Agree in Merge or the feature on the targeted maximal projection. Restrictions on pied-piping are poorly understood, but one limitation may be that the *wh*-REL must covertly move to a scopal position, and cannot do so if conditions on movement are violated (see e.g.

The elimination of feature-checking and percolation from the theory on the assumption that Merge is free as long as the output is interpretable at the interfaces obviates any appeal for operational triggers embedded in the application of Merge. Now consider the function of Agree, which has been appealed to play two roles in recent years (e.g., Chomsky 2000, 2001). One is to value features of the goal (and is still widely employed for this purpose), and the other, in some versions of minimalism, is to make the goal susceptible to iMerge (to 'activate' it, in some theories, for example, Bošković 2007) in combination with an EPP feature, that is, to trigger movement (e.g., in Pesetsky & Torrego 2001). On this point Chomsky has frequently been explicit, even as late as Chomsky (2004), where he first proposed that iMerge and eMerge are the same operation.

> If there is no Spec-Head relation, then the EPP-feature OCC cannot be satisfied by Merge alone. It follows that internal Merge requires Agree. Therefore Move = Agree+pied-piping+Merge. (Chomsky 2004: 114)

Agree, however, cannot relate a head and a maximal projection in the VSA version because there are no maximal projections in VSA. Thus, Agree is a viable relation in VSA just in case it is mapped onto a c-command relationship between terminal nodes, and I will appeal to just such a relation more than once, but it cannot be used in VSA to activate non-terminals for movement. If Chomsky (2004) is right that Merge is not constrained by economy, and if Agree and pied-piping are not added to Merge, then iMerge and eMerge should be equally possible at any point in a derivation. The result may be uninterpretable, however, and so the descriptive and explanatory burden must be borne by the architecture of syntax-sensitive interpretation.

Most minimalist accounts posit that, in addition to Merge, every derivation begins with the selection of a numeration, a fixed set of selections from the lexicon to be used in the course of a derivation and to be depleted until the set is empty. Appealing to the numeration, one could still maintain that iMerge is less economical that eMerge just in case an operation that reduces the numeration is more economical than one that does not (see e.g. Lasnik & Uriagereka with Boeckx 2005: 166 and Safir 2008: 331). However, the only reason to stipulate a numeration as part of a derivation is to insure that the most efficient derivation is computed. The actual selection of numerations is unprincipled. If interpretation, rather than economy, is the only arbiter of the well-formedness of a derivation, then pre-derivational numerations are superfluous.[12]

Sauerland & Heck 2003 for intervention effects within pied-piped constituents). This suggests that displacement is not restricted by conditions on Merge, but the result must be interpretable (a hypothesis that is fleshed out for VSA in sections 5 and 7).

[12] In rejecting economy calculations on forms that are independent of their semantic import, I am also rejecting approaches to syntax such as that of Optimality Theory, which begin from the assumption that GEN takes *a given input of forms* and generates all the representations that are competitors for the optimal derivation or representation. I do assume that the output of Merge submitted to the interpretive component contains a set of lexical items which are mapped onto the relevant prominence relations. The 'post-derivational numeration' and the structure built on it could be used to compute optimal form-interpretation matches, as in Safir (2004a), where substitutions into the post-derivational numeration can compete with the actual output (in the spirit of Reinhart's 2006 'reference set computation').

The elimination of projection, percolation, operational triggers distinguishing eMerge and iMerge, Agree as part of Merge, Spec-Head feature-checking, and numerations clears away a lot of what compromises Merge as a model for the KF, but leaves Merge insufficiently expressive. We must ask now whether the Mapping Principle meets the empirical burdens these accretions to Merge were supposed to address, while still respecting the burden and the logic of the KF.

3. Major Design Features of Viable Syntax Architecture

There are at least two properties of syntactic structure that do not appear to be required by a compositional semantics defined to interpret recursive structures. These include geometrically-defined locality restrictions, some of which appear absolute (nothing outside of the local domain can be accessed for interpretation), and in other cases, the distance between two potentially related nodes A and B can be unbounded, but any node of the wrong sort intervening between them will be enough to prevent A from being related to B. Both of these phenomena will be treated as forms of intervention effects in VSA.

Moreover, every theory of syntax must be able to resolve conflicts that arise between the mapping of lexical argument structure onto syntax and the mapping of scopal relations onto syntax. A typical conflict of this sort arises when argument A of predicate P in the argument structure of P(A, B) must be more prominent than argument B, but B must have scope over A for some quantificational relation. The usual syntactic approach to this conflict is to assume that argument structure projects onto prominence relations in syntax and then movement or its analog reorders arguments or their parts to achieve scopal prominence. The resolution of these conflicts may differ cross-linguistically. For example, the c-command relations that must exist to support a given word order may not appear consonant with those that must express scopal relations, as in the case of *in situ* phenomena. Any syntactic theory that aspires to adequacy must be able to both characterize and, hopefully, predict, the class of possible prominence conflicts and the class of possible resolutions. The burden of VSA is to insure that all such prominence relations can be best expressed as interpretative relations mapped onto c-command relations.

I am assuming that the prominence relations just mentioned are probably *not* part of what is introduced by the KF, but rather that the expression of them in structure is what is new. Whatever determines that agents are more prominent than patients for a given lexical argument structure, or for such argument structures generally, is not assumed here to be part of what Merge or c-command contributes. The assumption that lexical argument prominence relations for particular lexical predicates predates the advent of c-command is a strong claim, not defended here, that could easily be false. If prominence relations *only emerged* with HLF, then there is a greater burden to show that notional prominence could only emerge as a result of the KF. For example, it is not obvious that scopal

---

However, no form–interpretation competition can be part of GEN unless a great deal of semantics is built into it. See Safir (2004a: 234-237).

relations, especially relative scopal relations, are effable without sufficient syntax, since they are not prominence relations directly associated with lexical items. In other words, it is possible that relative scope could be an emergent property in VSA, but argument prominence for verbs or perhaps some asyntactic notion of domain associated with particular quantifiers may predate the KF.

3.1. The Place of Interpretation

From the VSA perspective, Merge must provide the structures from which prominence relations can be read by the syntax sensitive interpretive component. VSA is thus in the tradition of theories that generate syntax freely and filter the output, where in this theory, the phonological and interpretive components do all of the filtering. Insofar as setting of argument prominence relations will be relative to semantic classes of heads, argument prominence will be distinct from scopal prominence, but apart from the classes of heads involved, the formal prominence-setting mechanisms will be the same for both. If argument prominence is thought of, in GB-traditional terms, as 'A-relations' and scopal prominence as 'A'-relations', then the theory of A/A'-distinction will reduce to a difference between classes of heads, but will not reside in the way prominence domains are set or trees are generated, nor will there be distinct versions of movement or adjunction that require elaboration of Merge (such as pair-Merge, as in Chomsky 2004, or probe-contingent iMerge, as in Chomsky 2007a, cited above).

For example, if we assume that for every class of quantifier, there is a scopal position (or one of a set of scopal positions) to which it must move to receive an interpretation, as in the theory of Beghelli & Stowell (1997), then a quantifier without a compatible scope domain will fail to have a proper interpretation (see section 7.1).[13] I also assume that for every predicate P with more than one argument, there is a prominence ordering between the arguments of P (e.g., *killer* and *killee* for *kill*) interpreted from syntax as a c-command relationship between those arguments. If a verb cannot recover prominence relations from the tree that match the prominence relations in its lexical entry, the verb cannot have a proper interpretation. In this way, semantic conditions on scope and argument structure filter out uninterpretable trees from amongst those that can be generated by Merge applying freely.

Another design feature that is adapted here from existing accounts is that there is only one syntactic movement component: quantifier movements must take place amongst the movements that result in overt reordering as part of the

[13] The approach pursued here shares much with that of Bošković (2007) insofar as an insufficiency on a constituent is frequently what drives that constituent to move, rather than a trigger that may not have yet entered the derivation, particularly in cases of successive cyclic movement. Bošković's mechanism is different, in that he explicitly assumes that the insufficiency is an unvalued feature which Last Resort permits to move in the interest of convergence, but Last Resort is a form of look-ahead, which Bošković is trying to avoid. Still, this use of Last Resort permits Bošković to dispense with appeal to intermediate triggers for successive cyclic movement. In VSA the movement is optional, but a moved constituent must find the right domain by the time it reaches interpretation, and must avoid local interpretation by escaping to the intermediate zone (see section 6), so intermediate triggers, and triggers in general, are also unnecessary (and unstateable) in VSA.

same process. The phonological component will determine which copies generated by movement are to be pronounced. In such theories, notably that of Bobaljik (2002) and references cited there, most, if not all of the movements that are observed in any language are found in all languages, but in those languages that appear not to have a movement found in some other language, the phonology masks the presence of movement by pronouncing the lowest copy. From this perspective, wh-*in-situ* phenomena are just instances where the lowest copy has been pronounced instead of the highest one. I will not adopt these assumptions whole, as my discussion of pied-piping and intervention will show, but my design follows the same leading idea.

I also draw some of VSA design from a theory that has been regarded, in part, as a competitor to the approach just described. I assume that prominence interpretation proceeds in tandem with the formation of structure to *set prominence* and *assign domains* and that it does so in ways that later operations in a derivation cannot revise. It is convenient to give this general design hypothesis a name, as in (12).

(12) *Derivational Drag Hypothesis*

Certain relations, once established at a given point in a derivation, are never revised.

In an architecture with this design, every operation that results in, or contributes to, prominence-fixing for interpretation or for phonology limits the possible interpretations of structure introduced subsequently. In this respect, my approach adapts ways of thinking introduced by Chomsky (2000) as 'cyclic Spell-Out' and adapted in Fox & Pesetsky (2005), where precedence relations introduced in the course of a derivation constrain possible linearizations that could result from further derivational steps. My approach will also use cyclic Spell-Out to (indirectly) freeze overt displacement across more than one cycle, but also to distinguish the cyclic Spell-Out of terminals vs. non-terminals. Derivational drag effects are also posited to arise from the assignment of unerasable Case and unerasable scope. Essentially, once a node is assigned a non-Nominative structural Case, no occurrence of that node can receive any other Case assignment without inducing Case conflict that causes the derivation to crash. Only a node assigned Nominative permits subsequent copies of that node to be reassigned some other Case. If Case is assigned cyclically, only Nominative nominals can move. This system is developed in section 4 and illustrated with several derivations.

3.2. Setting and Assigning a Domain

Derivational Drag insures the preservation of interpretive relations once they are introduced, but another key notion that resonates throughout the approach developed here is that many interpretive relations are established at the first point in the derivation where that interpretive relation can be interpreted. This means that there will be a special importance to points in a derivation where a new c-command relation is established that permits mapping of an interpretive

relation onto it. Principles for setting and assigning domains are presented in (13).

(13) a. *Setting a Domain*

If H is terminal, every node H c-commands when it is (first) merged is in the domain H_D.

b. *Assigning a Domain*

The domain H_D of head H is assigned to the first X merged such that X c-commands H_D and H.[14]

On this analysis, domain assignment is thus a two-step process for phrasal nodes. First a head H *sets* the domain it c-commands (H_D) at the first point where H is merged. The first 'X' that c-commands H_D and H has the domain of that head *assigned* to it. Scope is assigned, for example, when a scope-marking head H sets a scopal domain H_D, and then a QP is merged to [H H_D] (where H_D is shorthand for the set of nodes c-commanded by H, that is, H_D is not the name of the node that is a sister to H).

Although there may be reasons to doubt the parallel later, an instance of Merge that induces domain assignment may be thought of as the derivational equivalent of the specifier relation. However in this account, the uniqueness of 'specifiers' is a derived consequence of the role of the *first* c-commander of a domain, which presents the first opportunity in a derivation to map an interpretive relation, namely, domain assignment. Firstness could be factored out

[14] Note that (13b) does not stipulate that the setting of a domain is always an asymmetric relation, although it is possible that it always is. Domain-setting is asymmetric because a terminal and a non-terminal are merged, but as it stands, (13b) allows that merger of a non-terminal with another non-terminal that could result in both constituents being assigned a domain by first c-commanding a domain in its counterpart, e.g., where merger of [X, X_D] and [Y, Y_D] would allow that Y_D is assigned X_D and X_D is simultaneously assigned Y_D. To (13b) could be added the following condition: No more than one domain is assigned on any given instance of Merge.

As Michal Starke (p.c.) points out, there is otherwise too much symmetry for certain instances of Merge, yielding uncertainty about what sort of constituent results from the merger. I am not convinced that symmetry in such cases could not be ruled out on independent interpretive grounds, but if not, then this stipulation is necessary and recovers part of the information formerly expressed in terms of projection relations when non-terminals merge. When non-terminals merge and no domain is assigned (e.g., neither non-terminal immediately dominates a domain set by a head), it is not clear that any asymmetry is necessary or relevant in VSA. Thus there is no 'adjunct-argument' distinction in syntax at all, but merely interpretive differences that arise on account of Firstness. Adjunctions can be described as Merge operations that do not result in domain setting or domain assignment. Compare Hornstein & Nunes (2008), where adjunction is formally distinguished in the syntax from other structure-building, based on a difference in how structures come to be labeled. Since Hornstein & Nunes treat Merge as two operations, one of which (concatenation) can apply without the other (labeling), it is not clear how their approach could be squared with KF reasoning.

Notice also that I have included c-command of the head H as well as H_D, which is to distinguish cases where an operator Merges to H_D by 'tucking in' such that H c-commands the operator. It is not obvious that this possibility should be excluded on empirical grounds, but I exclude it here to simplify presentation of the theory.

to serve as a particular statement of Derivational Drag.[15]

(14) *Firstness*

Interpretive relations in a derivation are uniquely assigned at the first point in a derivation where they are viable.

The mechanisms for domain setting and assignment meet the desideratum of being structurally identical for argument prominence relations and non-argument prominence relations. For example, the definition of 'external argument' can now be thought of as the assignment of v^*_D to X, which arises when X is the first c-commander of v^* and v^*_D (where v^* is 'little v'). Thus when *John* in the sentence *John hit Bill* merges to [v^* [*hit Bill*]], *John* is assigned v^*_D, and as a result *John* is an argument more prominent than anything v^*_D. The relation between v^* and the verb it selects will not be explored here, but the theory of domain assignment is consistent with the view that V raises to v^*, adjoining to v^* at the point in the derivation where v^* and v^*_D are immediately dominated by the undominated node. Subsequent merger of the external argument to the undominated node to [[V v^*] v^*_D] will still be the first constituent to c-command both v^* and v^*_D, and thereby will meet the definition of domain assignment (for assumptions about head-to-head adjunction, see sections 5.1 and 7.2).

The nature of H_D is determined by the properties (e.g., features) of H. Setting of a domain may or may not involve sensitivity to a label in that domain. As mentioned earlier, the theory of c-selection requires that the selecting head find a particular label within its domain if selection is to be successful. However, heads that normally agree with a label in their domain are sometimes permitted a default form if their domain is empty of such a label, a situation without parallel in complement selection. Heads setting scopal domains do not appear to have any relation with a node in their domain, but the burden of successful interpretation is then on whether or not a scopal element, which requires a domain, has the right domain assigned to it. In addition, it is argued in section 4 that some domain-setting heads also assign prominence ordering to multiple labels in their domain where there is no intervention between them. Domain setting in (13a) is thus the most general statement of domain setting, in that it

[15] Firstness, along asymmetric c-command in (5) and the assumptions about domain assignment in (13), will derive much of what Collins (2002) introduces his Locus Principle, insofar as assignments occur as soon as they can, but without the problem of relativized intervention pointed out by Seely (2006). Interventions stated on the heads that introduce domains (enabling assignments) are relativized across different sorts of relations, such that c-selecting heads do not intervene between a probe and goal for agreement relations.

Starke (2004) (see also Jayaseelan 2008) has suggested that specifiers are unnecessary insofar as they are so often in complementary distribution with the heads they correspond to, hence only one of the two is really necessary to satisfy a fixed sequence or template of functional projections. However, heads and the elements that they set domains for in VSA are not always in complementary distribution. Moreover, insofar as Starke assumes that the functional sequence is fixed as a template, rather than as recursive selection as permitted by VSA, a new linguistically specific device, the functional sequence, is added to the theory along with Merge and whatever else is needed to generate the full range of structures. This appears to be a departure from KF reasoning and in spite of some interesting issues that are raised, I will not explore it further.

does not require selection for a label and, moreover, it does not specify what sorts of interveners limit the domain, apart from the hope that (at least the weak form of) the natural intervention hypothesis (10) will provide a leading strategy. It remains to be seen whether the properties assigned to heads in the varieties of domain-setting relations provide sufficient descriptive power for both lexical and crosslinguistic variation, but this power appears to be ample.

3.3. VSA and the KF

Given this design, most instances of Merge in a derivation, whether they are internal or external, add unalterable information to interpretation, either by introducing a head that sets a domain, or one that orders the labels in its domain for prominence, or else by merging a term for which a domain is assigned. Interpretation proceeding in tandem with tree construction thus initiates considerable derivational drag. The reach of a head setting a domain is limited by the intervention of a more locally c-commanding head, though I have touched on this only lightly so far. That is the essence of VSA design in a nutshell.

How well do these assumptions so far respect the KF? Insofar as domain setting and domain assignment respect c-command, I have not compromised the Mapping Principle for these notions, which are central to all that follows. I am, however, committed to some pre-linguistic relation between certain kinds of potentially discrete notional content and terms that cover those contents. For example, if there is a pre-linguistic notion of an eating event and as a condition of it that something must be consumed, that is, something must undergo that action, then there is perhaps a pre-linguistic precursor to Agent-Patient relations. Such relations can become both more intricate and more generalizable when the Mapping Principle provides for viable expression of them, for example, in c-command structures that map argument prominence in a consistent way.[16]

4. Case Prominence

One of the practical functions of Agree that has been appealed to within

[16] Although I assume that features enter structure by virtue of the contributions of non-syntactic components, I do not explore here how lexical items bearing labels and features (phonological, semantic, morphological) come to have the labels and features they bear. One of the most dramatic differences between primates that have been taught sign vocabularies and human lexicons is that human lexicons are enormous by comparison and can be casually and productively extended. This difference should, in principle, also result from the KF, at least as a research strategy. There is reason to believe that structure internal of words is interpretively complex. The question is then whether or not it is the advent of the Mapping Principle that makes the lexicon fully viable. Whether one has the view that morphological structure *is* syntax, or that there is an independent, pre-keystone lexical component transformed by access to structure provided by the KF (e.g., being able to map meaningful relations onto word internal c-command found in rich word-internal representations), a theory consistent with VSA is possible. For a variety of recent theoretical approaches exploring the textured internal syntax of words, see Borer (2005a, 2005b), Ackema & Neeleman (2004), Ramchand (2009), amongst others. The exploration of this interesting question is beyond the scope of this essay.

minimalist architectures is to identify constituents that can be moved to A-positions (as in Chomsky 2007a: 25). In a theory without operational triggers that distinguish iMerge from eMerge, an alternative mechanism must indirectly determine which constituents participate in successful A-movement and which constituents do not. In this section I exploit recent developments in the theory of Case assignment and agreement stemming from work by Marantz (1991), Bittner & Hale (1996), and, especially, Bobaljik (2008), who suggest Case assignment arises from a mapping of Case arrays onto prominence relations (as determined by c-command). This independently motivated theory will not only permit us to dispense with triggers for A-movement, but it also turns out to derive syntactic distinction between A-movement and A'-movement, a consequence I postpone until section 5.

In a prominence-mapping (P-M) theory of Case assignment, Case-assigning heads can assign a single Case or a Case array—practically speaking, two Cases, where one of the Cases is more prominent than the other one. The most prominent Case is mapped onto the most prominent argument in the domain of the Case assigner ($CASE_D$), and the less prominent Case is assigned to the next most prominent argument in $CASE_D$. If there is only one nominal, K, in $CASE_D$, then the most prominent Case in the array of the Case assigner is assigned to K. If there is no nominal in $CASE_D$ to assign Case to, none is assigned.[17] A Case sometimes described as 'unmarked' is assigned to the most prominent nominal in $CASE_D$ and a marked Case is assigned to the nominal(s) of lesser prominence in $CASE_D$. No marked Case can be assigned Case again or a conflict I will call 'Case Clash' arises. This means that a marked Case cannot move into the $CASE_D$ of another Case assigner. A nominal with an unmarked Case, however, is still eligible to be assigned another Case. If the unmarked Case is identified as Nominative, for example, then only Nominatives can move into a higher Case domain and be reassigned Case. This is a form of derivational drag, where marked Case assignment has the effect of freezing nominals in place (not literally—movement is free, but Case Clash is a failure at the morphological interface). Only the most prominent nominals in $CASE_D$, when these are Nominative, can be promoted into the $CASE_D$ of a higher Case assigner.

Part of the independent appeal of a P-M Case theory is that it derives Burzio's Generalization, which is the observation that verbs that do not assign an external argument do not assign (structural) Accusative Case. In a P-M theory, this is simply due to the existence of only one eligible argument on the prominence scale.

P-M Case assignment also provides the basis for an elegant theory of finite verb agreement based on the idea that the target of verb agreement is the nearest nominal with an unmarked Case. In English, T assigns the array N-A, with Nominative (the unmarked Case) mapped onto the highest nominal. Empirically, agreement is thus typically with the subject of a transitive verb, but when a verb is intransitive, only Nominative is assigned and agreement with the intransitive

[17] This is one respect in which the Case-assigning ability of T is not like an uninterpretable feature, in that it does not cause ill-formedness if it is not assigned (a parallel with Bošković 2007). The Case-assigning potential is simply inert when there is no mapping onto nominals in its domain. The same is true of agreement.

verb argument is the result, regardless of whether or not the single argument is a complement or a subject, a matter elaborated on in section 4.3 for English *there* constructions.

If we add the assumption (as do other P-M Case theorists) that Oblique Cases are invisible to prominence marking, and thus invisible to Case Clash, then it will also be possible for verbs with two arguments, one Oblique, to assign just one Case, Nominative. In languages like Icelandic, there is a richer variety than in English of oblique Cases that are lexical or inherent. As a result, verbs that assign lexical Case to their most prominent argument, will leave only one argument for Nominative Case to be mapped onto. If we assume that T assigns the Nominative-Accusative array and that C assigns Nominative, then a Dative argument in Icelandic, immune to Case assignment by C (and hence immune to Case Clash), can then satisfy EPP (see next section). In (15), for example (from Jónsson 1996: 143), the experiencer argument is lexically determined to be Dative, so the verb complement *Þessir sokkar* gets Nominative Case. T agrees with the highest Nominative, which in this instance appears to be the less prominent argument.

(15) Jóni likuðu Þessir sokkar. *Icelandic*
Jon.DAT like.PL these socks.NOM
'Jon likes these socks.'

There are a variety of more complex Case-marking situations in Icelandic, but this theory follows the logic of Bobaljik (2008), and can claim the same sorts of empirical successes and problematic cases given the basic concordance with principles presented here (see, for example, Bobaljik's account of defective intervention and partial agreement).

A particularly attractive result of the P-M Case theory is that it permits, in Bobaljik's rendering of it, a unification of the rule for subject-verb agreement across languages that have Nominative-Accusative (N-A) Case patterns and languages that have Ergative-Absolutive (E-A) Case patterns. If the unmarked Case in E-A languages is Absolutive, then agreement is predicted to be with the less prominent of two arguments for typical active transitive verbs. Bobaljik is not specific about how the E-A array is assigned, but, in the spirit of Marantz' analysis, suppose that the mapping onto argument prominence is inverted, such that Absolutive is mapped first onto the argument assigned the lowest prominence, and then Ergative is mapped onto the more prominent argument, if there is one. For the sake of discussion, let us assume that either T or Aspect assigns the E-A array and that the Ergative Case has the status of an Oblique Case for subsequent assignment (it is immune to Case Clash).[18]

The charm of this account is especially evident for split Ergative languages

[18] I am aware that this brief determination that Ergative is a structural Case does not do justice to the literature, see for example, Legate (2008: 58) where a number of sources are cited in favor of the view that Ergative is an inherent Case. For the purposes of VSA, treating Ergative as inherent would simply make the E-A pattern similar to the Icelandic Dat-Nom pattern, but more general. Although my position on these matters is not crucial to VSA, it is notable that in split Ergative languages like Hindi, the same argument that is Nominative in one aspect but becomes Ergative in another, which suggests that the term 'inherent Case' is atypically applied to Ergative in such languages.

like Hindi, where only imperfective clauses show the E-A pattern. Suppose that a perfective aspect node can intervene (or merge with T) and count as an inverse Case mapping head, resulting in Erg >> Abs mapping to prominence, as in (16a) (examples from Bobaljik 2008).

(16) a. Raam–ne RoTii khaayii thii. *Hindi*
Ram–ERG bread–Ø.FEM eat.PERF.FEM be.PST.FEM
'Ram had eaten bread.'

b. Siitaa kelaa khaatii thii.
Sita–Ø.FEM banana–Ø.MASC eat.IMP.FEM be.PST.FEM
'Sita (habitually) ate bananas.'

c. Siita–ko larke pasand the.
Sita–DAT boys–Ø like be.PST.MASC.PL
'Sita likes the boys.'

Where the verb is not perfective, the Nom >> Acc pattern remains the default as in (16b). If the subject is Oblique for any other reason (i.e. where the subject is Oblique but not Ergative), then the subject is invisible for T mapping Nom >> Acc, and the object gets unmarked Case (Nominative), as in (16c), just as in Icelandic. Agreement follows the unmarked Case, as indicated by the plural marking on the verb in (16c).

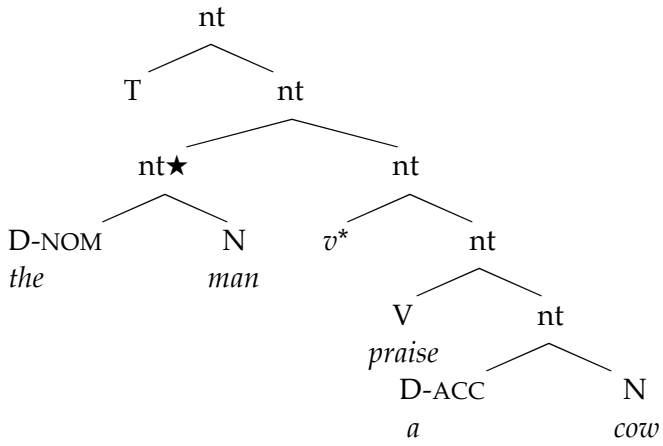
These results suggest that there is strong independent motivation for a P-M Case theory. The rest of this section develops a particular instantiation of a P-M Case theory that serves the goals of VSA.

4.1. Case Prominence in VSA

Adapting a P-M Case theory to VSA requires some non-trivial adjustments, but none that violate the core ideas of such theories.

For example, since there is no projection of categorial features in VSA, hence no DPs, it is necessary to be more precise about the description, 'nominals in the domain of a Case-assigning head'. The unit assigned Case is D, and so a Case-assigner must rank all of the Ds in its domain for prominence. Intervention will block access to a D that is in the domain of any other Case-assigning intervener, such as C, T, P, or another D. In (17), which diagrams a point in the derivation before *the man* merges above T (about which, more later), the Case domain of T includes two Ds (*the* and *a*), such that neither of the two Ds c-commands the other. Non-terminal nodes are simply marked 'nt'.

(17)



Yet T must rank *the* higher than *a*. The ranking proceeds as follows: T ranks D_A >> D_B if, within the domain of T, there is a node above D_A (illustratively marked with ★), that c-commands D_B. This will usually insure that what is commonly called DP will be the unit that contains D_A and c-commands D_B, and the ranking will proceed accordingly within the $Case_D$ of T. The non-Case-assigning heads between T and D_2 (*a*), such as *v** and V do not intervene for the $Case_D$ of T. The highest ranked D is then assigned NOM(inative) and the phi-features of that D-NOM determine the shape of agreement on T. In more richly inflected languages where N bears a Case that matches its determiner, I assume that N gets that assignment by virtue of being in the domain of D at PF, where the morphological Case assignments are spelled out. In this sense D is also a Case-assigner, and as such is an intervener blocking access to its domain by any higher Case-assigning head.

The core innovation of this Case system now rests on the assumption that any D assigned Nominative by a Case-assigning head can move to a higher domain and still be eligible for a replacement Case assignment, but a D assigned some other structural Case will accumulate Case assignments and be ruled out by the morphological filter in (18).[19]

(18) *Case Clash*

At the point of linearization, a D must not bear more than one Case.

I make four further assumptions about Case and one concerning EPP.

Oblique Case, once assigned (and by whatever assigns it), is invisible to Case prominence. Thus when T has two Ds in its $Case_D$, A and B, and A is more prominent in the argument structure of the verb, but has already been marked

[19] Languages with more than one Case affix on D are unexpected if (18) is universal, or at least, inflexible, but (18) is essentially a morphological filter, and may plausibly be flexible with respect to language-specific morphological properties. Other apparent counter-examples to (18) as a universal include Case-attraction phenomena, whereby a *wh*-phrase in a relative clause structure bears a different Case from the one it would receive *in situ*, even if that Case is ACC(usative). These effects raise issues for most Case theories, and so I will not explore the possibilities here. Notice that I do not assume a general condition that D must bear a Case, an issue that arises in section 7.

Dative, then only B is visible for Case prominence. This is just the VSA instantiation of the idea already introduced in P-M Case theory to account for agreement patterns like those in Icelandic.

All proper names that receive Case occur with (sometimes null) determiners. This is a theory-internal requirement, since I assume that N cannot be ranked for Case prominence, but determiners do co-occur with proper names in the world's languages, for example, in German dialects and Greek, and the existence D in nominals containing proper names has also been argued for by Longobardi (1994).

The generalization about 'unmarked' Case is probably more neutrally stated as 'reassignable' Case, or R-Case, since Nominative clearly has a marker in many languages (e.g., Icelandic). While this is often true of Nominative and Absolutive marking in the world's languages, it is by no means fully general, and so I will leave the morphological Spell-Out of R-Case to language-specific morphology (see Legate 2008, for a similar conclusion, although in a non-PM Case theory). Thus I would restate Bobaljik's agreement proposal as follows:

(19) Finite T agrees with highest R-Case in its domain.[20]

The nullity of PRO is not derived from Case assignment under prominence nor by the absence of Case; rather PRO is Case-marked and, even in control structures, bears a Case according to its context, just as other nominals do (see Landau 2006: 154-157 and references cited there). I assume that PRO is assigned Case by the C that introduces infinitives, unless the EPP is satisfied by an Oblique (as some of the concord phenomena with the Case of PRO show in the reference cited).[21]

Finally, I assume the EPP as in (20):

(20) *EPP*

T_D must be assigned to a non-terminal.

This stipulation is no more conceptually attractive than several others in the

[20] I abstract away from those languages where a verb agrees with more than one argument, as such cases are orthogonal to the discussion of VSA instantiations.

[21] Since the demise of the PRO theorem with the theory of government, no interesting alternative to account for the nullity of PRO has emerged, other than Hornstein's (1999) movement approach, which remains controversial (see, amongst others, Landau 2000, 2003, Hornstein 2003, Boeckx & Hornstein 2006, and van Urk, forthcoming) and which I will not address here. I will not attempt to provide a theory why PRO is null in this essay. For further critique of the place of PRO in traditional Case theory, see McFadden (2004) and references cited there.

In so far as the Case assigned to PRO is not realized in morphology, for whatever reason, I depart further from Bobaljik's (2008) view that Case assignment is assignment of 'm-Case', a form of Case that always has a phonological form. I have to assume that what morphological Spell-Out sees is somewhat more abstract, as I do for R-Case, but Case is not the key regulator of movement for passive or raising (or the distribution of PRO) as it is in GB-era abstract Case theories. Although I have not compared this account to that of Legate (2008) in any detail, my position on the relation to M-case is more in line with her position than with Bobaljik's.

literature, insofar as it involves an interpretive assignment with no interpretive content, but it does eliminate the last justification for an operational trigger for movement. As I will demonstrate below, this way of formulating the EPP will avoid many of the complications that other theories are forced to without introducing counter-cyclic instances of Merge (as are found, for example, in Chomsky 2007a, 2008, where T must inherit features from C before it can attract a constituent to its Spec-position) and without introducing features that are crucially uninterpretable in some other component. The treatment of EPP in (20) holds out the hope that the right theory of 'subjects' may yet reduce EPP to a contentful form of domain assignment, such as scope or argument prominence, perhaps as Rizzi (2006) has suggested (see the discussion of criterial positions in section 5).

As a means of fleshing these ideas out as they apply in specific cases, consider the annotated derivation in (21a–h) for *Mary hit John* and the resulting structure in (21i). In contrast to Chomsky (1995, 2001), I assume that there is no numeration specifying the input to structure-building because the only arbiter of what constitutes a well-formed output is whether or not the result of the derivation is semantically and phonologically interpretable at the interfaces.

(21) Mary hit John.

a. By hypothesis, D must merge with the name *John*. Similarly, a D merges with *Mary*.

b. If *hit* is merged with [D *John*], then [D *John*] is the selected domain for *hit*, hit_D, irrevocably formed at this point in the derivation because *hit* selects for D.

c. If v^* is Merged to [*hit* [D *John*]], then v^*_D is [*hit* [D *John*]] because v^* selects for V.

d. If [D *Mary*] is Merged to [v^* [*hit* [D *John*]]], then [D *Mary*] first c-commands v^*_D (and v^*) and [D *Mary*] is assigned v^*_D (i.e. [D *Mary*] is the external argument of *hit*).

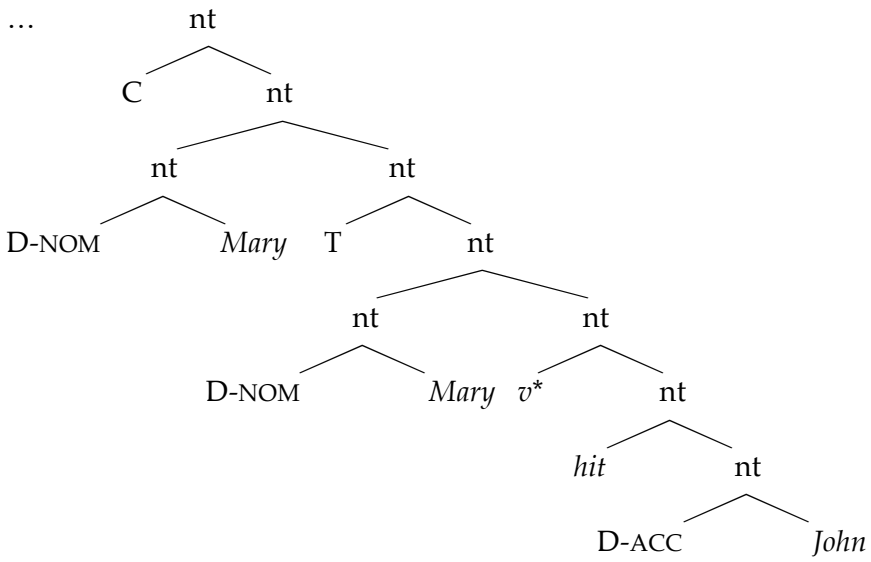
e. If T is Merged to [[D *Mary*] [v^* [*hit* [D *John*]]]] and T selects for v(*), then [[D *Mary*] [v^* [*hit* [D *John*]]]] is T_D. Since T sets prominence for Case in English NOM >> ACC, and the D of [D *Mary*] is more prominent than the D of [D *John*] in the $CASE_D$ of T, [D *John*] will be ACC and [D *Mary*] will be NOM.

f. EPP requires that a non-terminal must be assigned T_D and since [D *Mary*] has R-Case (Nominative), it is the only candidate to first c-command T and T_D by iMerge.

g. When C Merges to [[D *Mary*] [T [[D *Mary*] [v^* [*hit* [D *John*]]]]]], the $CASE_D$ of C includes the higher [D *Mary*] only, since T intervenes for the rest, and the [D Mary] is re-assigned NOM by C. If [D John] had satisfied EPP, it would fail by Case Clash.

h. The output: [C [[D *Mary*] [T [[D *Mary*] [v^* [*hit* [D *John*]]]]]]]

i. …



I assume for now that, given multiple occurrences, only the highest is pronounced (see section 6), thus *Mary* is pronounced outside of T. A few comments are necessary here concerning C. I am assuming first that T is always selected (i.e. by C or a lexical head), as seems necessary to say independently of any considerations peculiar to this theory. Notice that either iMerge or eMerge could satisfy EPP, but for reasons that remain murky for English, transitive expletive constructions are not allowed, and so *there* cannot be inserted at step (21f) (a problem shared by most other theories). In the absence of an eMerge option in this instance, movement to subject is forced by the domain assignment condition embodied in EPP. As remarked above, [D *Mary*] is marked NOM twice, but without Case Clash, because one Nom assignment (by C) simply replaces the other (by T). If [D *John*] had iMerged to first c-command T_D at step (21f), then [D *John*] would induce Case Clash, since ACC on [D *John*] cannot be reassigned any other Case (including another ACC).

The same system derives the sentence in (22), an instance of raising. Notice that I assume that infinitival T, like tensed T, assigns Nominative Case and, as in standard accounts, that there is no C in English raising infinitives.

(22) Several men appear to be leaning on the balcony.

a. The first steps include Merge of *balcony* and *the*, where *the* selects N, then Merge of *on* with *the balcony*, where *on* selects D. Since *on* is a P, it is a Case assigner and assigns Oblique to the most prominent (only) D in its domain which is *the*.

b. Merge of *leaning* (I ignore *–ing* here) to *on the balcony*, where *lean* selects P and takes *on the balcony* as its domain, depending on whether or not *on* is an intervener, since *the balcony* is also the domain of *on*. Merge of *v** then takes *leaning on the balcony* as its domain and Merge of (previously formed) *several men* fills the external argument (EA) slot for *v** (i.e. *several men* is assigned v^*_D).

c. *Be* is a V that Merges to [[*several men*] [*v** *leaning on the balcony*]],

setting the contents of the latter constituent as be_D.

d. T (*to*) Merges to [*be several men leaning on the balcony*] selecting the latter as its domain up to intervention. T detects the prominence relations in its domain, but since *on* is a Case assigner that intervenes between T and *the*, T orders only *several* and assigns it Nominative. Infinitival T lacks agreement features.

e. [*Several men*] iMerges to [*to be several men leaning on the balcony*]. This step is derivationally optional but necessary to satisfy the interpretive condition EPP.

f. Then *appear* Merges to [*several men* [*to* [*be several men leaning on the balcony*]]], and then *v* Merges to [*appear several men to be several men leaning on the balcony*].

g. When [+tense] T Merges to [*v appear several men to be several men leaning on the balcony*], two things happen. Since T is a Case assigner, it ranks all the Ds in its $CASE_D$. Since subordinate *to* is also a Case assigner, it intervenes, and so the only D in $CASE_D$ of [+tense] T is the higher occurrence of *several men* which is not harmed by being reassigned Nominative. Also, [+tense T] has agreement features, so it agrees with the highest Nominative in its agreement domain, *several men*.

h. Then *several men* iMerges to [T v *appear several men to be several men in the room*]. This satisfies EPP.

i. C merges to [*several men appear several men to be several men in the room*] because [+tense] T is always selected by C. Since step (h) moved *several men* into its domain, C orders *several* as its most prominent D in $CASE_D$ of C and reassigns it Nominative.

'Pronounce highest occurrence' (see (52) in section 6 below) will derive the right phonology, although I will have more to say about which of multiple occurrences are pro-nounced in later sections.

One major advantage of this account over many others in the literature is that it is never necessary to assume any counter-cyclic movement. Chomsky (2008), for example, requires that movement to matrix [Spec,TP] position is only triggered *after* C is merged to TP, the features of C are inherited by T, and then iMerge is triggered from the lowest position of *several men* to the highest position of *several men*, but stopping in the lowest [Spec,TP] to satisfy EPP and then in the highest [Spec,TP]. No stutter-step anti-cyclic movement is required in the derivation of (22).

In the next two subsections, further mechanisms for, and consequences of, mapping Case and agreement onto prominence relations will be fleshed out.

4.2. Case Prominence and Derivational Drag

I assume that C, T, P, and D are potential Case assigning heads and that the intervention of any one of these heads blocks Case prominence ordering and assignment by any higher head. Lexical or inherent Oblique cases are not assigned

prominence or Case by the Case-assigning heads, such that a nominal K without lexical or inherent Case will count as the most Case-prominent nominal in its domain even if a nominal J with inherent or lexical Case c-commands K in the domain of K's Case assigner. Any nominal that is assigned more than one Case without reassignment will be excluded by Case Clash (as stated in (18)). In languages where Case is more frequently inherent or lexical, hence immune to Case Clash under reassignment (because they won't be prominence-ordered for Case assignment by a head), nominals with non-structural Case may move to higher Case domains with impunity. Case assignment by a Case-assigning head is obligatory unless there are no eligible D nodes in its domain.[22]

As mentioned with respect to PRO, I assume that infinitival C also assigns Case. One way of blocking raising from the position of PRO is simply to assume that infinitival C always assigns Accusative, whether it is null C or English *for*. On this account, (23b) and (24b) are excluded for the same reason, namely, iMerge of *him* in a domain of higher Case assignment will result in Case Clash.

(23) a. It is important for him to leave.
b. * *He* is important [for *him* to leave].

(24) a. It is important [C_{NULL} PRO to leave].
b. * *He* is important [C_{NULL} *him* to leave].

It is instructive to see how this reasoning applies to passive structures. In English, when a transitive verb like *praise* is in its passive form, it takes no external argument (suppressed by a rule applying to the lexical entry) and is selected by *v*, which, unlike *v**, does not select a domain that can be assigned to an external argument. For the moment, let us assume that no argument is assigned to v_D, a result to be derived momentarily. As a result, there is only one nominal in the domain of T, the direct object, and that nominal gets R-Case. The nominal with R-Case is the only one eligible to be reassigned Case by C, and so it is that argument that satisfies EPP by being assigned T_D.[23] Where a locative expletive (*there*) is available to satisfy EPP, the object may remain *in situ* and receive NOM as the only visible nominal visible to Case assignment in $CASE_D$ of T.[24]

---

22 I do not have an explanatory way of addressing the difference between German and Icelandic, where Dative appears capable as acting as a subject in Icelandic, as expected, but not in German. Stipulations about the compatibility of C with a Dative D in its domain could perhaps express the difference, but I have not looked into any independent motivation for such an approach.

23 Interesting questions emerge about psych predicates from this perspective. For example, there could be a difference between oblique lexical or inherent Case arguments of psych verbs such that some occur in subject of v position where for other verbs they occur in object of V position, with resulting contrasts. One might attempt to express the contrast between classes of psych verbs in this architecture, but these are issues I will not explore.

24 Pseudo-passive, on this account, will become possible whenever Case assignment by a preposition is neutralized such that it is not an intervener for Case assignment by a higher Case-assigning head. Then the nominal complement of P will be ordered in the Case domain of the higher head, receiving Nominative if there is no other visible argument more prominent in the domain.

At this point we have seen that there is a derivational drag effect on arguments assigned structural Case that is not R-Case, such that only an R-Case nominal can move to a higher Case domain.[25] The effects of this restriction have been shown to be positive, insuring that all non-R-Case structural cases are frozen in place, while permitting only R-Case nominals, or Oblique Case nominals, to move to higher Case domains. However, even A'-movement of non-R-Case nominals that have structural Case is also apparently ruled out in this system, a matter to which I return in section 5.

Another interesting derivational drag effect is induced by these assumptions about Case assignment. If a nominal moves within in $CASE_D$, it will move to a position that c-commands the copy of it that is in its point of origin. In such instances, both nominal positions will be evaluated for prominence in Case, with the result that two occurrences (OCCs) will have different Case assignments, where the lower one cannot be reassigned. This will result in Case Clash.

(25) $[H_C$ [OCC … [… OCC …]]]

[25] Prenominal Genitive nominals (PGNs) are generally assumed to c-command the D of the nominal they are contained in (e.g., they are assumed to be in '[Spec,DP]'), which means the D of the PGN is in the $Case_D$ of the Case assigning head external to the nominal. That Case-assigning head should determine the prominence prenominal D >> containing D, resulting in the wrong Case array (e.g., for a D that should be Nominative, such as *mother* in *John's mother laughed*) and the Genitive D could also be interpreted as a potential complement by a selecting head.

The Case problem disappears if Genitive is an oblique Case, but Genitive appears also to be structurally assigned. Since Spec-Head agreement is not available in VSA, the only way to model Genitive assignment in this theory is to posit a Gen-assigning head H-Gen that c-commands and Case-assigns the D of the PGN. If H-Gen is itself a determiner (one that can bear a different Case than it assigns to a D in its domain) then PGNs must originate below H-Gen. The following contrast suggests that some PGNs are below H-Gen and others are above it.

(i) What kind of bread was she willing to buy a loaf of?

(ii) * What kind of bread was she willing to buy John's loaf of?

(iii) Which of his latest escapades was she willing to hear an explanation of?

(iv) ? Which of his latest escapades was she willing to hear John's explanation of?

It is pointed out in Safir (1987) that extraction out of nominals that have PGNs is improved if the PGN bears a thematic relation to the head N. In the latter situation, it appears that the PGN is optionally assigned a theta-role by merger above [N N_D] (or perhaps a nominal version of *v**). Thematic roles undetermined by the head N seem to be assigned a default theta-role by merger above [H-Gen $H\text{-}Gen_D$]. If so, the PGN initially emerges below H-Gen where it receive Case, but the PGNs not thematically related to the head N must raise to receive a theta role above H-Gen. High PGNs block extraction from the full nominals more completely, which would be explicable if the left periphery of the nominal domain is an escape hatch for extraction. This account requires a theory of low PGNs, an analysis, which happens to be consistent with VSA tenants of Case assignment. Moreover, the mechanisms described insure that PGNs are already theta-assigned before a higher selector can assign any theta-role to them, and perhaps in certain contexts more than one thematic assignment is possible (e.g., possessor raising constructions). A more fleshed out version of this proposal is too large for discussion here.

On this account, it is impossible for any nominal in the Case domain of a Case-assigning head, H_C, to move to another position within the domain of H_C. This will prevent iMerge of the direct object just above [v v_D] in passive structures, or else the configuration in (25) will arise within T_D. This would appear to rule out a great deal of possible movement, perhaps too much, especially in scrambling languages. I return to this issue in section 5, where the relation of A'-movement to Case assignment is examined.

4.3. Relativized Domains: Case and Agreement

Chomsky (2008: 154) contends that

> [uninterpretable features] must therefore be valued at the stage in computation where they are transferred—by definition, at the phase level—and this must be the same stage for both transfer operations, again supporting the optimal assumption that transfer to both interfaces is at the same stage of derivation.

This assumption is not necessary, it is not obvious that it is the optimal assumption, and, moreover, some of the stipulations required to support this hypothesis suggest it is misconceived. For example, Chomsky appeals to EPP features in raising structures to attract movement within a phase to potentially many subject positions, yet Agree only applies between the highest T and an argument embedded potentially several TPs below. This difference between attraction by EPP features and Agree suggests that phases are just cycles of a particular kind, not the *only* cycles. Rather it would appear that cycles of different kinds, and, in VSA, interveners of different kinds, are *relativized* to the sort of relation that they establish. In other words, the notion that there are windows (domains) in the course of a derivation when certain relations can or must be established is preserved, but it is not assumed to be the same window for all relations.[26]

Relativized intervention can be illustrated with the difference between Case assignment and agreement as modeled in VSA. For example, a Case-assigning node H will set a domain $CASE_D$, its sister, which includes all the nominals in $CASE_D$ outside the next lower Case-assigning head. If there are two nominals, then they will be ranked for prominence (ordered) and aligned with a Case array that is determined by H. $CASE_D$ for T_1 is everything dominated by its sister up to intervention, i.e. the $CASE_D$ for T_1 does not extend any lower than the sister to the next Case-assigning head that T_1 c-commands. If the closest Case-assigner to T_1 is another T, T_2, then anything $CASE_D$ for T_2 is not in the ordering that includes those nominals in $CASE_D$ for T_1.

[26] For arguments that A-movement and Agreement have different domains, see Bobaljik & Wurmbrand (2005), and for an argument that the domain of Principle A is not congruent to the domain of A-movement, see Safir (2004a: 147-156). Seely (2006: 202) points out that the domains for selection and Agree are different. Bošković (2007) argues in particular that phases and the domain of agreement are different, and that A'-movement does not involve intermediate triggers for a constituent which has features not satisfied by any structure in the tree, but could involve Agree applying across a larger domain when the probe is introduced.

We have seen how this works for simple cases like (21) where T assigns the Case array NOM >> ACC and that T agrees with the highest NOM in its domain, namely the subject when it is in EA position. In (26), however, agreement is with the nominal within the complement of *be*.

(26) There are three men missing.

In this example, T orders the only argument in its domain (*three men*) for Case prominence and then assigns it NOM Case. However, instead of raising *three men* to the matrix position to eventually satisfy the EPP requirement of matrix T, *there* has been inserted. Such cases as these recall the Icelandic DAT/NOM pattern, where agreement is with the highest NOM even when it is not a subject. Such a parallel suggests that *there* plays this role because it is locative, and thus an Oblique Case.[27] As in Icelandic, it may then be assumed that *there* satisfies the EPP requirement of matrix T, it is in the Case prominence domain for C, but it is not assigned Case by C. (After all, the use of a locative in subject position with postverbal agreement is also found in English examples like *In this village are found many fine woven goods*.) In other words, I am treating *there* as something like *D.Loc-here*, where locative meaning has been bleached away. This then provides a natural account for the distinction between *there* and Standard English *it*, which is not oblique and which agrees with the verb. However, examples like (26) still correlate Case and agreement, insofar as the highest NOM, the associate of agreement, is within $CASE_D$ of T.

To see how the setting of domains is relativized, consider, for example, ECM structures in English and let us suppose that T_1 is the matrix tense T.*past* which takes everything in its sister non-terminal node as its domain, and that the next Case ordering domain is that of T.*to* of the subordinate clause.

(27) a. He T.*past* expected her T.*to* hate them.
b. There T.*past* were expected T.*to* be several men in the room.

For (27a), *her* and *them* will be Case-ordered and assigned by T.*to*. The pronoun *her* receives R-Case from T.*to*, but *them* will be marked ACC and thereby frozen in place. When *her* raises up to be assigned the domain of T.*to*, as it must to satisfy EPP, it will then be embedded as *expect* merges to [T.*to* $T.to_D$] and selects T.*to*. When *he* is merged to *v** above *expect*, it becomes the external argument (EA) of *expect* and will be in position to receive R-Case from matrix T.*past* which will order *he* >> *her* for Case prominence and assign ACC to *her*, overwriting R-Case. *He* will then be the first c-commander of $T.past_D$ above T.*past* and satisfy the EPP. When C is Merged, *he*, which has R-Case from T.*past* will be reassigned R-Case (NOM) by C.[28]

[27] It may be necessary to assume that there is a locative head between T and v that assigns a LOC argument to its domain. See Linares (forthcoming) for a proposal along these lines.

[28] If the C of tensed clauses assigns R-Case, then in theory, the R-Case subject could raise into the next Case domain. This does not normally happen because movement to a position immediately above C_D is movement to a criterial position, which a non-*wh*-subject is inappropriate for (see section 5.2) and movement to a higher position will not be able to escape

In (27b), however, *several men* is the highest nominal in the lower domain (that of T.*to*) and receives NOM in situ. *There* satisfies the EPP by being the first c-commander of [T.*to* T.to_D]. Subsequent iMerge of *there* satisfies the EPP requirement for matrix T.*past*. The issue now concerns agreement of T.*past*, which is empirically plural. Insofar as locative *there* counts as an Oblique Cased nominal, neither matrix *there* nor its subordinate copy qualifies as the nearest R-Case (NOM) to T.*past*; rather T.*past* agrees with *several men*, even though *several men* is in the domain of T.*to*. However, in the relativized system assumed here, T.*to* is not an intervener for agreement because it does not bear agreement, i.e. it is not in an agreement relation with anything, so the domain of T.*past* with respect to agreement includes the contents of the lower clause and its highest NOM, which is *several men*.

In a theory where phases are not relativized to relations in this way, it is necessary to assume that *there* carries a plural agreement feature or that the lower phase is invisible selectively or generally in this sort of example. Rather than complicate the derivation in the latter way, intervention is relativized to the relation involved. A natural account of long distance agreement in *there*-sentences with raising is the result.

5. Criterial Positions and Reducing the A/A′-Distinction

What have been known since Chomsky (1981) as A′-movements, the quintessential example being *wh*-movement, have a different class of properties from A-movements, which seem to revolve around movement to subject(-like) positions. The A′-movements displace constituents in such a way that the determination of scopal properties or information-structure values are almost always involved (although *tough*-movement structures raise questions), as contended most recently and explicitly, for example, by Rizzi (2006), whereas A-movements appear to revolve around Case and agreement relations, *not necessarily* involving scope or information structure (though they often do have effects of this kind). Chomsky (2007a: 25) still distinguishes them as follows: "A-movement is IM contingent on probe by uninterpretable inflectional features, while A′-movement is M driven by [edge features] of P" (where IM = iMerge).

That Case requirements are involved for A-movement, but not for A′-movement, is a distinction that should be effaced or derived. The Case prominence theory permits us to achieve this result, insofar as the only difference between A-movement and A′-movement resides in the different strategies they exploit to avoid Case Clash, not with respect to features that attract movement, nor with respect to the requirements of interpretation they satisfy (e.g., edge features appear to be linked to the satisfaction of interpretive requirements).

Consider simple examples such as (28a), where it is clear that *whose brother* must reach the position where its scope is assigned, which would be the point in

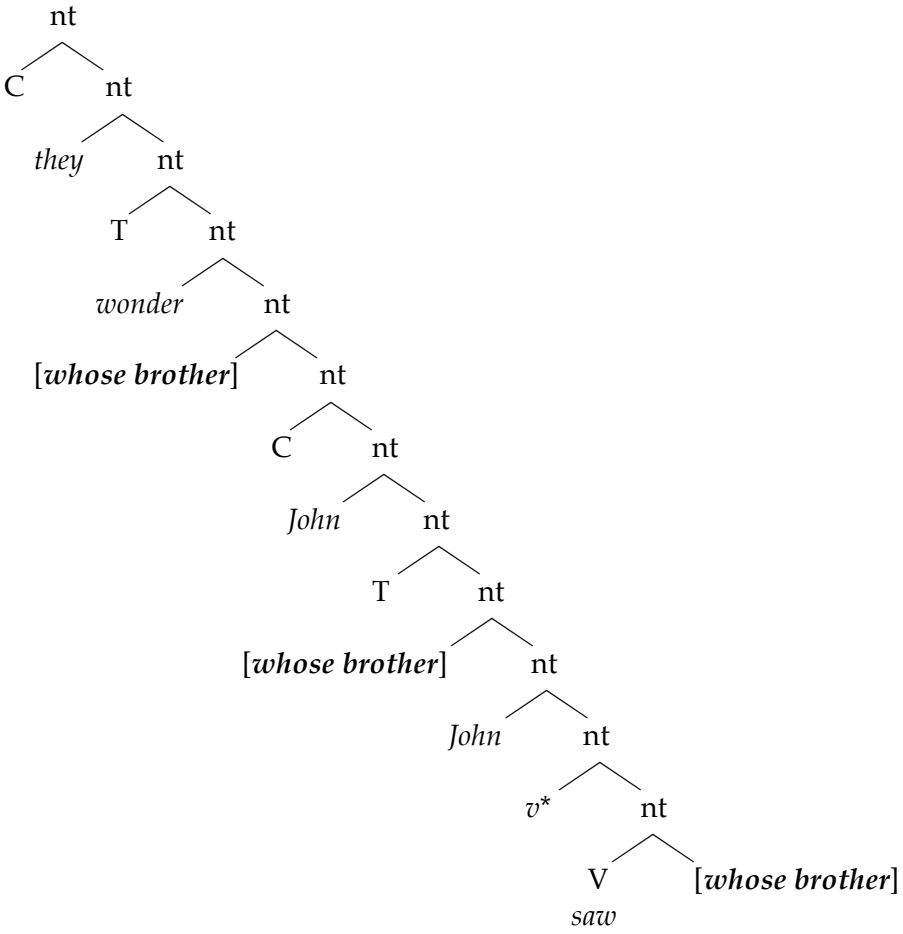
---

forced Spell-Out in domain C_D (see section 6). There are languages where the subject of a tensed clause can raise into the next higher domain and receive ACC (Korean), and for that to be possible, it must be assumed that there is a C (or some other domain-assigning head in the intermediate zone above C) which sets a domain compatible with a non-*wh*-phrase.

the derivation where it iMerges just above C. This it presumably must do in order to achieve the proper assignment for interpretation. However, it is not obvious that the Case theory proposed so far can produce the right result, especially given the schematic tree in (28b).

(28) a. They wonder whose brother John saw.

b.



The representation in (28b), where *whose brother* has 'adjoined to *v*P' en route to its final landing site, exposes *whose brother* to Case assignment by T in the lower clause, where the higher occurrence would get NOM and the lower one ACC, leading to Case Clash. Moreover, *John* would receive ACC from subordinate T instead of NOM, and should then induce Case Clash when it moves to subject position to satisfy. Even if we solve the problem below subordinate T, subsequent iMerge of *whose brother* above the embedded C exposes it to Case assignment from matrix T, which would assign it Accusative.

In this section, I will explain how the scope of overt wh-movement is assigned in VSA and how the solution to the problem raised by Case assignment resolves the A/A'-distinction without relying on pair-merge or any other special addition to VSA that is not required by other theories. I postpone to section 6 why it is that A'-movements do not typically take place directly to criterial positions, but must pause below their final destination to satisfy locality conditions associated with cyclicity.

5.1. Revising Extension

Before I present my account of the A/A'-distinction, it is necessary to begin by introducing a revision of Extension which *is* required by other theories, at least in some form. Extension has actually been abandoned in some recent accounts, more typically in practice than by any explicit rejection (e.g., counter-cyclic movement in Chomsky 2008). Although the revision required has been motivated independently of the theory proposed here, it is a revision my subsequent proposals will exploit.

Many accounts of anti-reconstruction effects depend on violations of Extension as stated in (2) (see Lebeaux 1991, Chomsky 1995, Bhatt & Pancheva 2004, Fox 2002, Stepanov 2001, and Safir 2004 for discussion). The variable-binding contrasts below are from Safir (1999: 601–602).

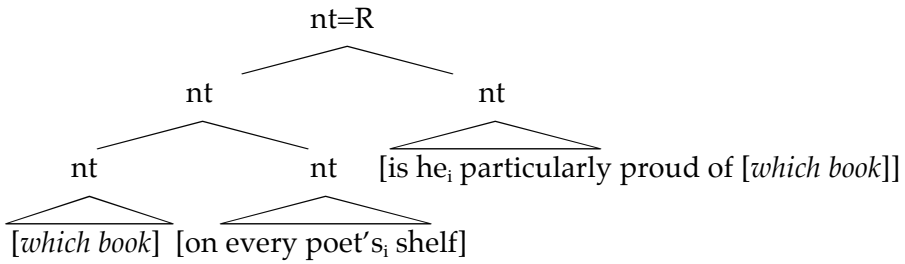
(29) a. * [Which reviews of every poet's$_i$ book]$_j$ does he$_i$ try to forget t$_j$?
b. ?? [Which analysis of every poet's$_i$ book]$_j$ is his$_i$ mother most afraid of t$_j$? {answer, for example, 'the Freudian one'}
c. ? [Which reviews of every poet's$_i$ book]$_j$ t$_j$ give him$_i$ the most satisfaction?

(30) a. [Which book on every poet's$_i$ shelf]$_j$ is he$_i$ particularly proud of t$_j$?
b. [Which book on every poet's$_i$ shelf]$_j$ is his$_i$ mother most proud of t$_j$? {'The one dedicated to her'}
c. [Which book on every poet's$_i$ shelf]$_j$ t$_j$ gives him$_i$ lasting satisfaction?

The bound reading of *he* in (29a) fails on the assumption that a copy of *which reviews of every poet's book* inhabits the position notated with a trace, hence *he* c-commands its antecedent, leading to a violation. Similarly, a copy in the position of the trace in (30b) would also induce a weak crossover effect, but not if the copy were in subject position (30c). It appears that certain prepositional phrases must be attached before iMerge of the whole *wh*-phrase takes place and these PPs leave a copy behind that results in variable binding violations. The effects disappear in (30), however, as if *on every poet's shelf* were not part of what is left in the trace position (following a line of reasoning from Lebeaux 1991). Safir (1999) proposes that if *on every poet's shelf* is merged to *which book* after *which book* moves, then strong and weak crossover effects are avoided. But merge of *on every poet's shelf* to *which book* after iMerge of *which book* violates Extension because *on every poet's shelf* does not extend the undominated node R, as illustrated schematically in (31b) and in the diagram in (31c).

(31) a. [$_R$ [*which book*] [is he$_i$ particularly proud of [*which book*]]
b. [$_R$ [[*which book*] [on every poet's$_i$ shelf]] [is he$_i$ particularly proud of [*which book*]]

c.



Every appeal to late attachment (or 'late merge') in the literature involves an Extension violation of this kind.

It would be unfortunate, however, to jettison Extension (or just stipulate that it does not restrict 'adjunction', as in Chomsky 1995: 189–190) from the theory entirely, as it is the condition that prevents a rich variety of counter-cyclic movements, possibilities that greatly enhance the descriptive power of grammars. It is still possible to capture the idea, however, that Extension is always sensitive to the position of the undominated node, if we provide a more articulated idea of what the top of the tree is.

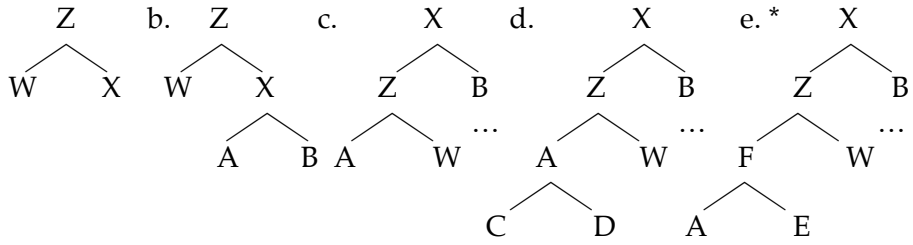
(32) *Revised Extension*

After every instance of Merge, M^i, the undominated node of the resulting structure immediately dominates a node it did not immediately dominate before M^i.

The structures in (33a–d) result if W has just been Merged (abstracting away from linear order) and Revised Extension has been respected. This definition exploits the long held assumption that Merge always produces binary branching trees. Cases where both terms of Merge are terminals (themselves undominated nodes), as in (33a) will be immediately dominated by a new node Z, which is the non-terminal formed by M^i. Since Z is new what it immediately dominates is newly dominated by it, and the same account extends to (33b), where one (or both) terms of Merge are non-terminal(s). However, a more novel possibility arises where W adjoins to A, as in (33c) or (33d) (where A is terminal or is not, respectively) in each case creating Z, a node newly immediately dominated by the undominated node X after M^i applies, consistent with (32).[29]

[29] Non-terminal nodes do not bear labels in VSA, so it is fair to ask how we would know that a particular non-terminal node that was not dominated 'before' was *in fact* not dominated before. As a technical matter, suppose that every node created by Merge is assigned a 'term index' and every term of Merge is assigned a new term index only if it does not already have one. In cases where Revised Extension is satisfied by submerging one immediate daughter of the undominated node, but not the other, the undominated node is not a term in that Merge operation; only the nodes that together form a new node are terms in that operation. The term index is then what is copied in iMerge and could then replace the very similarly employed 'numeration index', which will be eliminated in VSA, which has no numerations. The ability to describe copies of both non-terminals as sharing the same term index, just as terminal copies do, would simplify references to copy sets (for terminals and non-terminals), especially in cases where a terminal node has been extracted from a non-terminal copy.

(33) a. b. c. d. e. *



As a matter of useful terminology, let us say that *W subMerges A* in (33b–d) by alienating A from immediate domination by the undominated node, a status A loses after M^i applies. Notice further that Merge could not subsequently apply to some node E to create a new node F immediately dominating A and E in (33b–d) because that would violate Revised Extension: The undominated node X still dominates the same two nodes, Z and B, before and after E is merged.

Revised Extension permits more attachments than Extension in (2), but it still insures that Merge operations grow only the top of the tree.[30] Some instances where nodes are subMerged, however, permit structure-building crucial to a variety of current proposals that are technically excluded in the earlier formulation. For example, both late Merge, described above, and 'tucking in' as proposed by Richards (1999) are instances of (33d). The structure in (33c), where the terminal W subMerges the terminal A, models iMerge of a head to another head, an operation frequently appealed to in the literature, and one long known to be problematic for Extension (e.g., as noted in Chomsky 1995: 327 and Bobaljik & Brown 1997). The latter three possibilities are all excluded by Extension as in (2), but permitted by Revised Extension.[31]

---

[30] SubMerge does not change any assumptions about derivational c-command (see fn. 6) nor should it be confused with questions about sideward movement. The subMerged node does not fail to c-command any position it c-commanded before subMerge. Sideward movement occurs when one term of Merge, A, is contained in K, and the other term of Merge is J, where J is undominated and does not dominate K. If A Merges to J, then J is expanded by the operation, satisfying Extension in (2), but the resulting iMerged copy of A does not c-command anything in K, including the copy of A in its launching site, as noted by Bobaljik & Brown (1997). Unregulated, sideward movement could be intersentential. Epstein *et al.* (1998: 103) argue that such a movement could not satisfy economy conditions, that is, it could never be triggered. For example, Agree could never hold between a probe and a goal not contained in its complement, but this will not do if iMerge is not triggered, as argued here. In VSA, however, intersentential sideward movement leads to non-viable relations anyway, since the iMerged element L does not c-command its copy L'. Where L and L' are not in a viable relation, nothing would prevent them from receiving independent semantic assignments (e.g., the pronoun *he* could be assigned a value independent of its copy in another constituent). Issues concerning the Spell-Out of multiple copies would still have to be resolved where c-command cannot regulate Spell-Out (see fn. 35).

The output of sideward movement could be viable if there is subsequent movement to a position that c-commands both copies, permitting criterial interpretation of the copies in the relevant way. Some accounts of parasitic gaps have this character, but they must rely on chain formation and c-command (e.g. Nunes 2004: 91). For reasons to doubt the theoretical legitimacy of chain formation, see Safir 2008: 345–346).

[31] It is reasonable to ask how well Revised Extension fits with the KF reasoning. In principle, the revision of Extension costs nothing, because Merge, which Extension restricts, is assumed to be prior to the KF in any case. Just the same, an implicit claim has been made that subMerge as in (33c–d) is part of the pre-linguistic operation that generates syntactic structure, a complication beyond simple Merge in (1), and I have no evidence for this (nor any obvious way to search for such evidence).

5.2. Criterial Positions

Rizzi (2006) calls positions that must be reached for a phrase to be interpreted a *criterial position* for that phrase (along the lines of the *wh*-criterion or the *neg*-criterion, as in the references he cites). He suggests that the range of criterial positions may include focus and topic positions as well as those that participate in scope interactions, along the lines of Beghelli & Stowell (1997) (see section 7). Indeed Rizzi even holds out hope that EPP is a criterial movement, which would be more suitable for the VSA approach and perhaps others, but the evidence for this further extension is currently slim. All of these criterial assignments are instances of (13b) for a particular kind of domain assignment. In this section, I explore (13b) as it relates to assignment of scope for *wh*-Q-phrases.[32]

(34) *Assigning Scope for* Wh-*Q Phrase*

The scope of *wh*-Q is assigned when *wh*-Q is assigned a domain that matches its quantificational features/properties.

I assume that a questioned constituent, the *wh*-Q-phrase, can only receive a scope consistent with its interpretation if it is assigned WH_D which is set by C_{WH}, so unless this configuration is achieved, or some other way to interpret the *wh*-Q-phrase is introduced, then the result is a failure of interpretation.

(35) *Domain Assignment Failure*

An element Y that must be assigned a domain of type X is not assigned a domain of type X.

Thus, Domain Assignment Failure occurs when no domain is assigned to an element that needs one or the domain that is assigned does not match the element it is assigned to, that is, it is a form of incoherence. IMerge is optional in VSA, not syntactically triggered, so cyclic movement of a *wh*-Q-phrase is optional even if it is to a non-criterial position, as it is when it is Merged above *v* (higher than the external argument) or above C_{THAT}, which does not set a scopal domain. Thus no intermediate triggers are needed for long cyclic A'-movement, but failure of the phrase to move cyclically does not lead to an interpretable outcome.[33]

There is an important consequence of this formulation of domain setting and assignment for cases of head to head movement frequently found in V2 languages, for example. Notice that C_{WH} sets WH_D at the moment it enters the derivation c-commanding WH_D (an instantiation Firstness in (14)), and so if C_{WH} is subsequently subMerged by head adjunction of T, WH_D is still indelibly set and ready for assignment when *wh*-Q-phrase is merged to the undominated node.

---

[32] This will turn out to be a sub-case of scope assignment generally, as in (62).

[33] McCloskey (2002) argues that 'spurious features' are needed to trigger movement to intermediate positions for long cyclic movement. In Safir, in preparation, I argue that no intermediate triggers are necessary, and that a more elegant account than any other available is feasible precisely when triggers are rejected. It is also unnecessary to introduce an Activation Condition, as in Bošković (2007) to determine what is visible to iMerge, since every node is visible to iMerge.

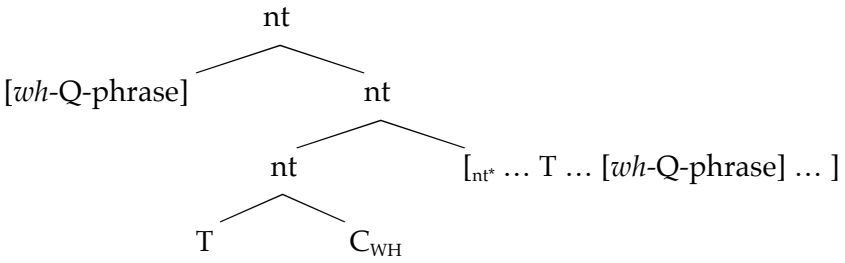
This is illustrated in the partial derivation in (36), where WH_D is not a real node label, but is the non-terminal (nt) marked with * that contains WH_D, as illustrated in the resulting structure in (36d).

(36) a. [C_{WH} [$_{WHD}$ … T … [*wh*-phrase] …]]
– Merge of C_{WH} results in the setting all it c-commands as WH_D

b. [[T C_{WH}] [$_{WHD}$ … T … [*wh*-phrase] …]]
– T subMerges C_{WH} (as in V2 languages)

c. [*wh*-Q-phrase [[T C_{WH}] [$_{WHD}$ … T … [*wh*-Q-phrase] …]]]
– *wh*-Q-phrase first c-commands WH_D (and c-commands C_{WH}), so WH_D is assigned as the scope of *wh*-Q-phrase

d.



It may not be the case that T-to-C in (36b) is required in every language (even covertly), but I include it to indicate that it does not change how scope is set and then assigned to the *wh*-Q-phrase. The same reasoning applies to the *wh*-Q-phrase in (31), which is assigned its WH_D before it is submerged by late attachment. Discussions of scope assignment to bare *wh*-Q and other quantifiers is reserved for section 7, where what have been called 'covert movements' are addressed.

5.3. A/A'-Distinction

Reconsider now (28a–b), where *wh*-movement resulted in Case Clash when copies were introduced by iMerge such that two occurrences of the same phrase could be assigned different cases, ruling out garden variety cases of *wh*-movement. Notice that the Case Clash issue for *wh*-movement, or A'-movements more generally, would disappear if the landing sites for A'-movement were invisible to Case assignment by higher heads, that is, once a phrase is A'-moved, it is no longer evaluated for Case mapped onto prominence. Something must insulate A'-moved nominals from further Case assignment or else typical utterances like (28a) would be disallowed.

As it turns out, VSA does not have to be revised in any way to permit *wh*-movement to avoid Case assignment, at least if a few plausible assumptions about the role of quantificational heads are adopted. For example, it is plausible to assume that *wh*-Q must have scope over its restriction, which consists of the larger phrase that moves with *wh*-Q, just as the *wh*-Q-phrase as a whole must be assigned a scope. Suppose scope over a restriction is achieved by iMerging *wh*-Q to the larger *wh*-Q-phrase it is embedded in where it was first merged. In what follows, I exploit the idea that extraction of *wh*-Q to the margin of the nominal in

which it is embedded insulates that nominal from Case assignment because the *wh*-Q acts as an intervener for Case assignment and is not itself a D that can receive Case by assignment. Consider the following assumptions:

(37) a. A *wh*-Q must locally c-command a restriction on *wh*-Q.
b. *Wh*-Q is typically embedded in the domain of D.

A tension now arises between (37a) and (37b), since (37b), a consequence of selection, usually, means that *wh*-Q be c-commanded by D while (37a) requires *wh*-Q to c-command D when a whole *wh*-Q-phrase is moved. Given (37a), we expect to see *wh*-Q outside of D in a nominal at the semantic interface, or else the *wh*-Q has not moved to form its restriction (normally resulting in an impossible interpretation). I shall assume that the same condition that applies to interrogatives to create scope assignment also applies to relative clauses, except that the domain is the open sentence formed as a property by iMerge of the *wh*-REL-phrase which is the first to c-command REL_{D}.

Since determiners and quantifiers tend to be mutually exclusive in English nominals, neither claim in (37) can be robustly supported by overt English phenomena. The only overtly moving *wh*-Q-phrases within nominals are instances where there is extraction of the *wh*-Q(-phrase) to the left edge of the pied-piped nominal, as in (38) (see Safir 1986: 679).

(38) Those reports which, the height of the lettering on, the government prescribes, are tedious.

Cases like (38) are relatively rare in overt syntax for reasons that will become apparent in section 7, but I will assume that extraction of just a quantifier head to form quantifier-restriction structures is the common case in covert (unpronounced) syntax. After all, one appeal of this analysis is that it directly feeds one of the most commonly employed representations of scope, namely, one where the quantifier has scope over its restriction and over the proposition that the restricted phrase originates in (i.e. the nuclear scope; see fn. 33). Moreover, it is an analysis consistent with a proposal independently made by Cable (2008), to which I will return.

Now let us consider the relevant steps in the derivation of the bracketed portion of (39), on the assumption that *wh*-Q is an intervener for Case assignment. The expansion of the tree can be tracked by the brackets on the right, and I pause after (39d) to take stock.

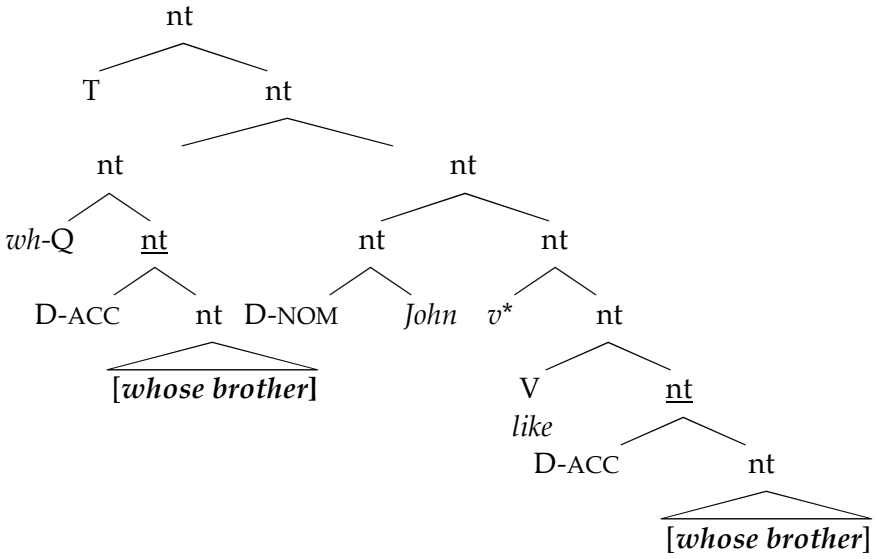
(39) I wonder [whose mother John likes]
a. [[D John] [*v** [like [D [whose mother]]]]]
– here the EA [D John] has just merged to set its domain as that of *v**
b. [[*whose mother*] [D John] [*v** [like [D [*whose mother*]]]]]]
– under assumptions dating back to Chomsky (1986), the A′-constituent must escape '*v*P' by adjunction (for the reason why in VSA, see section 6).

c. [[*wh*-Q [*D whose mother*]] [[D John] [v* [like [D [*whose mother*]]]]]]
– *wh*-Q subMerges its restriction and permits the moved phrase to be invisible for Case prominence

d. [T [[*wh*-Q [D *whose mother*]] [[D John] [v* [like [D [*whose mother*]]]]]]]
– T assigns Case prominence orders [D *John*] over [D *whose mother*] assigning R-Case (NOM) to *John* and ACC to [*D whose mother*], while T agrees with [D *John*](see (39d'))

The key move here is that in (39c), *wh*-Q, just the question quantifier, was extracted from [*whose mother*] and iMerged to the latter. This is the move that renders the displaced *wh*-Q-phrase invisible to T, so that the higher occurrence of *whose mother* is not ranked for Case prominence with respect to the lower occurrence of *whose mother*, with the result that Case Clash is avoided. The tree in (39d') clarifies the Case assignments.

(39) d'.



When *wh*-Q subMerges [D-ACC [*whose brother*]] before T is merged, a possibility allowed by Revised Extension, [D-ACC [*whose brother*]], sister to *wh*-Q, is then insulated from Case Clash. When T merges as in (39d), it orders only [D-NOM *John*] >> [D-ACC *whose brother*] for Case prominence, where the latter is sister to V, and no problem arises. If *wh*-Q does not subMerge its restriction at the first opportunity after movement, then Case Clash will end the derivation (as soon as cyclic Spell-Out applies, see section 6) because the occurrences of *whose brother* have more than one Case assignment between them.

The derivation then con-tinues with further movement of [*wh*-Q [D-ACC [*whose brother*]]] invisible to Case assignment in higher Case domains.

(39) e. [[*D John*] [T [*wh*-Q [D *whose mother*]] [[*D John*] [*v** [like [*whose mother*]]]]]]
– [D *John*] iMerges and satisfies EPP.

f. [C_{WH} [[*D John*] [T [*wh*-Q [D *whose mother*]] [[*D John*] [*v** [like [*whose mother*]]]]]]]

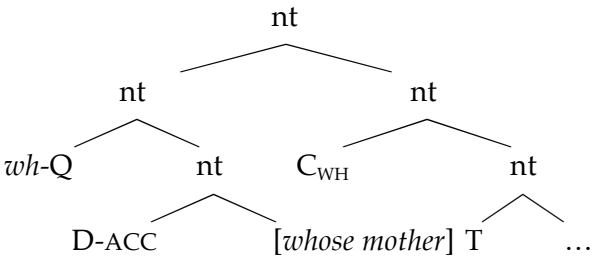
– Merge of C_{WH} assigns R-Case (NOM) to [D *John*], which can receive it without Case Clash, and C_{WH} sets WH_D

g. [[*wh*-Q [D *whose mother*]] [C_{WH} [[D John] [T [*whose* [D *whose mother*]] [D *John* [*v** [like [*whose mother*]]]]]]]

– the *wh*-phrase reaches its criterial position and is assigned WH_D as its scope

The top of the tree is diagramed in (39g'). If the derivation were to continue beyond the bracketed portion of (39), the Case domain of the matrix T would include the highest occurrence of [*wh*-Q [D *whose mother*]] in its criterial position in (39g), but once again, the *wh*-Q-phrase would be invisible to Case prominence.

(39) g'.



'Pronounce Highest' (see (51) in section 6 below) is the instruction for phonology, but for terminal nodes (that do not exclusively subMerge a terminal), 'Pronounce Lowest' is the order of the day (see (52) below). I will have more to say about how these principles apply in section 6 and, with respect to quantifier scope, in section 7. Thus the highest occurrence of [*wh*-Q [D-Acc *whose mother*]] is the highest non-terminal pronounced, and within that phrase, the second *wh*-Q morphologically embedded in *whose* is the one pronounced as the lowest occurrence of the terminal.

On these assumptions, simpler cases like *whom did John see* require an analysis of *whom* in situ before movement as [D [*wh*-Q *pro*]]. After [D [*wh*-Q *pro*]] moves to its criterial position or to any intermediate one, *wh*-Q will subMerge its restriction, thereby blocking Case from being assigned to D once a Case-assigning head is merged, as in (40).

(40) [H_{Case} … [*wh*-Q [D [*wh*-Q *pro*]]]]

Since D is silent in most languages preceding *wh*-Q and since Pronounce Lowest will apply, [D [*wh*-Q *pro*]] is likely to be heard only in its (highest) criterial position, since it is a higher occurrence of a non-terminal.[34] Notice that if a

[34] I am assuming that all lowest occurrences of a quantifier are treated as variables at LF by a rule that converts the copy into a variable (see Safir 2004b and Fox 2003). It is possible that further iMerge of *wh*-Q out of the constituent it forms with its restriction to form a new undominated node, as in (i), to allow the Q to have scope over both its restriction and its domain, but this depends on how the variable-forming rules are formulated.

particular Q is not an intervener for Case, then A'-movement cannot be protected, a consequence that will be explored in section 7.

To summarize, A'-movement of *wh*-Q-phrase is permitted to exist only when the phrase that moves is insulated from Case Clash by the intervention of a *wh*-Q that subMerges its restriction (and for further evidence based on intervention effects, see fn. 43 below). Phrases that are not insulated in this way will be susceptible to Case Clash if they move without R-Case.

Some potential support for this position might be gleaned from a slight reinterpretation of a proposal by Cable (2008), who explores a way to eliminate special mechanisms, like percolation, from playing a role in pied-piping phenomena. He is operating under the assumption that movement to a criterial position is triggered and that the goal probed by C_{WH} should only be a head that takes the 'pied-piped' constituent as its complement. He proposes that a Q-particle he identifies in Tlingit (and, drawing on earlier literature, for Japanese and Sinhala) takes the *wh*-phrase as its complement and it is the maximal projection of the Q-particle that is then moved. He extends this analysis to English, as in (41).

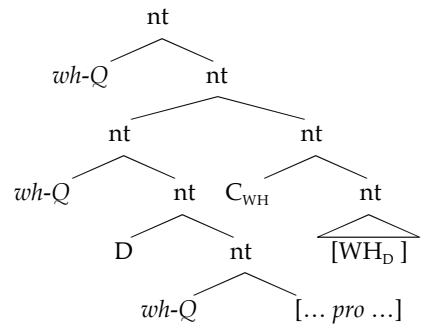
(41) a. Whose father's cousin's uncle did you meet at the party?
b. [**QP** [[[[**whose**] father's] cousin's] uncle] **Q**] did you meet at the party?

He continues (Cable 2008: 22):

> Under this analysis, a pied-piping structure in English is derived exactly like the pied-piping structures of Tlingit. In such sentences, the (null) Q-particle takes as sister a phrase properly containing the wh-word, which entails that the fronted phrase of the wh-question properly contains the wh-word.

The structure of Cable's analysis of the Q-particle is similar to the structure that emerges when the *wh*-Q in a *wh*-phrase is extracted to c-command its restriction, but is pronounced low. Except for the fact that the Q particle in Tlingit has distinct morphology from the wh-phrase (perhaps what permits it to be pronounced high), Cable's analysis suits VSA assumptions very well. The reason for the positioning of the Q-particle may either be an alternative Spell-Out of a *wh*-head, or it may be a head that functions, at least in part, to protect a *wh*-phrase from Case Clash in the course of cyclic movement. The second possibility is explored with respect to another phenomenon in the next section.[35]

(i)



[35] I will not enter further into Cable's analysis, where complications would surely arise. In particular, he notes that Tlingit pied-piping violates islands and pied-piping in English and

5.4. Differential Object Marking

This section closely follows Linares (2008), who demonstrates that VSA permits a much more natural account of differential object marking (DOM) than more conventional accounts based on special Case assignment or marking by a preposition. He proposes that DOM involves the introduction of a head to insulate a shifted direct object from Case Clash.

First, Linares describes the DOM phenomenon succinctly, and I repeat his description here (with examples renumbered to order with mine):

> The phenomenon known as Differential Object Marking (DOM) since Bossong (1985) involves languages in which direct objects appear in two different forms, depending on their intrinsic degree of specificity and/or animacy. In these languages, unspecific and/or inanimate direct objects (DOs) appear in an unmarked, nominative-like form, whereas specific and/or animate DOs are associated to a specialized particle or affix. Thus, for example, specific and animate direct objects in Spanish are associated to the particle *a*, often referred to as 'personal *a*'. By contrast, unspecific or inanimate DOs remain unmarked.
>
> (Linares 2008: 1)

(42) Veo *(**a**) la amiga de Pedro. *Spanish*
see.1ST A the friend of Pedro
'I see Pedro's (girl)friend.'

(43) Veo (***a**) una maquina.
see.1ST A a machine
'I see a machine.'

Linares points out that DOM is widely attested in the world's languages, including Sakha (Turkic), Hindi (Indo-European), Chaha (Semitic), and Miskitu (Misumalpan), many Romance languages, such Catalan and Romanian, as well as Spanish. Then Linares argues for the following generalization, which I name after him.

(44) *Linares' Generalization*

The marked version of a direct object in a DOM language is (also) a shifted object, while the unmarked version can (but need not) correspond to an object *in situ*.

For example, he shows that specific and/or animate direct objects are marked in Sakha, whereas unspecific and inanimate direct objects are not (46).

---

many other languages does not. He adapts a theory of wh-agreement from Kratzer & Shimoyama (2002) to make the distinction, but it seems that the extra assumptions that are involved in his account, including agreement limited by intervention, are essentially compatible with the Mapping Principle. I leave further exploration of this phenomenon for further work.

(45) a. Min Sardaana*(**–ni**) kördüm. *Sakha*
I Sardaana.ACC saw.1SG
'I saw Sardaana.' *(Sardaana: here a personal name, based on a flower name)*

b. Ali bir piano–**yu** kiralamak istiyor.
Ali one piano.ACC to-rent wants
'Ali wants to rent a certain piano.'

(46) a. Min sardaana kördüm.
I lily saw.1SG
'I saw a lily.' *(sardaana: denotes the flower in this case)*

b. Ali bir piano kiralamak istiyor.
Ali one piano to-rent wants
'Ali wants to rent a (non-specific) piano.'

(Vinokurova 1998)

Linares shows that unmarked objects must appear in strict preverbal position (47a–b), whereas ACC objects are placed to the left of a VP-adverb in the unmarked order (47c).

(47) a. Masha türgennik salamaat sie–te.
Masha quickly porridge eat.PAST.3
'Masha ate porridge quickly.'

b. * Masha salamaat türgennik sie–te.
Masha porridge quickly eat.PAST.3

c. Masha salamaat–y türgennik sie–te.
Masha porridge.ACC quickly eat.PAST.3
'Masha ate the porridge quickly.'

(Baker & Vinokurova 2008)

As Linares points out, however, it is not always easy to know if a differentially marked object has indeed undergone object shift (OS). Linares continues (p. 4):

> In Spanish, for example, in which word order is quite free, distributional tests fail to diagnose such short movements as OS. Indirect evidence of vacuous movement of marked DOs is nonetheless available, in control configurations involving gerundival adjuncts. In these contexts, marked objects can control PRO in secondary predicates, but unmarked objects fail to do so. As a result, [(48a)] is ambiguous, but [(48b)] is not.

(48) a. Besé$_i$ a una niña$_j$ [PRO$_{i/j}$ llorando]. *Spanish*
kissed.1ST A a girl crying
'I kissed a girl while I/she was crying.'

b. Besé$_i$ una niña$_j$ [PRO$_{i/*j}$ llorando].
kissed.1ST a girl crying
'I kissed a girl while I (*she) was crying.'

(Torrego 1998)

Further controls support his argument, but I leave these aside. If Linares' Generalization is correct, and for the sake of argument I will assume that it can be defended over the enormous range of cases that would have to be tested, then the question arises as to why such a correlation between OS and DOM should exist.

Linares proposes that the presence of DOM results from the presence of an intervening head that is inserted to avoid Case Clash when a specific/animate nominal is moved within the domain of T by OS. For example, the insertion of the Spanish *a* seems to play the role of insulating head, that is, a head that blocks Case assignment by T. The difference between DOM and *wh*-movement is that the movement is within the Case domain of T, assuming movement to a position that is perhaps higher than *v**, but still to the right of the overt position of the verb in Spanish. Still, by generating two occurrences within the Case domain of T, something must insulate the higher occurrence. Linares suggests that the differential object marker (DOMa) subMerges the shifted object in a derivation like (49) (showing just the relevant steps, where the external argument is *Joe*) and the resulting tree is presented in (49f').

(49) a. [*v** [see [D Mary]]]

b. [[D Joe] [*v** [see [D Mary]]]]

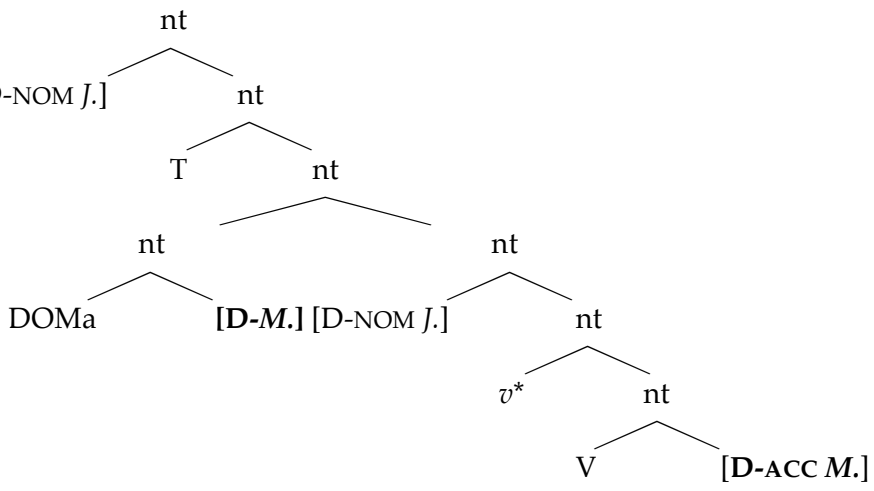
c. [[D Mary] [[D Joe] [*v** [see [D Mary]]]]]

d. [[**DOMa** [D Mary]] [[D Joe] [*v** [see [D Mary]]]]]

e. [T [[[**DOMa** [D Mary]] [[D-NOM Joe] [*v** [see [D-ACC Mary]]]]]]

f. [[D-NOM Joe] [T [[[**DOMa** [D Mary]] [[D-NOM Joe] [*v** [see [D-ACC Mary]]]]]]]

f'.



Step (49c) is OS. In (49d), [*D Mary*] is subMerged by the DOMa. Notice that (49d) does not increase the size of the tree, but by subMerging [D *Mary*], DOMa insulates it from Case assignment by T (by hypothesis). Thus T only ranks the external argument [D *Joe*] and below it, [D *Mary*], in its lowest, first-merged position, assigning [D-ACC *Mary*] and [D-NOM *Joe*]. The D of [D *Mary*] embedded in [DOMa [D *Mary*]] receives no Case, so it does not cause Case Clash with the D-ACC in direct object position.

Notice that it would not suit VSA architecture for the DOMa to be, itself, a

Case marker, or else Case Clash would result for [D *Mary*] insofar as the higher occurrence would receive two assignments. Linares (2008) argues that the DOMa is not a Case marker, although I will not review his arguments here. However, one may support his view with the contrast between the *a* of the DOMa and the *a* of Romance Dative marking. The Dative *a* argument always corresponds to a clitic in the Dative series, as seen in the double object configuration in (50a) and the causative structure in (50b). By contrast, the clitic corresponding to the DOMa argument is typically from the Accusative clitic paradigm, as in (50c) ((50a–c) from Carlo Linares, p.c.).

(50) a. Le dí una medulla a la gimnasta / a Michael.
CL.DAT gave.1SG a medal to the gymnast.FEM to Michael
'I gave the gymnast/Michael a medal.'

b. Le hice limpiar la piscine a Michael.
CL.DAT made clean the pool to Michael
'I made Michael clean the pool.'

c. Lo ví a Michael.
CL.ACC saw.1SG A Michael
'I saw Michael.'

It seems that the clitic paradigms can remain true to the Case of the DOMa argument, and this is to be expected if the DOMa is not, itself, a Case assigner or a Case marker.[36]

This account does not explain why specific and/or animate direct objects should have to undergo OS. A natural way to model OS in VSA would be to introduce a head below T that sets OS_D and then to treat the phrase that first c-commands OS_D as having OS_D as its scope. In other words, OS would be iMerge to a criterial position. It is not at all clear what the semantic value for this head would be, however, beyond just restating the semantic properties of the phrases that undergo OS, but in this respect, the analysis is no more stipulative than most other OS analyses. For OS, however, there is no obvious parallel to quantifier-restriction formation for *wh*-Q to insulate a second occurrence of a nominal under T from Case Clash. In such situations, I have adopted Linares' proposal that first

[36] A systematic aspect of the configuration in (i)-(ii) deserves further mention.

(i) [T [$_{CASED}$ … D …]]

(ii) [[*D* T] [$_{CASED}$ … *D* …]]

In (i) and (ii), T is a Case assigning head, so D in $Case_D$ of T will be subject to Case prominence and assignment by T, but D adjoined to T will not be, since it is outside $Case_D$. Now suppose cliticization is subMerge of D to T, and that T protects the D from Case Clash (from above, e.g., from C). For this to work it must be assumed that the clitic is a non-terminal, for example, [D *pro*], as has occasionally been proposed. This would appear to predict the Romance pattern of cliticization, where all of the nominal clitics associated with a verb surface in a clitic row on the highest tensed auxiliary or verb. The next likely target for cliticization would be C, which is also a Case assigner in VSA, and this pattern is also frequently found. Further assumptions are needed to make this work, but the matter deserves more study.

Merge of the DOMa insulates OS nominals from Case Clash.[37]

5.5. Summary

The upshot of this discussion is (i) that iMerge is optional in syntax, but (ii) that iMerge typically serves interpretive requirements associated with presence in criterial positions, and (iii) that movement must in general protect multiple occurrences of D-heads from Case Clash. The term A'-movement is now a merely descriptive term for iMerge of constituents that avoid Case Clash in a particular way, that is, by restriction formation, whereby a fronted *wh*-Q head intervenes to block Case assignment from above. OS appears to avoid Case Clash by subMerging a phrasal branch with a special head (DOMa) that also insulates the submerged branch. A-movements typically avoid Case Clash by bearing R-Case or non-structural Case, as discussed in section 4. There is no further syntactic distinction between these iMerge constructions that inherently defines them. No syntactic trigger is appealed to, no stipulated distinction between iMerge and eMerge, and no special appeal is made to 'pair-Merge', as in Chomsky (2004), to account for differences between adjuncts and elements moved to criterial positions.[38]

6. Spell-Out and Locality

The program so far has been to establish that movements that are treated as triggered by Agree in standard minimalist architecture arise in VSA by free Merge, indirectly driven by interpretation and constrained by Case Clash and other forms of derivational drag. In this section I introduce some of the locality restrictions that regulate long distance movement based on cyclic Spell-Out. A

---

[37] The distribution of multiple *wh*-interrogatives raise issues for VSA, only some of which get any further mention here. All 'multi-specifier' patterns are problematic for this theory because a domain can only be initially c-commanded once, and so multiple specifiers require further assumptions. Perhaps the second *wh*-phrase tucks in (e.g., as in Richards 1999) under the c-command domain of the first moved *wh*-phrase to permit local c-command of WH_D, or else the second-moved *wh*-phrase subMerges the first moved one, piggy-backing on its domain assignment. Extensions to allow for additional fronted positions are required in every current account, and, except for the VSA advantage with respect to the variety of permitted landing sites under Revised Extension. I leave investigation of these matters for future work, but see Appendix B.

Moreover, as Jonathan Bobaljik (p.c.) reminds me, domain assignment, as I have formulated it, does not apply to create domains for modal verbs, which can scopally interact with certain QPs. Similar remarks would extend to negation particles in some languages. In the literature on negation, a head+operator analysis has been proposed, for example, Haegeman & Zanuttini (1996) and Haegeman (1995). If I posit a null modal operator that can first c-command the domain M_D set by a modal M, nothing extra has to be said, but I have no evidence for such a claim. An extra statement about quantificational heads may turn out to be necessary, but this I also leave for further work.

[38] The discussion of criterial positions is incomplete, however, and will remain so, insofar as I have not provided any criterial motivation for scrambling. If scrambling is not criterial, it still must be insulated from Case Clash, and if it is, then it is necessary to posit heads that set scrambling domains. Although both strategies are plausible, I do not pursue the matter further here.

key aspect of the account is that forms that have not been moved soon enough can never be pronounced high, which is another form of derivational drag.

Fox & Pesetsky (2005), developing an idea from Chomsky (2000), suggest that when a certain point is reached in a derivation where cyclic node is merged, every lexical item below that node becomes unalterable with respect to crucial aspects of Spell-Out. Fox & Pesetsky couch this in terms of linearization, which is to say that the linear order of nodes below the relevant cyclic node cannot be altered by subsequent operations. For example, if a cyclic node is merged above A and B such that A is linearized before B (A >> B), and if a subsequent operation were to iMerge B such that a later cyclic node linearizes B >> A, a conflict arises that cannot be resolved, and the sentence crashes in phonology. Overt displacement is then impossible unless the theory allows for lower occurrences not to be pronounced (linearized) in certain contexts. If, however, iMerge precedes linearization, then the lower occurrence can be treated, in some relevant sense, as invisible to linearization. In what follows, I will adopt the essential mechanism just described, but I will not follow Fox & Pesetsky's theory in its specifics.

The version of cyclic Spell-Out developed here distinguishes movement by terminals from movement by non-terminals with respect to how they are treated by Spell-Out under cyclic nodes.

(51) Only higher occurrences of non-terminals are visible to linearization.

(52) Only lower occurrences of terminals are visible to linearization.

Only (52) is actually novel here. A node is a lower occurrence if there is a point in the derivation where it is c-commanded by a higher occurrence (unique occurrences are visible according to both (51) and (52)). (52) intrinsically bleeds (51), insofar as (51) will not apply between terminals embedded in non-terminal occurrences, since the terminal in the moved phrase will not c-command its copy in the position of the non-terminal 'trace'. However, if the phrase *which man* is iMerged such that it c-commands its first-merged position, then the first-merged position is a lower occurrence of a non-terminal and not visible to linearization. Further discussion of iMerge and Spell-Out of terminals (heads) is deferred to section 7.

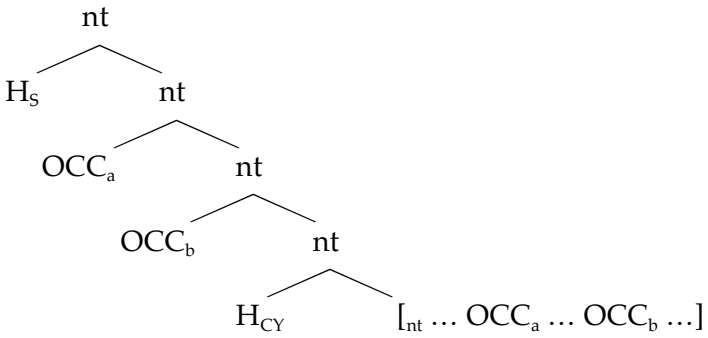
Now, let us assume that the cyclic nodes are C and *v*, which means that when cyclic Spell-Out applies, it applies to everything in the domains of these nodes. However, I do not assume that cyclic Spell-Out applies at the point when C or v is merged to its complement.

(53) a. Spell-Out below *v* only occurs when a head selecting *v* is merged.
b. Spell-Out below C only occurs when a head selecting C is merged or the derivation ends.

The practical import of this Spell-Out timing is that occurrences (OCCs) below the selectors but above the complement domains of the cyclic nodes (in the *intermediate zone*) are within the purview of (51) and (52). The illustrative indices on

the OCCs in (54) may be thought of as term indices, in the sense of fn. 28.

(54)



In (54), once the head (H_S) selecting the cyclic head (H_{CY}) is merged, (51) and (52) can determine which OCCs in the domain of H_{CY} are to be designated for pronunciation. If OCC_a is terminal, then its lower occurrence is linearized. If it is non-terminal, then the lower occurrence is not linearized. OCCs in the intermediate zone (IZ)—above H_{CY} but below H_S—are not evaluated for linearization until the next cycle. IMerge will always distinguish higher occurrences from lower ones when a tree is extended in the traditional sense, that is, when iMerge applies to the undominated node (but see the discussion of overt head movement in section 7). As shown in (54), there is nothing in the architecture that limits the number of occurrences that may be in the IZ.

The sort of movement described in (54) will result in well-formed structures (a) if the non-terminal is pronounced high, (b) if it reaches a criterial position that satisfies interpretive requirements, and (c) if it is properly insulated from Case Clash. Consider, however, what would happen if movement of a non-terminal did not stop in the IZ on its way to a criterial position.

(55) a. We wonder which boy John saw.
b. we [T [*v* [wonder [*which boy* [C_{WH} John [T [*v* [see *which boy*]]]]

In this derivation (irrelevant A-movement suppressed), the OCC of *which boy* in direct object position has been designated for pronunciation because merger of T above *v* has occurred without creation of more than one OCC of *which boy*. However, subsequent iMerge of *which boy* in its criterial position requires that the highest OCC be pronounced, with the result that more than one OCC of *which boy* is pronounced. Using the filter on linearization suggested by Nunes (1999, 2004) and others, locality is enforced by the need to avoid the linearization violation.

(56) No occurrence can precede itself in phonology.

I do not determine here how cyclic nodes determine what counts as a syntactic island, but limiting access to the IZ will result linearization/realization violations for subsequent extraction of phrases given (51). For example, if no OCC other than the one assigned REL_D is permitted to inhabit the IZ between a relative clause complementizer C_{REL} (a cyclic node) and the head that selects C_{REL}

(D or N, depending on other theoretical choices), then extraction of anything but the relative operator from below C_{REL} will result in a linearization violation, because the lowest OCC of any phrase below the cyclic node will have already been designated for pronunciation. The Merge operation is not intrinsically local; locality is enforced by cyclic linearization (and concomitant realization). Such an approach to islands would be compatible with VSA, but it cannot be explored here for reasons of space.

On this account, phrasal movement is always pronounced high, and so there is no space in this theory for covert phrasal movement, that is, phrasal movement pronounced low, as discussed in Appendix B. Covert head movement is discussed at length in section 7.

Finally, a remark is in order about whether or not the assumptions surrounding linearization are consistent with the logic of the KF and VSA. Spell-Out is triggered in this system by the introduction of a head that activates the cyclic node. Since this latter activation is a c-command relation that determines that another c-command-determined domain undergoes an operation outside of syntax (linearization), the only stipulations relevant to KF reasoning are (i) that certain linguistic labels and not others are the ones that set Spell-Out domains, and (ii) that the nodes that select for these labels are the ones that trigger linearization. One might argue that there must be cyclic nodes to limit demand on working memory, for example, but the stipulation as to which nodes are cyclic for Spell-Out is (so far) not derived from the KF and its interaction with the interfaces.[39]

7. Head Movement and LF Interpretation

Now let us return to the question of head movement, which, according to (52), requires that only its lower occurrence be pronounced. First I will show that head movement to a criterial position, where it iMerges to a non-terminal, need not be cyclic, and therefore is insensitive to cyclic effects like island restrictions. Second, while *wh*-Q is a case intervener, and thus enables A'-movement of the *wh*-Q-phrase, quantifiers that do not block Case will never permit phrasal movement, only head movement, as is the case for most quantifiers. Furthermore, head movement to a criterial position moves without its restriction by definition, and this turns out to be what makes it sensitive to intervention effects, in comparison with overt phrasal movement. Finally, overt head movement is also possible, but only when (52) does not apply. Such cases arise if heads adjoin to heads, such that the iMerged head never c-commands its 'lower' occurrence. Thus this systematic theory of head movement derives many key features that distinguish '*in situ*' quantifiers from the overtly moved sort.

First, consider the fact that the (52) insures that head movement (movement of a terminal) to a position where it c-commands its copy will always require that its lower copy be the pronounced copy. Now consider what would

[39] Bošković (2007) contends that all phrasal nodes are cyclic, and if so, a version of VSA empirically compatible with this broader application would come closer to satisfying the burden of the KF, but I have not explored this possibility for reasons of space.

happen if *wh*-Q were to move just as phrases do, first to an intermediate position, and then to a criterial position, as in (57).

(57) a. [T … [*wh*-Q [EA [*v** [V [D … *wh*-Q …]]]]]]
b. [*wh*-Q [C_{WH} [… T … [*wh*-Q [EA [*v** [V [D … *wh*-Q …]]]]]]]]

Under at least one interpretation of 'Pronounce lowest occurrence', head movement will always be restricted to swoop movement in this account, that is, a single movement from first merge position to its criterial position. If head movement (to a c-commanding position) were cyclic, then even after the first merged position is linearized, higher occurrences of the head would still be lower than subsequent head movements, with the result that 'Pronounce lowest occurrence' would pronounce the intermediate head, with a resulting linearization violation with respect to the lowest occurrence. Thus the lowest terminal node copy will be spelled out in the first linearized cycle that contains it. But the empirical results of the theory of cyclicity in the last section, i.e. all the locality restrictions, arise because Spell-Out must be avoided when it is too low. The game is over before it begins for lowest terminals, given (52). Nothing prevents a terminal from moving any distance at all as long as it ends in a criterial position. Thus the theory makes the prediction in (58):

(58) Movement of quantifier heads is
a. pronounced low (*in situ*),
b. insensitive to islands, and
c. moves without intermediate stops.

These results are largely consistent with what has been observed about *wh-in-situ* languages (e.g., Huang 1982 on island effects), although the usual puzzle concerning the clause-boundedness of most non-*wh* quantifiers remains underived and unresolved (as it is for most current theories), apart from assumptions about Beghelli & Stowell's (1997) clausal architecture discussed in the next section. Whether or not (58c) is correct would depend on showing that *in situ* quantifiers never have intermediate scope, a prediction that depends on too many other assumptions to be examined further here.

Now recall that I have also assumed that *wh*-Q acts as an intervener for Case assignment. This intervention prevents Case assignment to *wh*-phrases in intermediate positions, thereby avoiding Case Clash, thanks to restriction formation involving subMerge on the left branch. This configuration was illustrated in (40), repeated here.

(40) [H_{Case} … [*wh*-Q [D [*wh*-Q *pro*]]]]

Suppose, however, that only *wh*-Q is an intervener for Case and other Q heads are not. In this system, there are three immediate consequences for Q heads that are not Case-interveners (call them 'Q_O'). First, restriction formation in a left branch will not prevent Case assignment from applying across Q_O to D, creating a Case Clash if there is more than one occurrence containing D in $Case_D$. This

entails that phrasal movement to a criterial position for such quantified nominals is normally impossible, since the D in the higher copy (bolded) will get a different Case assignment from that of the lower copy.

(59) *[H_{Case} … **[[Q_O [D [Q_O *pro*]]]]**] H_o[… [D [Q_O *pro*] …]]]

Thus all Q_O must find their criterial positions by head movement, never by phrasal movement, and be pronounced low by (52). These are the cases known in the literature as 'covert movement' and they will be further discussed in section 7.2.

The immediate empirical consequence for quantifiers that are pronounced low is that their restrictions are never high. As a result, Principle C effects are predicted in cases like (60), where the quantifiers in the Beghelli & Stowell classification (see section 7.1) would find criterial positions above *v*P by head movement, but direct objects would still c-command the restrictions on the quantifiers (e.g., *books that criticized Noonan* and *book about Mary* in (60a–b), respectively).

(60) a. * Sheelagh gave *him* [many books that criticized *Noonan*].
b. * Arthur sent *her* [every book about *Mary*].
c. *Richard finally told *them* about three critiques of *the teachers*.
d. Which critique of *the teachers* did Richard finally tell *them* about?

Judgements about the success of coreference in (60d) are not uniform, but even those who reject the overt phrasal movement example in (60d) typically report it as less deviant than the *in situ* cases in (60a–c) (see Safir 1999). The difference is that overt movement has moved its restriction out of the c-command domain of the pronoun, and if the copied restriction is interpreted high, then it will not induce c-command effects. For (60a–c), which in this theory are predicted to be derived by head movement (since the quantifier is pronounced low), the restriction never moves with the Q_O, hence it can only be interpreted low, predicting the Principle C effects.

Now consider contexts where heads appear to have overtly moved, as, for example, in V2 constructions and in constructions involving head-to-head movement such as T to C. We have just established that head movement to the undominated node in this system cannot be overt because of (52), which will insure that the lower occurrence of the head will always be favored for pronunciation. However, it is possible for head movement to be overt if there is no way to determine which of two occurrences is higher. This situation will arise when head movement proceeds by submerging a head without submerging its sister, that is, in contexts of head-to-head movement.

(61) a. [H_1 [… H_2 …]]
b. [[H_2 H_1] [… H_2 …]]

Although head movement resulting in (61b) is the most widely assumed analysis of the overt displacement of heads, only Revised Extension permits such

structures to arise from (61a). In (61a), H_1 is higher than H_2 because H_1 c-commands H_2, but in (61b), neither occurrence of H_2 is higher than the other because there is no c-command relation between the two occurrences at any point in the derivation. When this occurs, the system does not predict which node is pronounced, but something must insure that only one of the nodes is linearized. Although I will treat cases like (61b) as the source of overt head movement structures, I regard such cases as outside the core of the linearization theory (embodied in (51-2)), and hence a matter for particular morphologies, with a certain amount of linguistic variation as a result.[40]

7.1. Domain-Setting for Scope

One of the major contentions built into VSA is that all prominence relations are keyed to domain-setting and domain assignment in the course of a derivation and this has been illustrated for Case, argument structure and agreement. Domain assignment for *wh*-Q-phrases has also been instantiated as first c-command of a domain set by C_{WH} to license the interpretation of a *wh*-Q-phrase. Scope interaction can be established by the same mechanisms.

Suppose, for example, that the domain Q_D for a certain quantifier Q is set by a functional head specific to that (class of) quantifier, as in Beghelli & Stowell (1997). The domain of a quantifier is assigned in the same way that domain assignment applies generally, i.e. like external argument assignment, and assignment of *wh*-scope as in (34) is a sub-case of (62)).

(62) *Assigning Scope*

The scope of a quantifier Q is assigned when Q is assigned a domain that matches its quantificational features/properties.

---

[40] I am avoiding discussion of the voluminous literature on head-to-head movement here. I concur with Matushansky (2006) (a useful re–assessment of the issues) that the movement is not phonological, but I do not regard it as different in nature from phrasal movement insofar as both can involve subMerge, given Revised Extension. Thus I do not appeal to a special morphological merger rule that must apply to deform the tree after normal iMerge, as she proposes. Nothing in my theory predicts that head-to-head movement must be local, although linking it to intervention for c-selection, as Matushansky does, would suit VSA for movement of heads to positions that are not criterial, yet c-command issues remain. Like Matushansky, I abstract away from some aspects of word internal structure such that what I am calling a head is the node that bears a label under which further morphological analysis is hidden from Merge in syntax, but I have no principled account of that divide. Most distinctly from Matushansky, I do not distinguish head vs. phrasal movements by their triggers, since there are no operational triggers in VSA. It could be that there is morphological longing for affixation that head-to-head movement satisfies or fails to for particular morphologies, but this again leads us into assumptions about particular morphologies and the phonological nature of affixes.

With respect to situations where more than one copy is pronounced, Nunes (2004) appeals to morphological reanalysis in cases where atypical Spell-Out results, as in certain cases of copy doubling. It is also possible that the lack of c-command between copies plays a role in conditioning these atypical outcomes. Sideward movement of any kind would be ruled out by (56) if there is no way to reduce pronounced occurrences to just one where (51–52) do not apply.

Note that on this account, a quantifier can only be ranked for scope after it has moved to a scopal position (not *in situ*), a point to which I will return. Relative scope arises when the domain of one quantifier is in the domain of another, as in (63), where $Q_1 >> Q_2$ because the domain of Q_2 is inside the domain of Q_1.

(63) $[Q_1 [H_1 [Q_{D1} \ldots Q_2 [H_2 [Q_{D2} \ldots]]]]]$

The quantifier last to reach its criterial scopal position will thus have widest scope.

Notice that iMerge of Q_1 cannot fall short or bypass the position where its scope domain is assigned, or interpretation fails. Similarly, interpretation fails if Q_1 or Q_2 is assigned a domain set by a head that is inappropriate for it (i.e. this could be thought of as the VSA analog of an interpretable feature that is not interpreted). For example, if H is interrogative C_{WH} which sets WH_D, then iMerge of *every* to $[C_{WH}\ WH_D]$ will mean that *every* is assigned WH_D. An interpretive mismatch results, since *every* is not interrogative.

(64) # $[$*every* $[C_{WH} [_{WH_D} \ldots$ *every* man $\ldots]]]$

The mismatch in question is only at the level of interpretation, however; the syntactic structure in (64) is well-formed.

It must be noted, however, that the account offered so far does not insure that a Case-intervening Q (like English *wh*-Q) will achieve its criterial position by overt phrasal movement—head movement by *wh*-Q should also be possible. For languages like English, where at least one overt wh-phrase must be iMerged, an additional stipulation must be made, namely, (65).

(65) WH_D must be assigned to a non-terminal.

(65) resembles the EPP in (20) (and is no more explanatory), but while the EPP may be universal (accounts differ), it is clear that (65) is not.[41] As long as (65) is satisfied, subsequent *wh*-movement, as in multiple interrogations in English, can be by head movement. Unlike iMerge trigger theories, however, this approach is consistent with Rizzi's (1990: 46–48, 2006) suggestion that *why* is generated directly in its criterial position, without the application of Agree to activate an unvalued feature in its domain, consistent with his approach to criterial assignment in Rizzi (2006). With respect to other language specific stipulations, see Appendix B.

7.2. Scope Ambiguity

As mentioned, VSA is consistent with clausal architectures that generate heads associated with scopal positions, as in Beghelli & Stowell (1997) (henceforth, B&S) and references cited there. B&S provide a typology of quantifiers and they

[41] Jonathan Bobaljik (p.c.) points out that first position in V2 languages may require a language-class specific EPP-like statement as well.

suggest that the scopal positions that a quantifier can be interpreted in are determined by the type of quantifier it is. In their way of putting things, the quantifier Q must be found in the specifier of the head that provides a domain Q is licensed to have scope over. This model lends itself to VSA architecture, in which a head sets the domain and the first c-commanding phrase merged above it is assigned the domain-scope by (62).

Prominence for scope will arise in VSA when the domain of one quantifier contains the domain of another one. If neither of two scope-interacting Q heads are interveners for Case, then both will be assigned scope after head movement to a criterial position and both will be pronounced low. However, in the B&S architecture, some quantifiers have more than one criterial position, and when a given quantifier can move to a criterial position either above or below some second quantifier, then scopal ambiguity will arise, since the derivational outputs will allow two different structures for a string. The output of phonology will not distinguish the two structures in this respect, though their structures lead to distinct interpretations. This is the source of scope ambiguity in VSA.

To get a practical sense of how VSA represents relative prominence for scope, consider the two scope possibilities permitted by (66).

(66) Two soldiers praised every general.

When *two soldiers* is merged to *praise every general* to form *two soldiers praise every general*, *two soldiers* is the more prominent argument of *praise* because *two soldiers* is assigned external argument status when it first c-commands the selected domain of v^*, v^*_D, and v^*_D contains *every general*. This argument prominence relation, that is, *two soldiers >> every general* for *praise* is thus fixed before any further operation, and no subsequent Merge operation can reverse any prominence relation once it is set. If the scope-domain-setting (SDS) head H_1 of the right type is merged to the constituent containing the full argument structure for *praise*, then for the quantifier *every* to receive the proper scope assignment, iMerge must apply to the terminal *every* to yield (67a). This movement is optional in syntax, but if it does not occur then the quantifier *every* will not be assigned a proper scopal interpretation.

(67) a. [*every* [H_1 [two soldiers [v^* [praise *every* general]]]]]
b. [*two* [H_2 … [*every* [H_1 [*two soldiers* [v^* [*praise every general*]]]]]]]

When H_1 is merged it sets the contents of its sister, H_{1D}, as its domain, and when *every* is iMerged so that it first c-commands H_{1D}, *every* is assigned H_{1D}, as it is in (67a). This means that *every* has scope over every other quantifier assigned a domain that is contained in H_{1D} (up to intervention), but in (67a), no such assignment has been made (*two* has not been assigned a scope yet). Now *two* must be assigned a scope if it is to be interpretable. Just as with *every* of *every general* and for quantifiers in general, *two* of *two soldiers* can receive an interpretation only if *two* can be assigned the right sort of domain. So if the right sort of scope-domain setting head H_2 is merged above the higher occurrence of *every*. Then iMerge of *two* will assign to it $H_{2D,}$ the domain set by H_2. Since H_{2D}

contains *every* and the domain that has been assigned to it, scope is established as *two* >> *every*.

Once scope prominence has been assigned in this way it cannot be reversed later in the derivation, or Scope conflict arises, and no scopal interpretation is possible (i.e. another form of incoherence). In other words, relative scope assignment becomes a form of derivational drag like Case Clash and linearization paradoxes. If *two* had moved to a criterial position first, and *every* after, then the prominence relation would have been reversed.[42]

What has been schematically presented in (67) would be instantiated in the B&S system where the head corresponding to H_1 is Dist (the head just below T that sets the domain for 'distributed' quantifiers like *every*) and the head corresponding to H_2 is Ref (one possible landing site for 'group' quantifiers, like *two soldiers*). The interpretation for (66) with wide scope for the universal would be achieved by merging a SDS head, Share (= H_1), above *v**, where Share would set its domain ($Share_D$). When *two* is iMerged above Share, as in (68a), *two* first c-commands $Share_D$, thereby assigning the scope domain for *two* as $Share_D$.

(68) a. [*two* [Share [*two soldiers* [*v** [*praise every general*]]]]]
b. [*every* [Dist [*two* [Share [*two soldiers* [*v** [*praise every general*]]]]]]]

Every gets its scope in the same manner that it did in (67a), that is, Dist is merged to the constituent in (67a), such that $Dist_D$ is set, and then $Dist_D$ is the scope domain assigned to *every*. Since Dist >> Share (universally, in the B&S theory), *every* has scope over *two*.

The key to the ambiguity of scope in the B&S analysis is that group quantifiers have more than one landing site because they can generally be associated with more than one scope domain, whereas distributed quantifiers are not so flexible. The derivations in (67) and (68) assume that the first quantified phrase that first c-commands a domain set by the head appropriate to it, will end up having the narrower scope. Extension will conspire with iMerge to insure that any other quantified phrase will have to be Merged to a criterial position above the first one. Any theory that distinguishes quantifiers in this way is suitable for instantiation in VSA, and so I will not further explore the particulars of the B&S system.

7.3. Scope and Intervention

In other sections I have discussed a variety of intervention effects where domain-setting heads are interveners for other domain setting heads of the same kind: Case assigning heads and Q are interveners for other Case assigning heads, selecting heads for other selectors, agreeing heads of one sort or another are

[42] Actually, B&S assume that 'group' quantifiers, of which *two soldiers* is an exemplar, can be interpreted in situ. If so, then there is an interpretation for (67a) for which *every general* has scope over *two soldiers* without movement of *two*. I have nothing to say about in situ interpretation of certain classes of quantifiers, but it is less natural in my system. I will assume for the sake of presentation that all quantifiers must establish scope by means of a domain set by a compatible head.

interveners for heads of that sort, and cyclic heads intervene to render their domains opaque to linearization by higher cyclic heads. This section explores intervention effects that arise when certain *in situ* quantifiers are in the domain of certain intervening quantificational heads. Given the approach to the overt/ covert displacement developed here, scope intervention effects typically arise where the lowest occurrence of a scope-bearing head is separated from its highest occurrence by an intervener (typically a certain class of SDS head). Beck & Kim (1997: 370) state the relevant effect as follows:

(69) a. *Quantifier Induced Barrier*

The first node that dominates a quantifier, its restriction, and its nuclear scope is a Quantifier Induced Barrier.

b. *Minimal Quantified Structure Constraint*

If an LF trace β is dominated by a quantifier induced barrier α, then the binder of β must also be dominated by α.

Following (and interpreting) the literature, it will be argued that pied-piping is a means of 'smuggling' a lowest Q-head occurrence past an intervener, thereby neutralizing the intervention effect, given a plausible treatment of the relation between copy theory and variable formation (see fn. 33 above).

Pied-piped constituents, since they are phrasal, will always be pronounced high in VSA, so this theory is designed to predict that intervention effects are only found in cases of head movement to a criterial position, where the lower occurrence is pronounced, as is generally the case for intervention effects (but see (77-78)).

Consider the following intervention effects that have been noted for Korean by Beck & Kim (1997) for (70)–(72) and by Kim (forthcoming) for (73)–(74).

(70) a. * Amuto muô–s–ûl ilk-chi anh–ass–ni? *Korean*
anyone what.ACC read-CHI not-do-PST-Q

b. Muô–s–ûl amuto __ ilk–chi anh–ass–ni?
what-ACC anyone read-CHI not-do-PST-Q
'What did no one read?'

(71) a. *Minsu–man nuku–lûl po–ass–ni?
Minsu-only who-ACC see-PST-Q

b. Nuku–lûl Minsu–man __ po–ass–ni?
who-ACC Minsu-only see-PST-Q
'Who did only Minsu see?'

(72) a. *Minsu–to nuku–lûl po–ass–ni?
Minsu-also who-ACC see-PST-Q

b. Nuku–lûl Minsu–to __ po–ass–ni?
who-ACC Minsu-also see-PST-Q
'Who did Minsu, too, see?'

(73) a. ?*Nwukwun–ka–ka mwues–ul ilk–ess–ni?
who-some-NOM *what-ACC* *read-PST-Q*

b. Mwues–ul nwukwun–ka–ka __ ilk–ess–ni?
what-ACC *who-some-NOM* *read-PST-Q*
'What did someone read?'

(74) a. *JOHN–I mwues–ul sa–ss–ni?
John-NOM *what-ACC* *buy-PST-Q*

b. Mwues–ul JOHN–I sa–ss–ni?
what-ACC *John-NOM* __ *buy-PST-Q*
'What did JOHN (but not someone else) buy?'

In all of these cases, the a-examples show that an intervener blocks wide scope interpretation for the in situ quantifier and the b-examples show that wide scope is available for the same quantifier if it moves overtly to a position higher than the intervener. Negation intervenes for an in situ question in (70a), but scrambling of the *wh*-phrase to initial position in (70b) is acceptable. The interveners for (71)–(74) are 'only', 'also', existential-indefinite *-ka* marking, and contrastive focus, respectively, and in each case, if the *wh*-phrase is scrambled to initial position, as in all the (b) examples, the intervention effect disappears. The contrast is schematically modeled for (70a–b) in VSA in (75a–b) where Neg (underlined) is the intervener for Q.

(75) a. * [Q [C_Q [$_{Dq}$ … [[Neg-XP] [Neg [$_{Dneg}$ …[D [Q N]] …]]]]]]
b. [[Q [D [Q N]]] [C_Q [$_{Dq}$ … [[Neg-XP] [Neg [$_{Dneg}$ … [D [Q N]] …]]]]]]

Notice that the negative quantifier phrase is assigned scope in the usual way according to VSA, that is, an SDS head (Neg in (75a)) has set a domain, Neg_D, that has been assigned to the quantifier or quantified phrase (for a clausal architecture for Korean compatible with these assumptions, see Kim, forthcoming).

Three points are of particular significance here with respect to the contrast between the a- and b-examples in (70)–(74), respectively. First, the a-examples involve direct head movement to a criterial position, leaving a lower occurrence in the form of a head, while the b-examples involve overt movement of a phrase (with head movement within the phrase of the Case-clash-insulating head Q). 'Smuggling' movement in (75b) does not have to be directly to a criterial position, as long as it is outside of Neg_D. Second, the rest of the quantified phrase in the a-examples, the part that would count as its restriction if the quantifier were moved, does not have an occurrence above the intervener. Third, the intervener, underlined in (75a–b), is an intervening SDS head that has set a domain for some other quantifier; the quantifier associated with the intervener is not the actual intervener in this system. Interpreting proposals in the literature (e.g., Pesetsky 2000: 67) in terms of VSA, I describe the effect as follows:

(76) The occurrence of Q in its criterial position cannot be separated from at least one occurrence of its restriction by an intervening SDS head of type Y.

The reason, then, that the b-examples in (70)–(74) permit wide scope is because phrasal movement has moved Q's restriction higher than the intervening SDS head, so interpretation has access to the restriction of Q.[43]

It is important to keep in mind that not all intervention effects involve the failure of wide scope readings for *in situ* quantifiers.[44] Overt movement that strands part of the restriction of a quantifier also falls under the generalization in (76). Indeed the first known effects of this kind, such as that reported for French by Obenauer (1984) in (77) and that reported for German by Beck (1996) in (78) (from Pesetsky 2000: 68) involve overt movement of a portion of the quantified phrase that may or may not be an instance of movement by a head.

(77) a. Combien de livres a–t–il beaucoup consultés __?
how-many of books did-he a-lot consult
'How many books did he consult a lot?'

b. Combien a–t–il (*beaucoup) consultés [__ de livres]?
how-many did-he a-lot consult of books
'How many books did he consult a lot?'

(Obenauer 1984)

(78) a. [Wen von dem Musikern] hat keine Studentin __ getroffen?
whom of the musicians has no student met
'Who among the musicians has no student met?'

b. * Wen hat keine Studentin [__ von dem Musikern] getroffen?
whom has no student of the musicians met
'Who among the musicians has no student met?'

(Beck 1996, Pesetsky 2000: 68)

The movements in (77b) and (78b) are exceptions to (52) if they are head movements pronounced high, but they are consistent with (51) and (52) if they are partial phrasal movements with part of the moved phrase silent. Even in the latter case, (76) applies to rule out (77a) and (78a), but not (77b) and (78b).

Much of what I have said so far is a translation of scope intervention into VSA, but the correlation proposed here between quantifier head movement, pronunciation low (*in situ*) and the potential absence of island effects makes a prediction within VSA that is distinct from theories that rely only on the *Agree*

---

[43] I have identified the SDS head as the intervener rather than the domain of that head or the quantifier that is assigned that domain because the SDS head, as a terminal node, has a label, but I have no evidence that slightly more complex statements in terms of intervening quantifiers (or quantified phrases) or containing domains (perhaps derived as a semantic consequence) could not achieve the same result. The application of (76) depends on choices for Q and Y, and in the absence of a way to predict which interveners block which quantifiers (i.e. no strong adherence to the Natural Intervention Hypothesis), I will not further explore this matter here (but see Appendix B).

[44] Sauerland & Heck (2003) show that there are intervention effects internal to pied-piped phrases when what I am calling restriction formation needs to take place. They show that overt movement within the pied-piped phrase or the avoidance of pied-piping altogether, results in well-formedness. In other words, intervention effects place a limitation on the embedding of the scope-bearing head in pied-piped phrases. See their paper for details.

relation to predict locality. Even when an *in situ* quantifier can be construed outside an island in the absence of an intervener, the presence of an intervener outside of an island is still predicted to cause an intervention effect. Consider the Korean examples in (79) (provided by Hyunjoo Kim, p.c.).

(79) a. Mina–ka [[nwukwu–lul al–nun] salam]–ul po–ass–ni?
Mina-NOM who-ACC know-REL person-ACC see-PST-Q
'Mina saw the man who knows who?'

b. *??Mina–man [[nwukwu–lul al–nun] salam]–ul po–ass–ni?
Mina-ONLY who-ACC know-REL person-ACC see-PST-Q
'Only Mina saw the man who knows who?'

c. *??MINA–KA [[nwukwu–lul al–nun] salam]–ul po–ass–ni?
MINA-KA who-ACC know-REL person-ACC see-PST-Q
'Mina (and no one else) saw the man who knows who?'

The *in-situ wh*-quantifier in (79a) is embedded in a relative clause (marked by brackets), from which overt movement is not possible, yet (79a) is successfully construed as a non-echo direct question. In (79b), however, the presence of the intervener hosting *Mina-man* outside of the island, but below the high scope position required for a matrix interrogative, results in an intervention effect. In (79c), it is the contrastive subject that causes intervention. The persistence of the quantificational intervention effect where cyclic movement is not possible suggests that *intervention applies over a potentially unbounded domain*, that is, no matter how large a span a single iMerge head movement can cover, intervention will hold across that potentially unbounded distance. If Agree were responsible for the quantificational intervention effect in (79), then the assumption that Agree is bounded by a cyclic domain must be abandoned. However, such a move is tantamount to conceding that an interpretive relation R stated on a c-command relation is the basic sort of viable relation, and that bounded c-command relations, including those that involve Case and agreement, simply involve more local interventions in most instances. Further inadequacies of Agree as an alternative to c-command are discussed in Appendix A.

7.4. Summary

The approach to head movement developed in this section aligns it with what has been called covert movement or LF-movement in the literature, but not with covert *phrasal* movement. In VSA, head movement to a c-commanding position is always pronounced low and typically is movement to a criterial position establishing scope. Overt head movement arises when one head subMerges another head, an outcome that is not regulated by (52), since the iMerged occurrence does not c-command the 'lower' one. Head movement to a c-commanding position is unbounded movement insensitive to islands, as opposed to overt movement, because islands, insofar as they are understood, are dealt with by cyclic Spell-Out, not by quantifier intervention. The inapplicability of cyclicity to head movement is principled, because cyclicity only leads to violations that arise from lineari-

zation conflicts, and movement uniformly pronounced low will never cause linearization conflicts.

Head movement to a criterial position strands its restriction, which is then unavoidably susceptible to Principle C effects and intervention effects, superiority perhaps being one of the latter. Intervention effects are neutralized by overt movement above the intervener because overt movement is necessarily phrasal (by (51)). Phrasal movement to a criterial position is typically a form of pied-piping in this system, and it is always an option if Case Conflict or interpretation of quantifiers or linearization restrictions do not rule it out. An EPP-like feature in C can have the effect of forcing movement of a non-terminal, as in English questions, but how large a non-terminal moves is not determined by any device. The latter result leaves 'heavy pied-piping', such as pied-piping of more than a nominal, unexplained, as it is in other accounts—but unlike Spec-Head feature checking theories, no ad hoc feature-percolation is needed to justify it (movement is always optional) and indeed, no Spec-Head checking can even be formulated in VSA.

8. Conclusion

If VSA can be feasibly maintained, then the wide range of unselected-for structural complexity that occurs in natural language can be traced to interaction of two simple factors, pre-HLF capacities, including Merge, and the Mapping Principle. The case for the existence of a KF is thus more plausible, and even if the Mapping Principle proves insufficient, it has been part of the goal of this essay to demonstrate how high the bar must be set if we are to meet the burden of the KF.

I have argued, however, that Merge and interface interpretive relations mapped on to c-command, interacting with the properties of lexical items, inherent prominence relations, and relative closeness, are indeed sufficient to generate the complexity we observe in natural language syntax. Attempts to streamline the theory so that only these notions were appealed to required solutions to prominence conflicts without appealing to operational distinctions between iMerge and eMerge, or labels projected to non-terminal nodes, or Spec-Head feature-checking or numerations.

Instead, I have featured the role of c-command as the template for all mapping relations interpreted at the interfaces. Heads set domains by c-commanding into them up to intervention, sometimes by necessarily selecting heads in domains (c-selection), sometimes by interacting with other heads in their domain if there are any (agreement), sometimes by prominence ordering of elements in their domains (Case prominence), and sometimes by setting the domains necessary for criterial assignments, as in the case of scope assignment. The special role of unique specifiers in other theories has essentially been derived by first c-command (as in domain assignment for scope) which is a continual source of derivational drag. A further role for c-command is post-derivational, as illustrated by the Korean island-insensitive intervention effects, which also demonstrate the insufficiency of Agree as an alternative addition to Merge. Further

evidence that Agree cannot replace appeal to c-command for the statement of structurally sensitive interpretive relations is provided in Appendix A.

As VSA emerged, I did appeal to special notions regulating Case (for example, (18)) and Spell-Out (for example, (51-52)) that do not follow from Merge and c-command mapping, but that do have to do with the way in which the output of Merge is evaluated at the phonological interface. The theory of interpretation in criterial positions does not appear to need anything novel in this account, except revising Extension to allow instances of subMerge on an immediate daughter of the undominated node, a revision that is independently necessary for overt head movement, tucking in, and late attachment. Failures of (appropriate) scope assignment also result in ill-formedness, where failure of interpretations is a filter on derivations.

Most of the component parts of VSA have been adapted from existing proposals and analyses, and thus rely heavily on generalizations and mechanisms that others have explored, including the prominence theory of Case assignment (particularly Marantz 1991 and Bobalijk 2006), the theory of criterial positions (particularly Rizzi 2006), the theory of cyclic linearization (particularly Chomsky 2000 and Fox & Pesetsky 2005), the nature of scope assignment (Beghelli & Stowell 1996) intervention (particularly Beck & Kim 1997 and Pesetsky 2000). The Mapping Principle, however, is what makes sense of why these components have the form that they do and how they interact. The interactions have produced some novel results, among them the dissolution of any theoretical A/A'-distinction, which is now only an artifact of the strategy by which Case Clash is avoided, and an explanation of why the scope of *in situ* quantifiers is acyclic, unless intervention effects on heads or restrictions keep it local. All of these components and results have been sewn together as manifestations of the role of the Mapping Principle in the structure of linguistic architecture.

Many issues not touched on here will have to be explored if VSA is to succeed as a leading idea informing the relationship between evolutionary boundary conditions on the emergence of human cognition, on the one hand, and the rich world of linguistic structure and diversity, on the other. This essay is only the first step.

Appendix A: C-Command vs. Operational Agree

In recent work, it has been suggested that c-command may not need to be independently stated in linguistic theory because all of the cases for which it is required can be handled by operational Agree (e.g., Chomsky 2007a, 2008). Note that Agree is assumed in these works to be independently necessary in addition to Merge, and thus positing Agree is, for those who believe that Merge is the KF, a respect in which the burden of the KF is not met. The only role of this section is to show that the very few arguments put forth as evidence for this view, all based on anaphora, do not go through unless (i) operational Agree relations are multiplied for local relations and (ii) an additional device is introduced as a c-

command based interpretive relation to apply over long distances.

With respect to relations that appeal to unbounded c-command, Boeckx (2003) employs 'Match' from Chomsky (2000: 122),[45] a device similar to many in the literature in this respect, to account for the binding of resumptive pronouns in weak islands, along with Agree, and also 'intrusion', which permits resumption into strong islands (and must be c-command sensitive), as in a typical Irish example from McCloskey (1979: 34).

(A1) Sin teanga *aN* mbeadh meas agam ar duinear bith aL tá
that a-language aN would be respect at me on person aL is
ábalta í a labhairt.
able it to speak
'That's a language that I would respect anyone that could speak it.'
(McCloskey 1979: 34)

Moreover, intervention effects must have the same open-ended character to cross islands as we have already seen with respect to the Korean cases in (79), where intervention effects hold even if the moved element could not have moved to a position locally below the intervener (in the same cycle), so the intervention must hold over distances greater than those permitted by Agree. Moreover, many languages employ a logophoric form of a pronoun which must be used when it is anteceded by the reported speaker in the matrix clause. This is true even when the pronoun is embedded in an island. Clements (1975: 156) reports that a relative clause complement to the verb meaning 'remember' in Ewe cannot embed the logophoric form referring to the one who remembers (*yè* is ill-formed in place of *e* in (A2a)). However, *yè* becomes possible if the whole structure including the relative clause is embedded in the CP complement to a verb of saying as in (A2b).

(A2) a. Ama do nku nyɔnuvi hi dze e gbɔ dyi.
Ama set eye girl wh stay 3.PS side on

b. Ama gblɔ be yè-donku nyɔnuvi hi dze yè gbɔ dyi.
Ama say that yè set eye girl wh stay yè side on
'*Ama* said that *she* remembered the girl who stayed with *her*.'
(Clements 1975)

These arguments from resumption, intervention and logophoricity provide a wealth of evidence that unbounded c-command still plays an important role at the interpretive level.

Nonetheless, Chomsky (2007a, 2008), Reuland (2005), and Kratzer (2009) have suggested that appeal to Agree could displace the need for c-command. They argue that the locality principle for anaphora known as Principle A follows from Agree, either because Agree can do all that is required (e.g., Kratzer) or that

[45] Match is supposed to involve matching features, but as Seely (2006: 202) points out, selection cannot be Agree, but neither is it Match, since the selected feature does not match the categorial feature of the probe in any obvious way.

Agree can achieve binding relations that standard c-command cannot (Chomsky 2007a, 2008, following Reuland 2005).

In VSA, Case and agreement relations are established by heads that establish c-command domains bounded by interventions. In section 4.3, the distinctions between Case and agreement relations for T meant that T was in a local relation of one sort for agreement, and of another sort for Case. In other words, not all local relations are Agree relations. The same sort of argument can be made with respect to the separable locality relations that must be appealed to account for morphological verb agreement and anaphora relations. Since Agree is a local relation and it is a relation between a head and a maximal projection, it is not obviously a *binding* relation (see Safir 2008 for discussion), though there are theories consistent with this view for local anaphora, and without going into details, Kratzer's theory is one of them. She proposes that all binding relations are effected by heads that induce property formation by lambda extraction, such that the specifier of the head must bind a variable in the sister to the head. For every binding relation, there is a head of this kind.

However, Agree was originally formulated as a relation between a head and a phrase, with morphological consequences in the case of subject-verb agreement. In many Icelandic cases, an oblique argument is bound by a non-nominative subject, while the verb in the same clause agrees with a Nominative non-subject, as illustrated in (A3a–b).

(A3) a. Henni þykir broðr sinn/*hennar leiðinlegar.
she.DAT thinks brother SIN/her boring
'She finds her brother boring.'

b. Konunginum voru gefnar ambáttir í höll sinni/?hans.
the-king.DAT were given slaves in palace SIN/his
'The king was given slaves in his palace.'

(Zaenen *et al.* 1990: 102, 112)

In these examples, the structural subject is Dative, so the verb agrees with the post-verbal Nominative. The possessive anaphor *SIN* in Icelandic (inflected for agreement with what it modifies), which must normally be bound by a structural subject, is bound by the Dative subject. If Agree is crucially identifying which elements may move, then it is already failing to select between the possibilities in this case, since it is the Dative that fills the structural subject position (presumably to satisfy EPP). Moreover, it is also clear that verb agreement is not establishing the anaphoric relation (since the verb usually agrees with a structural subject that is Nominative). One could add an additional ad hoc head $H_{ANAPHOR}$ that forms an agreement relation between the Dative subject and something in the complement of that head, but this would be an antecedent agreement relation, not verb agreement.

In short, if Agree on a functional node determines verb agreement, then a different functional node must be responsible for anaphor-binding, with the result that all that is left of Agree is that it is a kind of local c-command mapped onto one interpretive or morphological relation or another. Rather than giving Agree rhetorical priority, it seems more natural to see agreement relations as one

of the family of local interpretive relations mapped onto c-command and bounded by intervention, which is exactly the theory proposed here. It is too great a task to provide an alternative account of local anaphora here, but one such alternative theory might be based on prominence relations within a domain that is defined by intervention, a model similar to that suggested for Case. In such an account, the role of heads may very well emerge as compatible with an account like Kratzer's, but the key notion is then c-command for an interpretive relation—not operational Agree.

A more interesting argument for Agree as the effector of anaphoric relations is presented by Reuland (2005) and adopted by Chomsky (2007a, 2008). Reuland argues that Agree is crucial to anaphoric relations because there are cases where a probe can effect an anaphoric interpretation when the antecedent of an anaphor is not c-commanded by its antecedent. In these cases, it is argued, the probe c-commands both anaphor and antecedent locally. The argument clearly does not go through for English.

(A4) a. There appear to be **a lot of gifted children** getting advantages that others deserve.

b. ?***A lot of gifted children*** seem to *each other's* parents to be getting advantages that others deserve.

c. *There appear to *each other's* parents to be ***a lot of gifted children*** getting advantages that others deserve.

In these instances, an Agree relation is supposed to hold between T and the bolded nominal, but where *each other's* and *a lot of gifted children* are co-construed, the presence of T agreeing with both nominals in (A4c) is not enough to effect anaphora. However, the arguments that Reuland makes are for Norwegian and Icelandic.

Reuland suggests that (A5a) in Norwegian provides the right configuration and I have added (A5b), which does not require an awkward context, but shows the same effect.

(A5) a. Det ble introdusert *en mann* for *seg selv*/**ham selv*. *Norwegian*
it became introduced a man to SEG-self
'A man was introduced to himself.'

b. Det ble introdusert *en mann* for læreren *sin*/**hans*.
it became introduced a man to teacher-the SIN/his
'A man was introduced to his teacher'

In these examples, the form of the anaphors *seg selv* and *sin* appear to indicate that they are bound by a structural subject, yet the only candidate antecedent is expletive *det*, which, Reuland assumes (and I concur) could not bind the subject-oriented antecedent even if *en mann* does c-command *seg selv*. For data such as these, however, it appears that the choice of verb and idiolect are crucial. The consultant whose judgments Reuland reports, Øystein Nilsen, also accepts (A6) (thanks to Øystein Nilsen, p.c., for supplying the Norwegian examples in this

section and his enlightening discussion of them).

(A6) a. Psykiateren introduserte en forvirret mann for seg/*ham selv.
psychiatrist-the introduced a confused man to SEG/him self
'The psychiatrist introduced the confused man to himself.'

b. Rektoren introduserte student-en/en student for
headmaster-the introduced student-the/a student to
læreren [sin/hans].
teacher-the SIN/his
'The headmaster introduced a/the student to his teacher.'

Øystein Nilsen (p.c.) remarks:

> The grammatical version of [(A6)] with *seg* is ambiguous for me, even though the subject-bound construal requires perhaps a confused shrink, rather than a confused man […]. Other Norwegians have a strictly [subject]-oriented *seg selv*, so I may be 'special' in not requiring subject-orientation.

He reports the same ambiguity for *sin* in (A6b), but *hans* cannot refer to either of the potential antecedents. If the active form of (A5a) in (A6) permits the direct object to bind the *seg self* or *sin* anaphors, then there is no reason to suppose that Agree is crucial to permitting the 'subject oriented' form in this case. Unless an additional Agree relation is added here for the complementation of 'introduce' in Nilsen's dialect (and incidentally, Nilsen's dialect does not even show subject-verb agreement in (A5)–(A6)), the agreement theory seems to be exactly wrong, and if such a relation is added, then no evidence for the Agree theory of anaphor antecedence is to be gained by it.[46]

---

[46] Nilsen suggests (and in this note I paraphrase what he told me) that *introdusere* is special, in that other ditransitives allow both the pronominal and the anaphoric possessor in examples parallel to (A6b).

(i) Jeg sendte Jens til faren sin/hans. *Norwegian*
I sent Jens to father-the SIN/his
'I sent Jens to his father.'

He notes that these other ditransitives allow for VP topicalization, leaving the dative PP behind. In that case, only the pronominal possessor is possible. *Introdusere* doesn't allow such partial VP fronting in the first place (the verb in VP topicalizations in Norwegian retains finite inflection).

(ii) [Sendte Jens] gjorde jeg til faren hans/*sin.
sent Jens did I to father-the his/SIN

(iii) * Introduserte Jens gjorde jeg for Per.
introduced Jens did I for Per

Nilsen interprets this to mean that the PP in *send*-type VPs is allowed in two different hierarchical positions, one 'adjoined' outside the constituent made up by the verb and the direct object, and another, 'Larsonian' position, c-commanded by the direct object. *Introdusere* would only allow the PP to occur in the lower, Larsonian position. Only the higher position would be able to feed partial VP fronting of the relevant sort, and possessive SIN bound by the object would only occur in the Larsonian position, while possessive 'his'

The purported Icelandic evidence Reuland presents is based on (86) elicited from Halldor Sigurðsson (p.c.).

(A7) Thað kom maður með börnin sin/*hans. *Icelandic*
there arrived a-man with children SIN/his
'A man arrived with his children.'

Reuland suggests that this is another example where the *sin* anaphor is only possible because of antecedency facilitated by Agree, not by the expletive *thað* or c-command from *maður*. However, the use of *sin* is freer in Icelandic than it is in Norwegian, and it is not always the case that it requires a structural subject, as in (A8), pointed out to me by Halldor Sigurðsson (p.c.).

(A8) Eftir vinnu var bara farið heim til sín.
after work was just gone home to self
'After work you/they/we/people just went home (to X-selves).'

For (A8) one might argue that Agree is doing the work in the absence of an antecedent, or perhaps that Agree has no role to play at all since a structural antecedent is not always required. In any case, examples like (A7) are not strong evidence for Agree as the effector of anaphoric relations, and the English and Norwegian examples seem to suggest that is not.

These arguments that Agree can replace Principle A are not persuasive, but even if there are other such arguments, the need for c-command to regulate resumption, scope intervention effects, and logophoric pronoun distributions remains unchallenged. The case for positing a structural relation more general than Agree as part of HLF seems secure.

Appendix B: The Varieties of Movement

VSA does not permit all the varieties of movement that Pesetsky (2000) has argued must exist. In particular, VSA does not permit any sort of covert phrasal movement, insofar as the highest occurrence of a non-terminal must be pronounced high. The question that must be addressed is whether or not it is necessary to weaken (51) or to seek different sorts of accounts for contexts where covert phrasal movement has been argued to exist.

Pesetsky (2000) argues that covert phrasal movement is necessary to account for antecedent-contained deletion within the family of analyses stemming from May (1985). These analyses appeal to covert phrasal movement of a quantified relative clause (adjunction to VP) to avoid infinite regress in the italicized portion in (B1a) that is not pronounced.

---

would be restricted to the higher adjunct position. Nilsen's reasoning here seems to justify the basic c-command story and would require ancillary stipulations in an Agree-based approach.

(B1) a. George trusts everyone (who) you do *trust everyone...*
b. T … [$_{VP}$ [$_{DP}$ everyone (who) you do [__]] [$_{VP}$ trust [*e*]]]

The VP [*trust* e] is then used to fill in the ellipsis. However, given copy theory, the 'empty' object of *trust* is still present as an infinitely regressing copy of what has moved (*everyone who you do*), and so Fox (2002) develops an alternative account whereby the clause portion of the relative is 'late attached' (consistent only with Revised Extension as proposed here). Without going into details, Fox's analysis of this construction does not appear to require any covert phrasal movement to take place because the relative clause coda with the ellipsis, attached late, is pronounced in its first-merged position as in (B2a), and is (discontinuously) associated with the relative clause nucleus (*everyone*).

(B2) a. George [*every* [[trusts [*every*one]] [who you do__]]]
b. [*everyone* … [*who* [C+WH-REL [*George* [T [*who* [*George* [$_{A}$ *v** [V *who*]]]

Interpreting Fox's approach slightly, the portion after *(who) you do*, which is elided at the bracket marked 'A', can be 'filled in' with the interpreted contents of the lower VP (*trust everyone*, where *everyone* must be the interpreted content of *who*). See Fox (2002) for discussion.

Pesetsky also posits a correlation between covert phrasal movement and the presence of superiority effects, such that *wh-in-situ* phrases undergo covert phrasal movement where superiority effects are present. Superiority is neutralized when the *wh*-phrases are D(iscourse)-linked.

(B3) a. Who saw who?
b. * Who did who see?
c. Which person did which person see?

It is not clear to me whether or not head movement pronounced low in VSA is empirically equivalent to Pesetsky's 'feature movement', but the optimal analytic result from the VSA perspective would be to argue that head/feature movement is the same for both D-linked and non-D-linked *wh-in-situ* and the superiority effect is a result of an intervention or interpretive condition neutralized by D-linking (structurally or semantically) (such as a choice function; e.g., Dayal 2002 or Reinhart 2006).

With respect to the linguistic variation in the realization of multiple interrogation structures, Pesetsky posits different complementizers that are stipulated to trigger single or multiple overt movement (or neither). These stipulations are only attractive if they are embedded in the rich set of assumptions that Pesetsky employs in order to draw together generalizations about movement, and so it is not clear that any real translation of his theory into VSA is possible. Richards (1999) and Pesetsky appeal to 'Shortest Move' which favors iMerge of a closer *wh*-Q-phrase rather than a more distant one to satisfy a trigger, but since I do not permit operational triggers, I cannot appeal to such a distinction. Moreover, Shortest Move only works in Pesetsky's system if covert phrasal movement unrelated to the ACD analysis exists, which is not established independent of

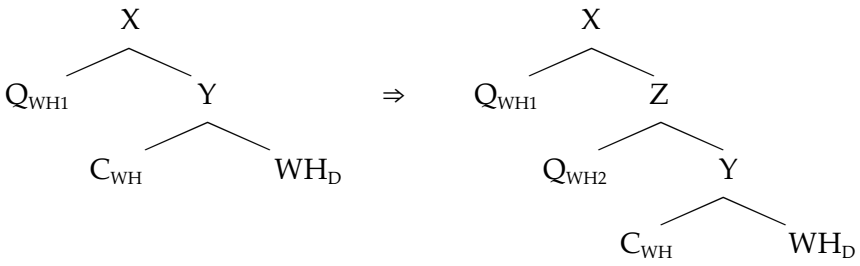
superiority effects. In VSA, the superiority effect can only arise as an intervention effect, such as the one described in (B4).

(B4) *Superiority as Intervention*[47]

The *wh*-phrase Q_{WH1} that is assigned WH_{D1} cannot have a lowest occurrence that is intervened by an occurrence of a Q_{WH2} if Q_{WH2} shares the scope of the Q_{WH1}.

For Q_{WH2} to share the scope of Q_{WH1}, it is necessary to introduce a way in which two quantifiers could share the same domain assignment. In VSA, scope assignment occurs at the unique point in a derivation where the Q or QP first c-commands WH_D, where WH_D is the domain set by the C_{WH} complementizer. Since domain assignment is unique, I need to add a device that allows for two quantifiers to share a domain for (83) to make sense. Notice that VSA permits 'tucking in' as in Richards (1999).

(B5)



In such configurations, an interpretive rule must be introduced to permit the scope of Q_{WH2} to be interpreted as parasitic on the scope of Q_{WH1}. A solution based on an instance of subMerge in (B5) has the virtue of preserving the primacy of initial wh-movement, while (B4) captures the relevant phenomenon in terms of intervention. If so, tucking in is then movement to a (parasitic) criterial position. A full treatment of superiority, one that reduces it to intervention without appealing to covert phrasal movement, is beyond the scope of this article, but these remarks suggest a plausible avenue to explore within VSA.[48]

---

[47] Hornstein (2001) has suggested that superiority may be a form of weak crossover, and a correlation between the two phenomena has been established, for example, by Adesola (2006). If so, a dependency relation holds between the extracted wh-phrase and the in situ one, and the intervention effect may be a subcase of the Independence Principle, proposed in Safir (2004b), where it is argued that a variable or the constituent that contains it cannot c-command anything the variable depends on.

[48] Challenges for VSA remain. Where phrasal movement is required, as in English, I have already accepted the (inelegant) possibility that a head can require that it be assigned to a non-terminal node (65). Nothing in the system requires multiple fronting, and indeed that is not stateable in terms of how a head is subMerged. Languages such as Bulgarian that require multiple fronting in multiple interrogations will have to be addressed in some other way. It could be that whatever permits multiple phrases to be assigned the same scopal domain can also be manipulated to express these differences (though many typological issues, including those raised by Bošković 2002, need consideration). Such cases pose a challenge to VSA, but they are problems that seem tractable insofar as appeal to triggers distinguishing iMerge from eMerge are not required.

References

Ackema, Peter & Ad Neeleman. 2004. *Beyond Morphology: Interface Conditions on Word Formation*. Oxford: Oxford University Press.

Adesola, Oluseye. 2006. On the absence of superiority and weak crossover effects in Yoruba. *Linguistic Inquiry* 37, 309–318.

Baker, Mark & Nadya Vinokurova. 2008. Two modalities of case assignment in Sakha. Ms., Rutgers University.

Beck, Sigrid & Shin-Sook Kim. 1997. On *wh*- and operator scope in Korean. *Journal of East Asian Linguistics* 6, 339–384.

Beck, Sigrid. 1996. Quantified structures as barriers for LF-movement. *Natural Language Semantics* 4, 1–56.

Beghelli, Filippo & Tim Stowell. 1997. Distributivity and negation. In Anna Szabolcsi (ed.), *Ways of Scope Taking*, 71–107. Dordrecht: Kluwer.

Bhatt, Rajesh & Roumyana Pancheva. 2004. Late merger of degree clauses. *Linguistic Inquiry* 35, 1–45.

Bittner, Maria & Ken Hale. 1996. The structural determination of Case and agreement. *Linguistic Inquiry* 27, 1–68.

Bobaljik, Jonathan & Samuel Brown. 1997. Interarboreal operations: Head movement and the extension requirement. *Linguistic Inquiry* 28, 345–356.

Bobaljik, Jonathan & Susi Wurmbrand. 2005. The domain of agreement. *Natural Language and Linguistic Theory* 23, 809–865.

Bobaljik, Jonathan & Susi Wurmbrand. 2008. Case in GB/minimalism. In Andrej Malchukov & Andrew Spencer (eds), *The Oxford Handbook of Case*, 44–58. Oxford: Oxford University Press.

Bobaljik, Jonathan. 2002. A-chains at the PF-interface: Copies and 'covert' movement. *Natural Language and Linguistic Theory* 20, 197–267.

Bobaljik, Jonathan. 2008. Where's phi? Agreement as a post-syntactic operation. In David Adger, Susana Béjar, & Daniel Harbour (eds.), *Phi-Theory: Phi-Features across Interfaces and Modules*, 295–328. Oxford: Oxford University Press.

Boeckx, Cedric & Norbert Hornstein. 2006. The virtues of control as movement. *Syntax* 9, 118–130.

Boeckx, Cedric. 2003. *Islands and Chains*. Amsterdam: John Benjamins.

Borer, Hagit. 2005a. *In Name Only: Structuring Sense* vol. I. Oxford: Oxford University Press.

Further issues arise if the *in-situ* strategy for multiple wh-interrogatives is not uniform across languages. For example, Dayal (2002) contends that *in situ wh* in islands in Hindi only permits single pair answers to questions, not multiple pair answers. This difference does not follow from assuming that the Hindi *in situ* strategy is head movement pronounced low (Dayal proposes that islands require interpretation by choice function). Dayal attributes this locality effect from a (stipulated) clausemate restriction on covert *wh*-movement. The effect might be recast as an intervention effect if multiple interrogation domains are linked to event or tense domains in Hindi, such that the head setting one event/tense domain intervenes for another (a similar domain limits anaphora in Hindi, see Safir, 2004a: 166–170). Then multiple pair interpretations would be limited to a domain containing a single event/tense, with single pair interpretations possible where an event/tense head intervenes between criterial landing sites. Head movement would not then have to be restricted in an ad hoc way. The exploration of this possibility is beyond the scope of this essay.

Borer, Hagit. 2005b. *The Normal Course of Events: Structuring Sense*, vol. II. Oxford: Oxford University Press.
Bošković, Željko. 2002. On multiple *wh*-fronting. *Linguistic Inquiry* 33, 351–383.
Bošković, Željko. 2007. On the locality and motivation of Move and Agree: An even more minimal theory. *Linguistic Inquiry* 38, 589–644.
Bossong, Georg. 1985. *Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Gunter Narr Verlag.
Cable, Seth. 2008. Towards the elimination of pied-piping: Evidence from Tlingit. Ms., University of Massachusetts, Amherst.
Chomsky, Noam. 1981. *Lectures on Government and Binding: The Pisa Lectures*. Dordrecht: Foris.
Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.
Chomsky, Noam. 1995. *The Minimalist Program.* Cambridge, MA: MIT Press.
Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.
Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A Life in Language*, 1–52. Cambridge, MA: MIT Press.
Chomsky, Noam. 2004. Beyond explanatory adequacy. In Adriana Belletti (ed.), *Structures and Beyond: The Cartography of Syntactic Structures*, vol. 3, 104–131. New York: Oxford University Press.
Chomsky, Noam. 2007a. Approaching UG from below. In Uli Sauerland & Hans-Martin Gärtner (eds.), *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax–Semantics* (Studies in Generative Grammar 89), 1-29. Berlin: Mouton de Gruyter.
Chomsky, Noam. 2007b. Of minds and language. *Biolinguistics* 1, 9–27.
Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud,* 133–166. Cambridge, MA: MIT Press
Christiansen, Morten & Nick Chater. 2008. Language as shaped by the brain. *Behavioral and Brain Sciences* 31, 489–558.
Clements, George. 1975. The logophoric pronoun in Ewe: Its role in discourse. *Journal of West African Languages* 10, 141–177.
Collins, Chris. 2002. Eliminating labels. In Samuel David Epstein & T. Daniel Seely (eds.), *Derivation and Explanation in the Minimalist Program*, 43–64. Malden, MA: Blackwell.
Dayal, Veneeta. 2002. Single pair vs. multiple pair answers, *wh*-in situ and scope. *Linguistic Inquiry* 33, 512–520.
Epstein, Samuel David, Erich M. Groat, Ruriko Kawashima & Hisatsugu Kitahara. 1998. *A Derivational Approach to Syntactic Relations.* New York: Oxford University Press.
Fox, Danny & David Pesetsky. 2005. Cyclic linearization of syntactic structure. *Theoretical Linguistics* 31, 1–45.
Fox, Danny. 2002. ACD and the copy theory of movement. *Linguistic Inquiry* 33, 64–96.
Fox, Danny. 2003. On logical form. In Randall Hendrick (ed.), *Minimalist Syntax*, 82–123. Malden, MA: Blackwell.

Frank, Robert & K. Vijay-Shanker. 2001. Primitive c-command. *Syntax* 4, 164–204.
Haegeman, Liliane & Raffaella Zanuttini. 1996. Negative concord in West Flemish. In Adriana Belletti & Luigi Rizzi (eds.), *Parameters and Functional Heads: Essays in Comparative Syntax*, 117–179. New York: Oxford University Press.
Haegeman, Liliane. 1995. *The Syntax of Negation*. Cambridge: Cambridge University Press.
Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298, 1569–1579.
Heck, Fabian. 2004. A theory of pied piping. Tübingen: Eberhard Karls Universität Tübingen dissertation.
Heck, Fabian. 2007. On certain properties of pied piping. Ms., Universität Leipzig.
Hornstein, Norbert & Jairo Nunes. 2008. Adjunction, labeling, and bare phrase structure. *Biolinguistics* 2, 57–86.
Hornstein, Norbert. 1999. Movement and control. *Linguistic Inquiry* 30, 69–96.
Hornstein, Norbert. 2001. *Move! A Minimalist Theory of Construal*. Malden, MA: Blackwell.
Hornstein, Norbert. 2003. On control. In Randall Hendrick (ed.), *Minimalist Syntax*, 6–81. Malden, MA: Blackwell.
Huang, James. 1992. Logical relations in Chinese and the theory of grammar. Cambridge, MA: MIT dissertation.
Jayaseelan, K.A. 2008. Bare phrase structure and specifierless syntax. *Biolinguistics* 3, 87–106.
Jónsson, Jóhannes. 1996. Clausal architecture and case in Icelandic. Amherst, MA: University of Massachusetts dissertation.
Kim, Hyunjoo. Forthcoming. Intervention effects and movement in *wh*-in-situ constructions. New Brunswick, NJ: Rutgers University dissertation.
Koster, Jan. 1976. *Locality Principles in Syntax*. Dordrecht: Reidel.
Kratzer, Angelika & Junko Shimoyama. 2002. Indeterminate pronouns: The view from Japanese. *Proceedings of the 3rd Tokyo Conference on Psycholinguistic*, 1–25. [Available online at the Semantics Archive: http://www.semanticsarchive.net/Archive/WEwNjc4Z.]
Kratzer, Angelika. 2009. Making a pronoun: Fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry* 40, 187–237.
Landau, Idan. 2000. *Elements of Control: Structure and Meaning in Infinitival Constructions*. Dordrecht: Kluwer.
Landau, Idan. 2003. Movement out of control. *Linguistic Inquiry* 34, 471–498
Landau, Idan. 2006. Severing the distribution of PRO from case. *Syntax* 9, 155–170.
Lasnik, Howard & Juan Uriagereka with Cedric Boeckx. 2005. *A Course in Minimalist Syntax*. Malden, MA: Blackwell.
Lebeaux, David. 1991. Relative clauses, licensing and the nature of the derivation. In Susan Rothstein (ed.), *Perspectives on Phrase Structure Heads and Licensing*, 209–239. New York: Academic Press.
Legate, Julie Anne. 2008. Morphological and Abstract Case. *Linguistic Inquiry* 39,

55–102.
Linares, Carlo. 2008. Approaching DOM in Viable Syntax. Manuscript, Rutgers.
Linares, Carlo. Forthcoming. On the typology of movement [tentative title]. New Brunswick, NJ: Rutgers University dissertation.
Longobardi, Giuseppe. 1994. Reference and proper names: A theory of N-movement in syntax and logical form. *Linguistic Inquiry* 25, 609–665.
Marantz, Alec. 1991. Case and licensing. *ESCOL '91*, 234–253. [Reprinted 2000 in Eric Reuland (ed.), *Arguments and Case: Explaining Burzio's Generalization*, 11–30. Amsterdam: John Benjamins.]
Matushansky, Ora. 2006. Head movement in linguistic theory. *Linguistic Inquiry* 37, 69–109.
May, Robert. 1985. *Logical Form*. Cambridge, MA: MIT Press.
McCloskey, James. 1979. *Transformational Syntax and Model-Theoretic Semantics*: *A Case Study in Modern Irish*. Dordrecht: Reidel.
McCloskey, James. 2002. Resumption, successive cyclicity and the locality of operations. In Samuel David Epstein & T. Daniel Seely (eds.), *Derivation and the Explanation in the Minimalist Program*, 184–226. Malden, MA: Blackwell.
McFadden, Thomas. 2004. The position of morphological case in the derivation: A study on the syntax-morphology interface. Philadelphia, PA: University of Pennsylvania dissertation.
Neeleman, Ad, & Hank van de Koot. 2002. The configurational matrix. *Linguistic Inquiry* 33, 529–574.
Nunes, Jairo. 1999. Linearization of chains and phonetic realization of chain links. In Samuel David Epstein & Norbert Hornstein (eds.), *Working Minimalism*, 217–249. Cambridge, MA: MIT Press.
Nunes. Jairo. 2004. *Linearization of Chains and Sideward Movement*. Cambridge, MA: MIT Press.
Obenauer, Hans. 1984. On the identification of empty categories. *The Linguistic Review* 4: 153–202.
Pesetsky, David & Esther Torrego. 2001. T-to-C movement: Causes and consequences. In Michael Kenstowicz (ed.), *Ken Hale: A Life in Language*, 355–426. Cambridge, MA: MIT Press.
Pesetsky, David. 2000. *Phrasal Movement and Its Kin*. Cambridge, MA: MIT Press.
Ramchand, Gillian. 2008. *Verb Meaning and the Lexicon: A First Phase Syntax*. Cambridge: Cambridge University Press.
Reinhart. Tanya. 2006. *Interface Strategies*. Cambridge, MA: MIT Press.
Reuland, Eric. 2005. Agreeing to bind. In Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinerz, & Jan Koster (eds.), *Organizing Grammar: Studies in Honor of Henk van Riemsdijk*, 505–513. Berlin: Mouton de Gruyter.
Richards, Norvin. 1999. Featural cyclicity and the ordering of multiple specifiers. In Samuel David Epstein & Norbert Hornstein (eds.), *Working Minimalism*, 127–158. Cambridge, MA: MIT Press.
Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, MA: MIT Press.
Rizzi, Luigi. 2006. On the form of chains: Criterial positions and ECP effects. In Lisa Lai-Shen Cheng & Norbert Corver (eds.), *WH-Movement: Moving On*, 97–133. Cambridge, MA: MIT Press.
Ross, John R. 1967. Constraints on variables in syntax. Cambridge, MA: MIT

dissertation.
Safir, Ken. 1986. Relative clauses in a theory of binding and levels. *Linguistic Inquiry* 17, 663–689.
Safir, Ken. 1987. The syntactic projection of lexical thematic structure. *Natural Language and Linguistic Theory* 5, 561–601.
Safir, Ken. 1999. Vehicle change and reconstruction in A'-chains. *Linguistic Inquiry* 30, 587–620.
Safir, Ken. 2004a. *The Syntax of Anaphora*. Oxford: Oxford University Press.
Safir, Ken. 2004b. *The Syntax of (In)dependence*. Cambridge, MA: MIT Press.
Safir, Ken. 2005. Introducing exxtension. In Yukio Otsu (Ed.), *The Proceedings of the Fifth Tokyo Conference on Psycholinguistics*, 1–31. Tokyo: Hitsuzi Syobo.
Safir, Ken. 2008. Coconstrual and narrow syntax. *Syntax*: 330–355.
Safir, Ken. In preparation. Irish complementizers and extraction without triggers. Ms., Rutgers University.
Sauerland, Ulrich & Fabian Heck. 2003. LF-intervention effects in pied-piping. *NELS* 33, 347–366.
Seely, T. Dan. 2006. Merge, derivational c-command, and subcategorization in a label free syntax. In Cedric Boeckx (ed.), *Minimalist Essays*, 182–217. Amsterdam: John Benjamins.
Starke, Michal. 2004. On the inexistence of specifiers and the nature of heads. In Adriana Belletti (ed.), *Structures and Beyond: The Cartography of Syntactic Structures*, vol. 3, 252–268. New York: Oxford University Press.
Stepanov, Arthur. 2001. Late adjunction and minimalist phrase structure. *Syntax* 4, 94–125.
Torrego, Esther. 1998. *The Dependencies of Objects*. Cambridge, MA: MIT Press.
van Urk, Coppe. 2009. On obligatory control. Ms., University of Utrecht.
Vinokurova, N.1998. The internal structure of noun phrases in Sakha. Tromsoe: Universitetet i Tromsø M.Phil. thesis.
Watanabe, Akira. 2006. The pied piper feature. In Lisa Lai-Shen Cheng & Norbert Corver (eds.), *WH-Movement: Moving On*, 47–70. Cambridge, MA: MIT Press.
Zaenen, Annie, Joan Maling & Höskuldur Thráinsson. 1990. Case and grammatical relations: The Icelandic passive. In Joan Maling & Annie Zaenen (eds.), *Modern Icelandic Syntax*, 95–136. New York: Academic Press. [Reprinted from *Natural Language and Linguistic Theory* 3, 441–483, 1985].

Ken Safir
Rutgers University
Department of Linguistics
18 Seminary Place
New Brunswick, NJ 08901
USA
safir@ruccs.rutgers.edu