(Dis)Agree

(Dis)Agree:
Exploring Agreement Mechanisms

by

Pritha Chandra

CAMBRIDGE
SCHOLARS
PUBLISHING

(Dis)Agree: Exploring Agreement Mechanisms,
by Pritha Chandra

This book first published 2011

Cambridge Scholars Publishing

12 Back Chapman Street, Newcastle upon Tyne, NE6 2XX, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-4438-3146-8, ISBN (13): 978-1-4438-3146-8

FOR MY PARENTS AND MY DAUGHTER HOMEYRA

TABLE OF CONTENTS

PREFACE

The main thesis of this work is that one of the three fundamental operations of Universal Grammar, Agree underlying agreement in natural language, is not supported by conceptually well-established arguments and the empirical justification given for it is also not uncontestable. The advent of minimalism in linguistic theory necessitates doubting all attributes of the language faculty that seem unique to it. If language is part of cognition, the rest of cognition should be reflected in its workings, thus ruling out the possibility of the language organ standing out for being too idiosyncratic. Anything language-specific should be defended as necessary for the system to compute and on general principles of model-building such as economy, simplicity, symmetry and non-redundancy. Agree seems very language-specific and yet the literature that readily accepts it hardly ever tries to locate it within the cognitive domain. I have made an effort in this direction and my objective is to show that this operation is not conceptually necessary to the language system. It cannot be justified on general economy considerations and worse, the relation c-command that lies at its essence has itself come under the radar, thus weakening the very base for Agree. Talking of agreement mechanisms then, Agree could and should be replaced by conceptually necessary operations of the grammar like Merge (both internal and external). Local sisterhood relations between 'probes' and 'goals' are the ones where agreement thrives in the grammar.

This work began in the corridors of Mary Mount Hall, University of Maryland, College Park and would have not seen the light of the day had it not been for the inspirations that I gained from the linguists there. Those who directly contributed to the present work were also fortunately my thesis committee members and I am very grateful to Cedric Boeckx, Norbert Hornstein, Howard Lasnik and Juan Uriagereka for giving shape to my ideas but also generally for leading me into the fascinating world of agreement, so widely researched and yet so poorly understood. Others who deserve equal appreciation for helping me understand the complexities of the linguistic system are Ivano Caponigro, Stephen Crain, Paul Pietroski and Rosalind Thornton. Discussions with Rajesh Bhatt, Ilhan Cagri, Alice Davison, Tomohiro Fujii, Scott Fults, Soo-min Hong, Atakan Ince, Youngmi Jeong, Anoop Mahajan, Alan Munn, Andrew Nevins, Ivan Ortega-Santos, Maria Polinsky, Dominik Rus, Asad Sayeed,

Usama Soltan, K. V. Subbarao, Heather Taylor and Heather Wilson have helped clean the present work of many errors that would have otherwise gone unnoticed.

Back home, those who have been instrumental in keeping my interests in agreement and related phenomena alive are my colleagues and students in the Indian Institute of Technology Delhi and Jawaharlal Nehru University. Most notably, stimulating discussions with Ayesha Kidwai have driven me to rethink many of the architectural issues of natural language grammar.

Finally, thanks to Pratyush for helping me view my own research from a non-linguist's perspective; his critical assessments, exasperating at times, have forced me to seek answers for the quintessential 'why' question behind any research program. Without his constant encouragement, the present work faced the risk of getting shelved in some obscure corner of my office permanently.

CHAPTER ONE

AGREE AND THIRD FACTOR DETERMINANTS: A CONCEPTUAL REASSESSMENT

In the pre-minimalist era, the relation 'government' served as one of the primary players, with a number of grammatical phenomena and principles including Case, Empty Category Principle (ECP) defined in its terms. With the advent of minimalism, however, the notion has become suspect for being too peculiar to the language faculty. One of the primary guiding forces of the minimalist program has been the search for third factor determinants – physical and chemical laws governing growth, principles of computational efficiency that define other systems (physical, chemical and cognitive) – in the formation of the Faculty of Language (FL). A natural follow-up has been to substitute the seemingly language-specific relations like government with principles that are conceptually necessary, for instance, local X-bar theoretic relations to the head of a projection (Chomsky 1993). Locality being an ubiquitous feature present in many language-external systems, would count in as a third factor principle, thus helping the linguist to reduce or eliminate the structural peculiarities endowed to Universal Grammar (UG) – in our case, government, and in consequence make it easier for her to explain the story of its evolution. And yet, in recent works (Chomsky 2004a, b, 2005a, b among others), an operation – Agree – has been proposed for phi-feature agreement (with case valuation as its reflex) whose mechanism very closely resembles that of government. Agree is a long distance relation between a head and a Noun Phrase (NP), essentially dissociated from all kinds of movement. Chomsky defends Agree on two counts. The first and conceptual argument is based on a specific understanding of a search domain, assumedly also the most natural one. It is based on the relation c-command holding between a head and its sister complement and all that she contains. Agree is therefore an asymmetric relation that holds between a head (probe) and elements (goals) that it c-commands. On close heels to this conceptual argument come various empirical phenomena strongly supporting the

existence of long-distance agreement relations between elements placed at considerable distance from each other.

This chapter tries to debunk the first argument. It shows that if a bottom up approach to UG were to develop that requires us to seek third factor considerations in UG's formation before we start attributing much to the genetic endowment of natural language, Agree is as much suspect as government is. One specific concern is the dubious status of c-command on which this agreement mechanism is based. Since c-command has itself come under the radar recently, having an agreement mechanism built on it seems a definite drawback for the system. Moreover, Chomsky's idea of the c-command domain as the most optimal search domain for agreement may not ultimately be right, and as I show below, actually contradicts his conception of headedness in natural language. In light of these various problems, I present an alternative thesis: Agree should be eliminated and all agreement relations must be recast in local sisterhood relations (conventionally understood as specifier-head and head-complement relations). The notion of search domain also needs reframing in Extension Condition terms, as has also been suggested by Hornstein (2005), another modification that, I believe, will take us towards a grammar that is also 'beyond explanatory' adequate. As far as this is true, I contend that the faculty of language has nothing beyond Merge and Move/Internal Merge, the first being inevitable in any language-like system and the latter necessitated by interface exigencies. With this brief introduction about the main claims of the chapter in place, I move on to the details now. But before that, presented below is a rough roadmap of the chapter.

I initiate the discussion with a very brief overview of how minimalism assists in a better understanding of the language faculty (i) as located "within the array of cognitive systems of the mind/brain" and (ii) as embedded in the evolutionary process, conferring to "general considerations of conceptual naturalness" (Chomsky 1995:1). Assuming that language is an optimal solution to legibility conditions and adopting as the research guideline that any property specific to language calls for a "principled explanation", I set out to evaluate in section 2 the justifications hitherto provided in the literature for postulating the three fundamental operations in the grammar: Merge, Move (a.k.a. Internal Merge) and Agree. This section reveals that unlike the former two, Agree is unique to the language faculty and also cannot be justified on general economy considerations. An even greater challenge for Agree comes from its dependence on c-command, itself not a primitive of the grammar. Section 3 includes a discussion on what counts as optimal configurations for phi-agreement. Following that in section 4, I present the empirical evidence that various

researchers have previously adduced in its favor, the reanalysis of which constitute the agenda for the remaining chapters.

1.1 Minimalist Strivings

Piatelli-Palmarini, as cited in cited in Uriagereka (1998: xxi-xxii) hails minimalism as that which "allows us to participate in one of the most fascinating conceptual adventures of our time...a revolution within a revolution". Such endearing proclamations have invited, and naturally so, a lot of criticisms from critics (see Seuren 2004). However, despite that, one cannot ignore the shift that minimalism has brought into the study of language, a radical departure from the immediately preceding Principles and Parameters Theory, while keeping the primary task of linguistic theory intact.

Three trends in the short history (of a little more than 50 years) of generative transformational grammar are identified in Boeckx and Hornstein (2010). The first – the combinatoric era – started with Chomsky's (1957) Syntactic Structures, which laid out the formal tools to identify the grammatical structures from the ungrammatical ones. This differentiation, as Chomsky notes in other places (see for instance, his opening remarks 'Of Minds and Language: A Dialogue with Noam Chomsky in Basque Country'), was however simply in response to the conventional wisdom of the times. The engineering field was at that time, under the influence of information theory along the lines of Warren Weaver's essay (see Shannon and Weaver 1949), and a popular idea was that human language properties could very well be handled by Markovian processes. Not surprisingly, Chomsky's Syntactic Structures, as it was aimed mainly for engineering students of MIT, opened with the inadequacies of Markovian processes for human language. On similar lines, the task of linguistic theory was mentioned there as distinguishing grammatical from ungrammatical sentences, on an analogy of well-formedness in formal systems. Despite being mentioned in the first ever-published book on transformational grammar, one must not confuse them with the goals of linguistic theory. Importantly, "[i]n the much longer and more elaborate unpublished monograph of LSLT two years earlier, intended only for a few friends, there is no mention of finite automata, ... the task of the theory of language is to generate sound-meaning relations fully, whatever the status of an expression ... (Chomsky 2009: 15). Over the span of fifty years and more, that particular goal hasn't changed, though new tasks have been taken up from time to time, often leading to revolutionary ruptures in the way we think about language.

In the trajectory of scientific development there are crucial moments or "extraordinary episodes" constituting decisive shifts in "professional commitments" triggering full-fledged scientific revolutions. Kuhn (1962) calls these episodes "the tradition-shattering complements to the tradition-bound activity of normal science." The Principles and Parameters (P&P) approach (Chomsky 1981, 1986) can be rightly thought of bringing about one such epistemological rupture or paradigm shift within the generative tradition of linguistics. This second phase marks the cognitive era of linguistic theory (Boeckx and Hornstein ibid).

In the era preceding this "radical departure", the Faculty of language (FL) was construed with a rich and highly structured format. The "general theory of linguistic structures" (as in Chomsky 1957:50) was an arena of ever-increasing complex, language-specific and construction-specific rules and principles. A common assumption underlying works of that era was that theory-formation required restricting the domain of admissible hypotheses, so that the language learner may hit upon the right theories corresponding to the input. That is to say that the range of competing grammars or 'meta-languages' tested against the evaluation procedure by the language-learner were extremely restricted, making it possible to answer 'the logical problem of language acquisition' with some success.

With the crystallization of the P&P approach, there was a drastic change in this scenario, with the principles of language now liberated from acquisition, itself reduced to the matter of parameter-setting (presumably, with binary values). FL was conceived as consisting of a restricted set of invariant principles and a few parameters, which accounted for both language variance and acquisition. There was no apparent necessity "to satisfy the empirical conditions of rapid convergence on generative systems of the kind required to determine meaning and external manifestation" (Chomsky 2006:2).

This relative autonomy of FL from its own empirical realization allows linguists to investigate it in its being or essence. This is a necessary step of abstraction, for until and unless a phenomenon is identified and given a proper shape with its own internal essence, it continues to be overwhelmed by the externalities and can only be deductively traced. On the contrary, by its identification (through the knowledge of its essence), we render possible the making of the "thing in itself" - paving the way for a more comprehensive analysis of the system, its interactions with other systems and finally its organization or design in accordance with the general principles of nature. The same observations extend to the domain of FL, understood as a biological organ. With the P&P approach facilitating (these necessary and further) steps of abstraction, we can now conceptualize

FL as embodying principles of growth and development typical to other cognitive domains of the mind/brain that also play a vital role in the evolutionary process. The focus of inquiry shifts from the nature of UG and linguistic experience to the "third factor" determinants in language design (Chomsky 2005a) - principles of structural architecture, constraints on the development and evolution of organic forms and computational efficiency. Notions like simplicity, economy, symmetry, non-redundancy are no longer merely methodological issues, but are directly derived from the workings of nature (also see Uriagereka 1998, 2002). The linguist's job is to attempt to eliminate some of the language-specific technology (i.e. attribute less to the genetic endowment or UG) and, consequently, provide more "principled explanations", i.e. provide answers that help reduce properties of language to properties of the interface systems and/or follow from general economy considerations.

The Minimalist Program (Chomsky 1995) helps us in formulating the right research questions in this regard. It marks, what Boeckx and Hornstein call the third – minimalist – phase of linguistic theory. The substantive thesis of minimalism (or 'ontological minimalism', see Martin and Uriagereka 2001) is that language is an optimal solution to legibility or interface conditions (also dubbed as the Strong Minimalist Thesis/SMT). Language design is influenced by interface requirements, i.e. there is nothing in FL beyond what is required to produce optimal solutions for the interfaces in accordance to general principles of computational efficiency. The remaining task is then to attempt to validate this bold conjecture by minimizing the differences that exist between FL and SMT. Lesser the differences between them, less specific or unique becomes UG. "UG is what remains when the gap between SMT and the true nature of FL have been reduced to the minimum, when all third factor effects have been identified" (Chomsky 2006:3).

Noteworthy is the continuity one notices in the primary objective underlying the transformational linguistic theory despite numerous changes in the theoretical nuances and formal implementations the system has undergone over the years. As in the LSLT days, the objective of minimalism is still to relate sound-meaning pairs. But the nuances of the system determine how these pairings are accomplished. In minimalism, a lot of the architecture determining sound-meaning representations is justified in adherence to general laws of growth and not just by resorting to the rich genetic endowments of UG, a radical step away from previous approaches to language. The three phases that Boeckx and Hornstein observe are therefore not isolated from each other; rather they are bound together by a common goal – the search for a system that relates the sound

and meaning representations of an expression. The changes that one witnesses, unlike what Boeckx and Hornstein would like us to believe, are not necessarily drawn or influenced by better- developed sciences. They are natural fall-outs of the workings and contradictions of each of the systems we have seen so far, with the new theoretical vistas emerging from their predecessor theories.

This book is restricted to an operation in minimalism – Agree - and I strictly adhere to the assumptions of the program here to understand whether this operation is indeed indispensable. Therefore taking the SMT as a point of departure, I venture to ask at this point, to what extent the core grammatical operations – Merge, Move and Agree – satisfy these conditions; i.e. whether or not they can reduced to general architectural properties or whether they constitute part of the language-specific technology incorporated into the system to satisfy legibility conditions at the interfaces? In a nutshell, do we have a "principled explanation" for each of these operations?

1.2 Merge, Move & Agree: Their Conceptual Loci

1.2.1 Merge: An Indispensable Operation

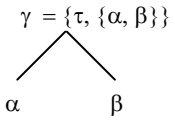
In terms of "naturalness", Merge stands out for being the least controversial among the set of operations. Note what Chomsky (2005a:11-12) has to say about it:

> "An elementary fact about the language faculty is that it is system of discrete infinity. Any such system is based on a primitive operation that takes n objects already constructed, and constructs from them a new object. Call that operation Merge. Either Merge or some equivalent is a minimal requirement…The simplest account of the "Great Leap Forward" in the evolution of humans would be that the brain was rewired perhaps by some slight mutation, to provide the operation Merge, at once laying a core part of the basis for what is found at that dramatic "moment" of human evolution…"[1]

To see the operation at work, consider (1):

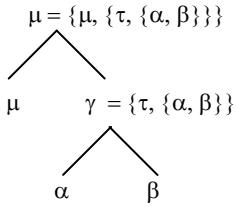
[1] To give an example of a symbolic system that uses concatenation, the language of propositional logic builds well-formed formulas by applying a fixed set of logical operators Ω = {not, and, or, conditional} on a finite set of terminal elements A= (p, q, r).

(1)



Merge takes two elements α and β to form a complex unit γ, with label τ projected from either of γ's daughter-nodes. Recursive Merge allows larger units to be built from the previously constructed γ by concatenating more items to the structure. The newly constructed tree in (2) is labeled μ and is derived by merging an extra item μ with γ

(2)



Merge provides two relations for free: (i) 'sisterhood' between two items merged to each other and (ii) 'immediately contain' between the newly constructed structure and the most recent concatenates. Composition of relations gives us a larger set of relations. I take up this matter in more detail in section 3.

In summary, Merge in natural language is not an anomaly or an imperfection; it is evident in all symbolic systems. We therefore have a "principled explanation" for positing it in the grammar.

Boeckx (2006a, 2009) makes a further attempt at decomposing Merge into a) a simple grouping procedure and b) labeling, both of which he feels are available in other cognitive systems and hence make up a more principled analysis for the existence of this operation in natural language. For Boeckx, though hierarchical structures are available in other systems, what is unique to human language is its headedness, i.e. the ability to take two items α and β and merge them to get a 'new' object by projecting one of the constituent items. Labeling, another ubiquitous operation in mental modules, plays a significant role in building recursive structures in human language.

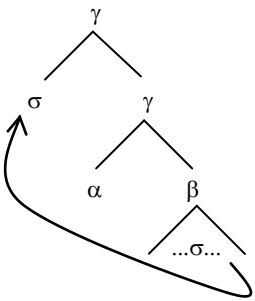
The problem with this analysis, that Chomsky (2009) among others raise, is that grouping, roughly amounting to concatenation has order inbuilt into it while merge does not. This makes Merge a simpler operation than concatenation, thus providing no reason to decompose it any further.

Secondly, headedness is not a unique property of FL, but rather an epiphenomenon that falls out naturally from the way unbounded merge works in natural language. This is a point I would like to come back to, when dealing with the Agree operation, as Chomsky's idea of headedness as an epiphenomenon is crucially linked to his idea of minimal or optimal search. Since the issue of minimal search is also closely linked to Agree, it is pertinent that we look it up more thoroughly in the context of the said operation.

1.2.2 Move: From Imperfection to Virtual Conceptual Necessity

There is a second operation in the grammar, Move, which displaces an element σ (or a copy of σ) within a phrase marker. Take (3) for illustration, where σ is shown to displace from its β-internal position to the specifier of γ.

(3)



In the initial days of the minimalist program, movement or dislocation was considered an "imperfection". The argument for this claim came from the fact that unlike Merge - this is a property very specific to language; it is never found as part of other symbolic systems. Chomsky (2000:120) seeks to justify its presence, along with that of the non-interpretable features (the driving force behind movement in the grammar) by suggesting that perhaps "they have to do with externally imposed legibility conditions" The Conceptual-Intentional systems require that outputs (SEM) to it from narrow syntax express interpretation roughly corresponding to (i) "deep" or theta-theoretic information and (ii) "surface" or scope and discourse properties. The former is of the general type present in all language-like systems whereas the latter appears very specific to human language. This distinction, Chomsky claims, is marked by language in a systematic way –

by including an operation like Move. Move or dislocation is incorporated into the language system with the objective of generating the second type of interpretation; namely "surface" or scope and discourse properties. It is therefore an anomaly but one that FL must possess in order to satisfy certain interpretative effects. As for the non-interpretable features, Chomsky conjectures that these too are only apparent imperfections as they are absolutely necessary for the system to be able to compute. Therefore, given the specific role of non-interpretable features in triggering dislocation in narrow syntax, it is conceivable that they are required by design specifications and therefore do not count as aberrations. Under this view, "the two imperfections might reduce to one, the dislocation property" (121).

Another problem noted with regard to Move was its (alleged) complex nature. It was assumed that the operation is a complex one composed of Merge and Agree (and delete). Good design conditions lead "us to expect that simpler operations are preferred to more complex ones, so that Merge or Agree (or their combination) preempt Move, which is a "last resort", chosen when nothing else is possible" (Chomsky 2000:101-2). Under this approach, Move is a glitch that the grammar employs when derivational convergence is at stake.

More recently however, Chomsky has taken a radically different step by upgrading 'Move' to the status of a "virtual conceptual necessity" of the system. This idea is most clearly articulated in the following quote from Chomsky:

> "Unless some stipulation is added, there are two sub-cases of the operation Merge. Given A, we can merge B to it from outside A or from within A; these are internal and external Merge, the latter [the sic] operation called "Move", which therefore also "comes free", yielding the familiar displacement property of language. That property had long been regarded, by me in particular, as an "imperfection" of language that has to be somehow explained, but in fact it is a virtual conceptual necessity; some version of transformational grammar seems to be the null hypothesis and any other mechanisms, beyond internal Merge, carry a burden of proof". (Chomsky 2005a:12-13)

This makes Move just another instance of Merge. It operates on items that are already part of a given phrase marker unlike Merge or External Merge that adds new items to it. In Chomsky's view, Move or Internal Merge yields different interpretive effects than External Merge; it serves topic/focus and discourse related properties at the C-I interface and can never target theta-domains, i.e. movement within a phrase-marker cannot

be triggered for theta-theoretic reasons.[2] However, the jury is still out on this issue. Hornstein (2001) for instance posits theta-features on heads that must be checked before the structure is shipped off to the interfaces. Theta-feature checking proceeds on lines very similar to regular phi-feature checking, i.e. these features can be checked by either externally merging DPs or by moving DPs from within a phrase marker. Under the second approach, the two types of Merge become even less distinct. Internal Merge seems to pattern alike with External Merge in almost all respects but one - i.e., in the former, computations take into consideration only elements that have already been introduced into the derivation. New elements from the numeration are not accessed by Internal Merge, unlike what we witness for External Merge.

1.2.3 Agree: "An engineering solution"?

The last fundamental operation attributed to UG is long-distance Agree that establishes agreement between two elements in a tree. Functional heads (v, T or C) are introduced into narrow syntax with a set of non-interpretable features that must be valued before the structure containing them are shipped off to the interfaces. Feature-values for these heads are obtained from DPs carrying identical features and this is accomplished via the operation Agree. DPs in turn must carry non-interpretable Case, which gets valued as a reflex of phi-feature checking with respective heads.[3]

The first crucial thing to note about this relation is its asymmetric nature. Only heads or probes can initiate the search for DPs or goals with appropriate features. The search is restricted to the domain of the probe's sister or complement, which is presumably also computationally least expensive. In the words of Chomsky (2005a): "the complement of a head H should be the only domain accessible to operations driven by H, by

[2] In Chomsky's theory, theta-positions also cannot serve as intermediate landing sites.

[3] In GB theory, Agree finds its precursor in the relation government as has been already alluded to. Government is defined as follows (i). (Chomsky 1986b)

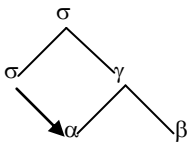
(i) α governs β iff (a) α m-commands β and (b) there is no barrier γ, γ a barrier for β, such that γ excludes α.

α excludes β if no segment of α dominates β.

The relation is at odds with minimalism as it is too unique to FL. Unsurprisingly, several attempts were made to explain all-government related phenomena in terms of other, more-natural sounding principles. One such attempt was to have all case phenomena recast in X-bar theoretic terms (specifier-head, head-complement relations).

conditions of minimal search, the core property of c-command…" (14). Agree is therefore restricted between a probe and a goal that it c-commands, as schematized in (4) where probe σ agrees with α that it also (asymmetrically) c-commands.

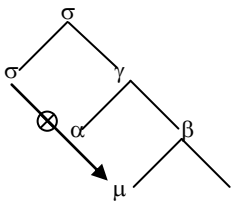
(4)



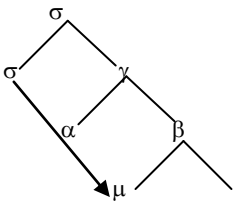
The second crucial thing to note about Agree is its dissociation from movement. Agreement between σ and α takes place long-distance; the goal is not required to move to the domain of the probe. It may however target a specifier of the probe (especially if the probe is also a phase head), but its movement will be triggered for reasons independent of phi-agreement (e.g., an EPP feature).

However, agreement cannot be arbitrarily long-distance. The probe must agree with the closest active (with an unvalued case) goal in its c-command domain. In case the closest DP turns out inactive for Agree, there are two possible scenarios: (i) the search terminates (a.k.a 'defective intervention' as will be detailed in chapter 4) with the probe either receiving a default value or no value at all[4] or, (ii) the probe seeks out the next active goal in the structure. These two options are given below in (5). In (i), probe σ finds α but fails to agree with it due to its inactive status. Even though α is defective for Agree, it blocks σ from seeking out the more appropriate goal μ. In (ii) on the other hand, σ looks past α and agrees with μ instead.

(5) (i)



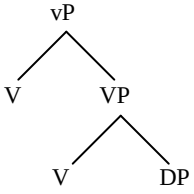
(ii)



[4] Receiving no values of course endangers the derivation at the interfaces.

The question to ask at this point is if an operation like Agree can receive a "principled explanation" like Merge and Move/Internal Merge, i.e. whether something similar is found in other domains and/or if it satisfies interface conditions and follows from general economy considerations. With regard to the first point, phi-features and agreement are generally specific to human language. It is absent from special purpose symbolic systems, which implies that any operation underlying this particular phenomenon must also be specific to FL. We can then ask the next question: does Agree fall out from general economy considerations? As discussed above, Chomsky has argued that search into the complement domain of a head is computationally least expensive. Under this approach, an operation like Agree – where the head agrees with an item which is either its complement or contained in its complement - seems like an optimal choice made by the grammar; its existence at least apparently comports with general economy constraints. However both ideas - (i) only heads must initiate the search and (ii) search into the complement domain is computationally optimal - are completely stipulated. For instance, it is not at all obvious why search cannot be undertaken by DPs seeking heads to check off their non-interpretable case. Bottom-up search, with nominals searching tree-tops for potential 'probes', also appears equally computationally efficient and a perfectly legitimate option that the grammar could use. Consider in this regard the following scenario:

(6)



The tree sketched in (6) has a functional verbal head with a non-interpretable phi-set and a DP with unvalued case that must be checked before the structure containing them are shipped off to the interface systems. In Chomsky's probe-goal system, the optimal situation is where v probes into its complement VP-domain to seek out the DP. A different and equally optimal solution would be one where the DP explores the top of the constructed tree (namely vP) for means to satisfy its case-requirements. In this approach, vP would bear the features of its verbal head, with the result that once constructed, the nominal will immediately seek it out for feature-checking. This option seems computationally as efficient as the first option. In light of this, it is hard to maintain that an

asymmetric probe-goal system like Agree can be justified on grounds of computational efficiency alone.

There is some more ground to believe that the idea of minimal search is not as straightforward as Chomsky would want us to believe. As mentioned above, Chomsky's response to Boeckx's proposal of decomposing Merge further was in the negative. The former contended that merge has no order inbuilt in it and hence must be simpler than simple grouping or concatenation. Secondly, headedness, for Boeckx was unique to human language and defined merge – take X and HP and merge them to get a 'new' object, but one that is projected from one of the constituent items. So to get the unique headedness property of human language, Boeckx's system would need to have labeling alongside concatenation. Quite contrary to Boeckx, Chomsky believes that headedness is a trivial or natural fall-out of the operation Unbounded Merge. In his own words,

> "...when you add head-XP, ... then headedness is a triviality; it comes from minimal search. If the element that you formed, the head-XP, is going to participate in further combinatorial operations, some information about it is relevant, and the simplest way to find the information – minimal search for the information – will be to take one of the two objects. Well, one of them has no information, because you have to find its head, and that is too deep down, so you find the other one. So the trivial consequence of an optimization procedure (presumably non-linguistic and non-organic, or maybe the law of nature) is in H-XP, take H." (Chomsky 2009: 52).

Let us try to unwrap the idea of minimal search emanating from the above paragraph. Chomsky believes that if H and XP merge and the new object formed has to continue merging with more items, generating recursive structures, then the information that is most readily available and most easily accessible is that provided by the head H. That information is 'percolated' up to the dominating node, which is thereby of the same type as the constituent item H. This trivially defines the headedness property of natural language. Looking into the XP constituent for the information is not optimal since its head is deep down and hence requires a non-local and less economical search. However note that if this is right, Agree that allows probes to look deep down into their complement's domains as long as there are no intervening items, in principle, is allowed by the grammar to use a non-optimal search procedure. Why Agree should have this special privilege when the system otherwise shows optimization elsewhere is a puzzle that is left unsolved in Chomsky's system. General economy considerations are one of the important hallmarks of minimalist inquiry and yet Agree seems to work on quite different and unexpected terms.

Moreover, Chomsky's quote also implies, very explicitly, that the top of the tree, which is as the same type as the head, is the most crucial information carrying node, one that holds the potential to keep the structure growing in the computational workspace. In that sense, searching the topmost node appears as the most economical thing to do when carrying out computations. External Merge seems to look for topmost nodes (very reminiscent of Chomsky's Extension Condition) so that it may have newer heads merged with older, already constructed phrases. Given that movement is nothing but another kind of Merge – internal Merge, it is expected that it too must adhere to the same economy principles in the process of building structures. If a head or an XP, computationally active, has to displace, it must, following general principles of economy, be allowed to access the top of the tree and target the nearest vacant position. This vacant position could turn out to be a head for the moving head or a specifier for the XP.

Other problems for Agree come from its dependence on c-command.[5] First, Chomsky himself identifies the dubious role of relations like c-command in narrow syntax: "The relation c-command is available, and expected, on very weak assumptions. The relation has played a large role in syntactic theory, though it may be that it does not function within narrow syntax but only in interpretation of the information it provides – that is, in mapping it to syntactic objects that belong to mental systems external to the language faculty itself." (Chomsky 2000:31) If true, then c-command, quite contra what one is led to believe while postulating an operation Agree in the system - should be irrelevant to all and any narrow syntactic phenomena. This makes it doubtful that an operation controlling phi-agreement in narrow syntax can be grounded in c-command like Agree is. Long-distance Agree therefore seems theoretically unrealistic when seen from this light.

Additionally, several researchers (cf. Epstein, Groat, Kawashima and Kitahara 1998; Hornstein 2005) have argued that c-command is not a primitive of the grammar. Rather, it is a direct consequence of the way the grammar operates. Below I summarize the main results of these two studies and then turn to show what implications they might hold for the operation Agree.[6]

---

[5] Unless specially mentioned, c-command is used here to imply asymmetric c-command. Mutual c-command is referred to as sisterhood.

[6] See Frank and Vijay-Shanker (1999) and Frank, Hagstrom and Vijay-Shanker (2002) who claim that c-command is a primitive of the grammar, from which are derived relations like dominance and notions like roots and constituents. However simply restating these relations in terms of c-command does not necessary imply

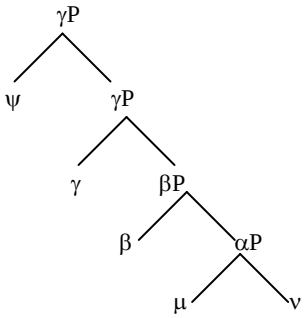
As was most clearly noted in the works of Epstein et al (1998), c-command falls out naturally from the derivational workings of the system. They dub it 'derivational c-command' (7):

7) ***Derivational C-command***

X c-commands all and only the terms of the category Y with which X was paired/concatenated by Merge or Move in the course of the derivation.

The assumption is that structures are built in a bottom-up fashion. As elements are introduced into the workspace, they end up derivationally c-commanding whatever items they are merged with and also their 'terms', i.e. all elements contained in them. Take (8) for illustration.

(8)



As in (8), head β, once merged with αP derivationally c-commands αP and all the elements that αP dominates (namely μ and ν). In a similar vein, when γ is merged with βP, it ends up derivationally c-commanding βP and all the elements that it dominates, namely head β, αP and its terms μ and ν. In Epstein et al's view, c-command is a relation that falls out naturally from the way Merge operates; it is not a primitive of the grammar. The problem with their account, as Hornstein (2005) explicitly recognizes, is that even though they justify c-command as a natural relation, Epstein et al run short of providing an explanation for why grammar chooses c-command for its computations, instead of some other relation(s) that also fall out from Merge. As briefly alluded to in section 1.2.1, there are at least two relations that are directly deduced from Merge (i) 'sisterhood' (between μ and ν in (8)) and (ii) 'immediately contain' (as between αP

---

that the latter is not derived from other properties of natural language (see below for more).

and μ). Ideally all syntactic phenomena must be restricted to these two configurations, which are most 'natural' in that they are provided for free by the indispensable operation Merge (more on this later). The elementary operation of composition of relations gives us some new relations (cf. Chomsky 2000). I have already talked about c-command, which could be interpreted as composed from 'sister (contain)'. In addition, there are numerous other relations that could be composed indirectly from Merge, including those corresponding to the transitive closure of 'immediately contain' (i.e. 'contain' as between βP and μ), 'sister (sister)' or the relation of 'identity' (eg. the sister of μ is ν whose sister is μ itself), 'immediately contain (sister)' (as between γ and ψ) or even 'immediately contain (sister) (immediately contain) (sister)' among others. Nothing in principle prevents the grammar from exploiting these relations as they are all indirectly obtained from Merge. In light of this, why c-command must be chosen for grammatical computations is a question that remains largely unresolved in Epstein et al's work. Hornstein's take on this issue, as I discuss below, is that the correct answer to this puzzle cannot be sought in the maneuverings of External Merge alone. Internal Merge and its application in different sub-theories are largely responsible for the systems's apparent reliance on c-command.

Hornstein considers each of the domains where c-command applies and demonstrates how by adopting a slightly different approach, c-command "pops out in a system that uses simple natural primitive operations, and deploys them in a computationally optimal way" (pp.28). Here I concentrate on his alternative analysis for Condition A of the Binding Theory, which also holds an answer for why c-command may be relevant in this domain. Consider:

(9) John-i likes himself-i.

Sentence (9) is a typical instance of anaphora binding. The reflexive *himself* must be A-bound (c-commanded and co-indexed) with a clause-mate antecedent *John*. The question we have to ask here is why c-command must play such an important role in anaphora binding. Hornstein's alternative analysis for binding (that also helps derive c-command in this domain) bases itself on two main assumptions: (a) computations are monotonic increasing (i.e. they obey the Extension Condition that permits adding material only to the roots of trees) and (b) movement or internal merge can also target theta-domains. Once we adopt (a)-(b), (9) is given the following underlying representation (10), with the relevant derivational steps sketched in 11(i)-(iii).

(10) John-i likes John-self-i.
(11) (i) [V likes John-self]
(ii) [vP John [v [V likes __-self]]]
(iii) [TP John [vP John [v [V likes __-self]]]]
(iv) [TP John [vP John [v [V likes himself]]]]

The derivation – as shown in successive steps in (11) - starts off with merging *John* with *self* as a complement to lexical V as in (i) and then moving it via the specifier of vP for a second theta-role (ii) to the specifier of TP for case and EPP (iii). The stranded affix *self* finally ends up getting a case from V. An overt pronoun is merged with *self* to avoid a violation of the Stranded Affix Filter in (iv).

In this alternative account for binding, c-command emerges naturally from the interplay of core grammatical operations like Move and conditions on computational efficiency. The reason why antecedents must c-command co-indexed reflexives is not because this relation is a primitive of the grammar or that it is built into the definition of binding, but rather derives from the fact that the antecedent itself must target the root of a given phrase marker (Extension Condition) for case/phi-feature requirements and in the process ends up in a specific relation (asymmetric c-command) with the copy it leaves behind.[7] In a nutshell, Hornstein demonstrates that not only is c-command not a primitive of the grammar, but that every syntactic phenomenon – where it apparently does play a role – relies crucially on the core transformation Move. C-command is merely an epiphenomenon that the grammar generates through the interplay of its core grammatical operations.[8]

There are two immediate implications that the above discussion might have for the operation Agree. The first is rather obvious: since Agree is dependent on c-command, and c-command is not a primitive relation, Agree (if it exists at all) is not a primitive transformation, unlike Merge or Move. Second, having c-command derived from Move, we are also forced

[7] Similarly for linearization and movement (where a necessary condition is that the higher copy alias the moved element c-command the lower copy or trace), the relation c-command comes into effect primarily for the same reasons that I discussed above. Elements must target the root of a phrase marker; this creates a configuration where the moved element asymmetrically c-commands its copy/trace.

[8] I must emphasize that this does not imply that c-command is not used in any grammatical phenomenon. Rather, all that can be drawn from the above discussion is that it is a consequence of narrow syntactic computations. Its use in various domains, for instance scope (see chapter 2 for movement driven scope effects) and binding is nowhere undermined by this approach to c-command.

to ask if this relation is actually essential for agreement; i.e., if it is necessary that there must exist a c-command relation between two items for them to agree with each other. This is a relevant question especially in light of Hornstein's observations that in every domain where c-command applies, it is derived from movement or displacement of the 'c-commanding' item. Given this, it is natural to question: if c-command applies in phi-computations, what derives it? I will take it one step further and ask in even bolder terms: is c-command even relevant for phi-feature computations? Previously I have noted that the 'complement domain as computationally efficient' argument for Agree is weak and cannot be maintained but as a stipulation. It has never been fully justified by Chomsky. A head probing a goal that it asymmetrically c-commands is not more computationally optimal than a nominal ('goal') seeking the root of the tree for feature checking. If this is right, then the answer to our second question is that c-command is not necessary to establish phi-agreement. There is nothing in the grammar that forces an asymmetric c-command relation between agreeing heads and DPs. For Agree, the implication is that it cannot be intrinsic to the grammar.

However, despite all these explicit conceptual conundrums related to its postulation, there exists a widespread acceptance of the notion of Agree, mainly based on empirical facts that I introduce below. It is therefore pertinent to ask if Agree remains merely as an engineering solution for these facts.[9,10]

1.3 Phi-agreement and the sisterhood relation

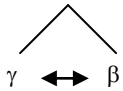
But before I proceed to discuss the empirical base for Agree, I must address the following question: if not Agree, then what must be the grammatical operation that underlies phi-agreement? From our discussion

[9] There is a second agreement relation 'concord' in the grammar. Concord holds locally between an NP and its determiner, adjective phrases etc. I will have nothing to say on this operation here, though I refer to it to at certain points of my analysis. I will simply assume that concord involves feature sharing between two items that do not fit into the probe-goal category. The details of this operation are left pending.

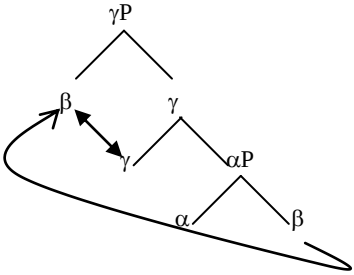
[10] Miyagawa (2009) proposes a novel approach to agreement, which is closer to my account here. He suggests that agreement brings the functional and the lexical domains together. He essentially tries to capture the Sportiche-Koopman intuition that agreement happens in local spec-head relations by suggesting that agreement happens in the grammar to establish functional relations between two otherwise disparate domains.

so far on minimalist goals, the optimal solution seems to be Merge or Move. Being core operations of the grammar, they must be able to determine all narrow syntactic phenomena, including agreement. If true, agreement must be carried out in one of the following configurations:

(12) (i)

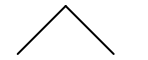


(ii)

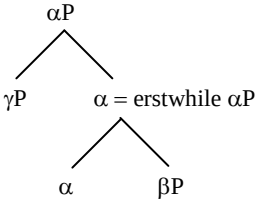


The diagram in 12(i) shows a head-complement relation formed by externally merging β to γ, whereas (ii) has γ and β in a specifier-head relationship, established by internally merging the latter into the domain of the former. I have noted previously that sisterhood emerges for free from Merge. That agreement must utilize this relation is therefore unsurprising. The status of a specifier-head relation is however much more controversial, as it necessitates positing a different relation – m-command – in the grammar. M-command does not follow directly from Merge (though a composition of relations can indirectly deduce it) and hence, its utilization for phi-agreement needs to be motivated. In this regard, 12(ii) is at least apparently not an ideal configuration for agreement. A possible solution to this problem is available once we adopt the basic tenets of Bare Phrase Structure Theory. According to Chomsky (1994), narrow syntax recognizes only maximal and minimal projections (X-0s and XPs). Intermediate bar-levels (X's) are ignored for narrow syntactic computations. A category that does not project further is a maximal projection while a category that is not a maximal projection is a minimal projection. Schematically, 13(i)-(ii). αP is the maximal projection in 13(i) after merging α with βP, but once a new item γ is added to αP, it transforms itself back to a minimal projection α as in 13(ii).

(13) (i)

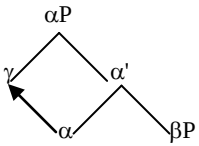


(ii)



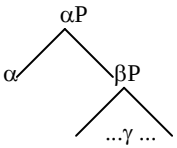
Once we seriously adopt these ideas, relations like m-command become superfluous. To put it in more concrete terms, relations like m-command work in a system where narrow syntactic computations first take into consideration intermediate bar-levels and secondly, only the original head carries the relevant features. This is shown in (14).

(14)



However, in a system where there are no bar-levels and features are percolated up to whatever is the next highest projection of that head, relations like m-command become irrelevant. For illustration, consider (15). In (i), γ targets the root of the phrase marker αP, (which carries the appropriate features) and subsequently merges with it as depicted in (ii). Once αP merges with γ, creating a larger structure in the process, it is once again deemed by narrow syntax as a minimal projection. It is this head, I assume, that enters into agreement with γ, and not the original head. Importantly, this implies that the actual agreement process takes place between a DP and a head and not the maximal projection of that head.

(15) (i)



(ii)