Locality and (minimal) search[1]

Kenyon Branan, Michael Yoshitaka Erlewine

branan@leibniz-zas.de, mitcho@nus.edu.sg

April 2022

For the *Cambridge Handbook of Minimalism*, Kleanthes K. Grohmann & Evelina Leivada (eds.)

1 Introduction: Declarative vs procedural descriptions of locality

Languages establish many "dependencies" between different parts of an utterance. For example, a verb might inflect for the φ-features of a particular argument, or an interrogative complementizer may attract a *wh*-phrase to its specifier. Operations of both agreement and movement are triggered by *probes*, which are themselves formal features on heads, like other structure-building instructions in the Minimalist Program. Probes trigger a search for a *goal* with a particular feature specification, which then feeds some interaction with the identified goal, such as the exchange of feature values or movement/attraction of the goal.[2]

Locality refers to the question of what configurations between probe and goal are allowed in grammar, as reflected empirically in what configurations of non-local dependencies (e.g. agreement and movement) are attested in languages of the world. Consider the case of *wh*-movement in a language such as English, which requires the movement of a single *wh*-phrase to the clause edge. In probe-goal terms, C bears a probe which seeks an active [WH]-bearing goal — which we annotate [PROBE:WH] for perspicuity[3] — and moves this goal to its specifier. In a situation such as (1) where there are multiple appropriate goals, the "closer" one must be moved (*Superiority*; see Kuno and Robinson, 1972; Chomsky, 1973).

[1] We thank Amy Rose Deal, Nick Huang, Alan Hezao Ke, Stefan Keine, Diego Krivochen, Omer Preminger, an anonymous reviewer, and the volume editors for helpful comments and discussion. We also especially thank Keng Ji Chow and Cara Leong for thoughtful comments on an earlier version of section 2. This research is supported by the Singapore Ministry of Education under grant MOE2017-T2-2-094 and by the National University of Singapore under grant R-103-000-178-133.

[2] Probes may be involved in other dependencies too, but here we concentrate on agreement and movement. In Chomsky 2000, 2001 and much subsequent work, Move/Attract is thought of as always parasitic on an earlier step of Agree between the attracting probe and the moving phrase. However, see e.g. Preminger 2014: ch. 8 for arguments against Agree being a precondition for movement. Here we simply describe both agreement and movement as reflecting a shared underlying process of "probing," which involves a search procedure. Differences in the locality profiles of agreement versus movement, if any, may result from differences in the probe specifications involved (see section 3) or other differences between these operations; see Bošković 2003, 2007; Bobaljik and Wurmbrand 2005; Fox and Pesetsky 2005; Richards 2012, among others.

Note too that, following recent work on *labeling* (Chomsky, 2013 *et seq*), *local* dependency formation — e.g. selection — may also be thought of as reflecting the results of this shared search process. See Ke 2019 ch. 2 for relevant discussion.

[3] Many authors following Chomsky 2000, 2001 describe probes as *uninterpretable* features, notated [uF], which are required to find a match (also: be checked or valued) for the derivation to converge. However, subsequent work has shown that a derivation can converge with one of its probes failing to find a match (see especially discussion in Preminger 2014), undermining the description of probes as due to "uninterpretable" features. Here we avoid such reference to "uninterpretable" features and [uF] notation in our description of probes, also following Erlewine 2018 and Deal to appear.

Chomsky (2001: 5) also claims that a feature is uninterpretable if and only if it is *unvalued*, but see Pesetsky and Torrego 2007 for further discussion of this distinction.

Many authors have discussed the conditions under which a potential goal is "active" and thus a licit target for Agree. See e.g. discussion in Chomsky 2001: 6ff and Bhatt 2005: 802ff. This question is orthogonal to the discussion here.

(1) ***Wh*-movement attracts the "closer" *wh*-phrase:**

C[PROBE:WH] you expect *who* to buy *what* ⇒

a. Who did you expect ____ to buy what?

b. *What did you expect who to buy ____?

The description of *relative locality* effects, beginning with Rizzi's (1990) *Relativized Minimality*, have overwhelmingly taken a *declarative* form, stating constraints such as (2):[4]

(2) **Closest:** A probe must target the closest goal. A potential goal G for probe P is *closest* if no other potential goal for P c-commands G.

This declarative constraint in (2) helps us (as analysts) explain the impossibility of a hypothetical derivation where the probe attracts the further goal *what* in (1), and therefore predict the ungrammaticality in (1b). The constraint, so formulated, can also be thought of as a *filter* on the results of structure-building operations that apply freely (see e.g. Lasnik and Saito, 1992), potentially as a result of transderivational competition as in the discussion of the Minimal Link Condition in Chomsky 1995: ch. 4 and Nakamura 1998.

However, around the turn of the century, Minimalist theorizing shifted towards explaining constraints on grammatical configurations as reflecting the behavior of syntactic operations, in more *procedural* terms. In the case of locality, such effects — such as the Superiority contrast in (1), as well as others we discuss below — have been thought to reflect the fact that probing involves an operation of "minimal search" (Chomsky, 2004: 113, *et seq*), where "minimal" means that the search procedure stops once it has found an appropriate goal (see also discussion in Ke, 2019: ch. 2). In (3), we sketch how the effects of the *declarative* constraint of Closest in (2) could plausibly follow from this *procedural* description of minimal search in (3):

(3) **Deriving Closest from "minimal search" (a sketch):**

A probe triggers a *search* for a goal — *using a particular search procedure* — and *stops* once a suitable goal is found. Therefore, nodes that "come after" the first suitable goal — based on the order in which nodes are considered by the search procedure — are never even considered as possible goals.

In particular, suppose the search procedure has the following property: For X and Y in the *search space* of probe P, if X c-commands Y, search initiated by probe P will consider node X before node Y. If this

[4] The statement in (2) is equivalent to the locality condition on Agree stated in Chomsky 2000: 122, assumed there to also be relevant for the behavior of movement, such as in (1); see note 2. Combined with a requirement that a probe must c-command any potential goal, (2) also becomes effectively equivalent to the formulation of Relativized Minimality from Rizzi 1990: 7, also restated in more contemporary terms in Rizzi 2004: 225.

property holds, we derive — and thus explain — the declarative generalization of Closest in (2) as a consequence of this more general nature of probing as minimal search.

In this chapter, we detail and discuss this *procedural* approach to the description and explanation of locality constraints in syntax.[5] We begin in section 2 by discussing possible formulations for the underlying search procedure. We then discuss empirical challenges for the "minimal" quality of the search procedure and introduce Amy Rose Deal's theory of *interaction vs satisfaction* as a promising extension to the procedural theory of probing, in section 3. We discuss the shape of the search space itself in section 4. We conclude in section 5 with reflections on the status of procedural explanations in syntactic theory.

2 Defining the search procedure

A procedural approach to locality effects requires us to make explicit the mechanics of the search procedure underlying probing. In particular, we care to identify the *order* in which a probe considers nodes within its search space, and tests them to see if they are a suitable goal for the probe. This ordering — combined with the "minimal" property, that probing terminates once one appropriate goal is found — may underly and explain *relative locality* effects, as per the logic in (3) above.

Here we first sketch and discuss two well-established options for the search procedure: *depth-first* versus *breadth-first* search.[6] Both procedures first check to see whether or not the *start node* — i.e. the root node of the *search space* — is a match for the probe. If not, and the start node has daughters:

- The depth-first algorithm chooses one of the start node's daughters to consider, to see if it is a match; if not, the search then considers one daughter of that node, and so on. If the depth-first algorithm reaches a non-branching node without returning a match, the algorithm backtracks just enough to a node with a not-yet-considered daughter, and repeats the process on that daughter, etc.
- The breadth-first algorithm checks if each daughter of the start node (call these nodes *depth 1*) is a match; if no match is found, the algorithm checks the daughters of each of the depth 1 nodes, i.e. each of the *depth 2* nodes, and so on.

The search proceeds until a match is found — the "minimal" property — or once all accessible nodes in the search space have been considered. In the latter case, the search will terminate with no match, which we take to be possible and not necessarily lead to ungrammaticality (see e.g. Preminger, 2014).

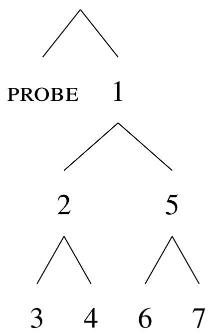
[5] The distinction we draw here between *declarative* and *procedural* descriptions is reminiscent of the distinction between *representational* and *derivational* modes of description and explanation. But declarative constraints such as (2) could be evaluated during the course of the derivation or only over final representations, and therefore the distinction we make here is not equivalent to questions of representations vs derivations. We therefore avoid discussion in terms of representations vs derivations here. See relevant discussion in the introduction to Epstein and Seely 2002.

[6] Our discussion here describes probing as a process that proceeds *downward* through the search space, as is a common assumption in current work. This top-down conception of probing has recently been challenged, in particular by Bjorkman and Zeijlstra (2019) based on facts in the domain of φ-agreement. See also Kush 2013: 21–23 for a concrete description of upwards search. We will not review this discussion here, although see note 30 below for a potential derivation of "upwards Agree"-like behavior using repeated applications of the search procedure described here, together with iterated upwards expansion of the search space.

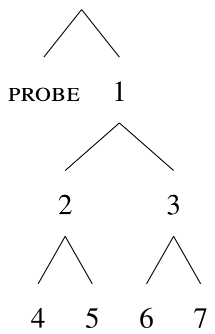
See Atlamaz 2019: ch. 3 and Ke 2019: ch. 2 for more detailed descriptions of depth-first and breadth-first search algorithms and discussion of their relevance for the notion of minimal search.[7] In recent work, Chow 2022 proposes a novel search algorithm which is distinct from depth-first and breadth-first search, but we first concentrate on these two more established algorithms and then briefly comment on Chow's proposal below.

The two search procedures are illustrated in (4–5) below for the case of search within an internally complex sister of the probe. Here we assume that sisters are ordered for search, but briefly return to this assumption later.[8] The numbering on the trees below reflects the order in which the nodes are considered by the probe, with nodes with lower numbers considered before those with higher numbers.

(4) **Depth-first search, left to right:**



(5) **Breadth-first search, left to right:**

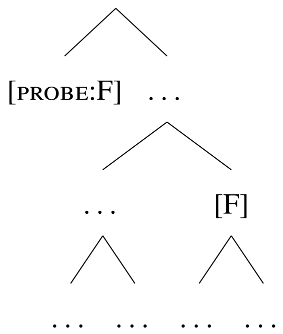


Suppose the probe seeks a goal which bears the feature [F] and there is just one node bearing [F] within its sister, as in (6). Both the depth-first and breadth-first algorithms allow for the probe to find and interact with this goal.

[7] Various authors have discussed tree traversal algorithms in detail for linearization purposes. The appendix of Kremers 2003 discusses depth-first versus breadth-first search and concludes that depth-first search is more appropriate for his approach to linearization, also summarized in Kremers 2009. The algorithms described in Yasui 2003 *et seq*, Kural 2005, de Vries 2009 are all depth-first, but vary between so-called *preorder*, *inorder*, and *postorder* traversal, which is an independent way in which tree traversal algorithms may vary. See these works for definitions and discussion. The search procedures we describe here are preorder traversals.

[8] The algorithms as described here, and illustrated in (4–5) below, require an ordering to be imposed on the structure to determine which of two daughters is to be considered first. One possibility is that something like the LCA (Kayne, 1994) is responsible for ordering structures for this purpose. This relative ordering between daughters may or may not be the same as a linearization. It is possible that the *raison d'être* for this ordering is simply to render a structure suitable for the search procedure; the interfaces then may or may not subsequently make use of this ordering for their own ends, e.g. for linearization at PF and also for processes at LF (Kayne 1994: §5.2, Bruening 2014, Branan and Sulemana 2019). See also Atlamaz 2019: 89–90 for a similar suggestion.

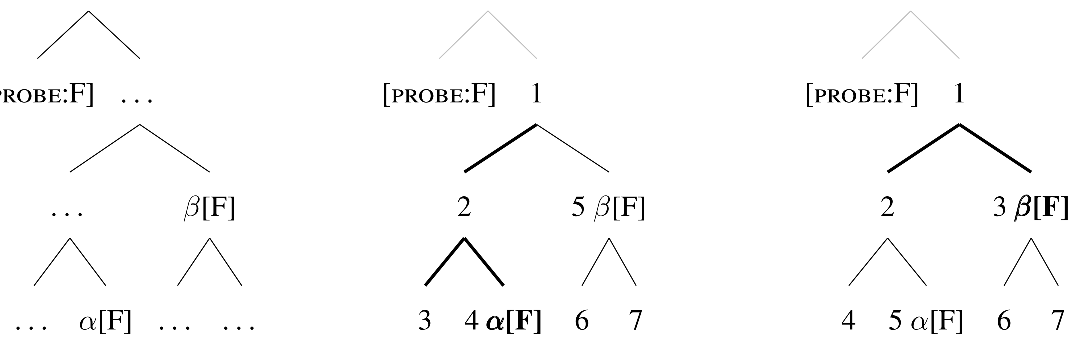
(6) **Probing for a unique potential goal:**



But now consider the configuration in (7) with two nodes with the feature [F] in the search space, labeled α and β. Here the two search algorithms yield different results. The depth-first search in (7a) descends down left daughters to node 3, resulting in no match; backtracking to node 2 and considering its other daughter α leads to a match, and the search terminates, never considering the nodes to the right. The breadth-first search in (7b) considers the start node (1) and its left daughter (2) before matching with β and then terminates, never considering the nodes deeper down. The paths to nodes that are considered in each search are bolded; the structure above the start node is gray as it is not part of the search space.

(7) **Probing with two potential goals:**

a. Depth-first, left to right: b. Breadth-first, left to right:

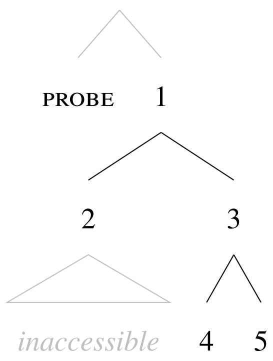


The illustration in (7) demonstrates that depth-first search will potentially lead to violations of the declarative Closest requirement on goals in (2). In (7a), the probe matches with α, despite the potential goal β c-commanding it. This is the case because α is within the left daughter of the start node, whereas β is (within) the right daughter of the start node, and we have chosen to follow a left to right search order at each level. This also highlights the outsized effects of the relative search order among sisters for depth-first search.

In contrast, breadth-first search in (7b) leads the probe to match with β, which c-commands α. As the minimal search terminates after finding β, α is never even considered by the probe. More generally, Ke 2019 demonstrates that a breadth-first search algorithm with unordered sisters derives the effects of the c-command-based Closest condition in (2).[9][10] For this reason, Ke proposes to identify the "minimal search" procedure of probing as a breadth-first search procedure.

We believe it may be premature to identify probing as involving breadth-first search as Ke proposes, for three reasons. First, it is possible that in many common grammatical configurations, depth-first and breadth-first searches are more difficult to distinguish than in the neat, abstract demonstration in (7). In particular, suppose that certain subparts of the search space are inaccessible for probing. (We discuss this general possibility in more detail in section 4.1.) Example (8) illustrates one such modification to the structures considered above, where the contents of the left daughter of the start node (possibly a specifier or adjunct) is made inaccessible for the search procedure. Notably, both depth-first and breadth-first searches now result in identical search order across the remaining, accessible nodes in the search space:[11]

(8) **Same order with depth-first and breadth-first, due to one part being made inaccessible:**



More generally, if the search space is strictly right-branching — binary branching with no accessible daughters under left sisters — depth-first and breadth-first searches yield provably equivalent results for search order and, in turn, relative locality constraints. Note that grammatical structures may indeed often take this form if the contents of (left) specifiers and adjuncts are generally made inaccessible for probing, as has been independently proposed in work such as Nunes and Uriagereka 2000 (see discussion in section 4.1 and note 23 below), inspired by Huang's (1982) Conditions on Extraction Domains.

Second, depth-first search may be motivated by derivations that involve *smuggling*, where it is important that elements contained within a derived specifier be preferred for movement over other

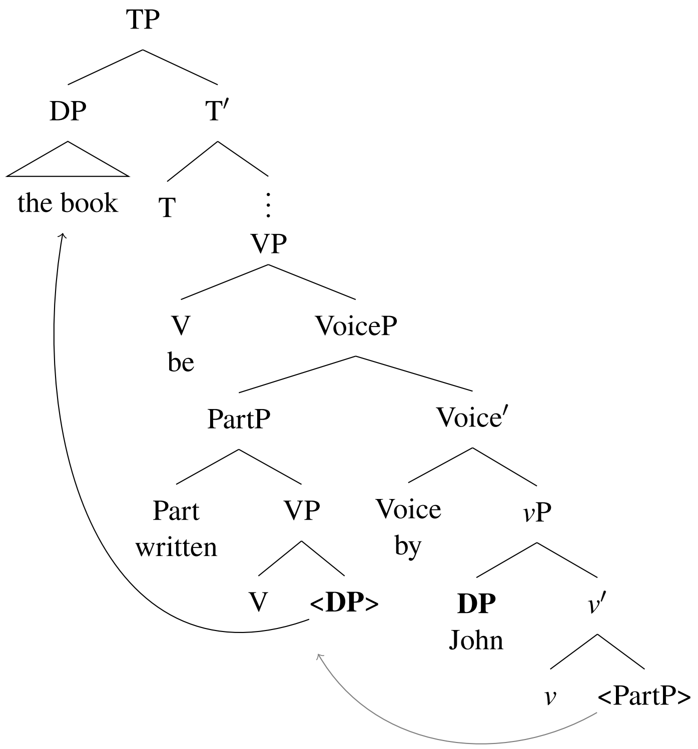
[9] The procedure with ordered sisters, illustrated here, derives the Closest condition for potential goals at different depths, but predicts left-right asymmetries at the same depth level. For example, in (7b), the node labeled 2 will block a match with 3 but not vice versa.

[10] Concentrating on its application for labeling (see note 2), Cao (2017) briefly describes minimal search as a breadth-first search.

[11] A reviewer suggests that the search procedure could potentially make reference to the different status of nodes in the structure. For example, the search algorithm may recognize certain nodes as adjuncts and therefore not consider them. Preminger (2019: 24) offers a recent proposal for minimal search along these lines, which makes reference to the status of different nodes in the structure as specifiers, complements, or adjuncts of a head, and which does not invite a clear classification as depth-first or breadth-first. Together with Atlamaz (2019), Ke (2019), and Krivochen (2021), we instead think it is productive to consider options for the search procedure in the most general case, without reference to such information, and to then consider the shape and size of the search space separately, as we discuss in section 2 below.

potential matches in that specifier's sister. For example, Collins 2005 proposes that in the derivation of an English passive as in (9), a portion of the extended verb phrase containing the theme (participle phrase, *PartP*) is first fronted across the agent but *not* "frozen," thereby "smuggling" the theme across the agent. It is crucial for this derivation that subject movement triggered by a probe on T target the theme DP within the fronted PartP rather than the agent in Spec,*v*P. Note that, in (9), the theme DP and agent DP (in bold) are at the same depth level within VoiceP.

(9) **Smuggling derivation for *The book was written by John* à la Collins 2005:**[12]



On an approach with depth-first search that traverses (left) specifiers before their sisters (see footnote 8 above), all elements contained within a (left) specifier (as long as its contents have not been made inaccessible) will be "closer" to a probe than any elements in the sister of that specifier. However, this is not generally the case on a breadth-first algorithm, which may incorrectly predict the availability of theme DP movement as in (9) to be sensitive to minute differences in the relative depth of the theme DP versus the agent DP.[13]

Third, while Ke and other authors discuss only breadth-first search and depth-first search as potentially

[12] This complete structure never appears explicitly in Collins 2005, but we understand this to reflect his final proposal. Following his argumentation for PartP movement (§3; see (22) on page 90), he argues for the analysis of *by* as a Voice head (§4; see (30) on page 95), reflected here. The theme subject moves first to the edge of PartP in (22) on page 90, but following discussion of the non-phase status of PartP on page 98, the theme DP is illustrated as moving directly from its base position within PartP in (46) on page 102. Finally, we note that Collins refers to Infl/IP for the projection here labeled TP.

[13] But we might also avoid this particular concern in the case of Collins' passive derivation, if Voice is a phase head which leads to *v*P being inaccessible for probing from above, as Collins suggests (page 98).

underlying minimal search, these two well-established algorithms are not the only possibilities. In particular, Chow 2022 has recently described a search algorithm that brings together some of the positive features of both breadth-first and depth-first search algorithms that we have reviewed above. Like breadth-first search but unlike depth-first search, Chow's search algorithm ensures that probing will satisfy the c-command-based Closest condition in (2).[14] But like depth-first search and unlike breadth-first search, Chow's search algorithm will exhaustively search within a specifier that is transparent for probing before considering the contents of the specifier's sister, thereby allowing for smuggling derivations of the form we describe above. We believe that Chow's algorithm is thus a strong candidate for adoption as the technical implementation for minimal search and refer readers to Chow 2022 for further details.

3 Distinguishing interaction and satisfaction

In the search procedure described above, the probe considers one node at a time to see if it is a match for the probe's feature specification; if it is a match, the search halts. The probe may then exchange information with the goal (leading to agreement and/or case-assignment) or attract the goal (movement). As discussed in particular by Ke (2019), this property of the search halting immediately upon encountering a match is the "minimal" property of *minimal search*, which allows for the procedural explanation for relative locality effects such as that described in (2).

A challenge for this "minimal" property is the existence of probes which may target multiple goals. In this section, we introduce this empirical possibility and the proposal in Deal 2015, to appear to address such facts, also informed by ideas in Keine 2016. Herself adopting a procedural description for probing and locality effects, Deal proposes that probes have two distinct specifications: an *interaction* condition, identifying the type of goal(s) the probe will find to later Agree with or to Attract, and a *satisfaction* condition, for what nodes will cause the search to halt. As we will see in this and the following section, such a dissociation allows for a richer typology of probing behaviors, which appears to be empirically motivated, but effectively undercuts the strong claim that all probing is minimal search.

Deal's original motivation for the distinction between interaction and satisfaction conditions comes from the behavior of Nez Perce complementizer agreement. In the general case, complementizers in Nez Perce inflect for first-person, second-person, and plural features, in case either the subject or object bear these features. This results in the same complementizer form *ke-pe-m* expressing plural and second-person features in examples (10a–c), where the plural and second-person features are both on the subject (10a), both on the object (10b), or contributed by two different arguments (10c). However, specifically when the subject is second person, there is no agreement with features of the object, for example explaining the lack of plural agreement on C with the object in (10d).

[14] There will, however, still be left-right asymmetries between sisters. See footnote 9 above.

(10) **Nez Perce complementizer agreement:** (Deal, 2015: 7–8)

a. ke-pe-m kaa *pro*$_{subj}$ 'e-cewcew-tée'nix A.-ne
C-PL-2 then PRO.2PL 3OBJ-telephone-TAM A.-ACC
'when you(pl) call A.' (2pl subject / 3sg object)

b. ke-pe-m kaa A.-nim hi-cewcew-téetu *pro*$_{obj}$
C-PL-2 then A.-ERG 3SUBJ-telephone-TAM PRO.2PL
'when A. calls you(pl)' (3sg subject / 2pl object)

c. ke-pe-m kaa A.-nim kaa T.-nm hi-cewcew-tée'nix *pro*$_{obj}$
C-PL-2 then A.-ERG and T.-ERG 3SUBJ-telephone-TAM PRO.2SG
'when A. and T. call you(sg)' (3pl subject / 2sg object)

d. ke-m kaa *pro*$_{subj}$ 'ee 'e-nees-cewcew-téetu *pro*$_{obj}$
C-2 then PRO.2SG 2SG.CL 3OBJ-O.PL-telephone-TAM PRO.3PL
'when you(sg) call them' (2sg subject / 3pl object)

Deal shows that this and other such asymmetries all follow from the generalization that "C does not probe past a second person argument" (Deal, 2015: 8). Deal proposes that the probe on Nez Perce C will *interact* with all φ-feature targets — exponing the first-person, second-person, and plural features of all such targets it finds on core arguments in the clause — but is *satisfied* specifically by the addressee feature [ADDR], where satisfaction refers to termination of the search. In other words, the probe will terminate only when [ADDR] is found or when the search space is exhausted. Deal (to appear) proposes the notation [INT:φ, SAT:ADDR] for this probe.

We might also imagine there to be cases of *insatiable* probes — i.e. probes which have no satisfaction condition, and which interact with any and all elements that match the interaction condition within the search space. One possible case of this is Japanese T, following the analysis of long-distance multiple nominative assignment developed in Hiraiwa 2001. Hiraiwa argues that the presence of finite T in a matrix clause in Japanese is able to license multiple nominative arguments, crossing both a finite clause boundary as well as intervening dative nominals. He proposes that certain heads — such as T in Japanese — may be specified to probe for multiple goals as part of a single Agree operation. For the theory sketched here, such a probe could be described as a probe without a satisfaction feature; for such probes, search would terminate only after the entire search space is exhausted. See also Deal to appear and citations there for discussion of other insatiable probes.

The possibility of a probe matching with multiple goals opens up a more general question of how the morphosyntax then handles the output of such a process. "Omnivorous" agreement of the kind observed in Nez Perce and multiple assignment of nominative case in Japanese are just two possibilities. In the domain of φ-agreement, see further discussion in Deal 2015: 11–13, as well as discussion of this question in relation to PCC effects in Coon and Keine 2021 and Deal to appear. Another possibility is that, following the identification of multiple goals by a probe, an independent heuristic is used to choose just one of these goals to then agree with or to move. See for example the *Specificity*-based proposal in Lahne 2012 and Hamann 2014, *Best Match* in Coon and Bale 2014, *Multitasking* in Van Urk and Richards 2015, as well as multiple matches simply leading to optionality in Halpert 2019. Note however that by allowing probes to first interact with multiple candidate goals and then later choosing just one of them to privilege for visible interaction, we effectively undo the core procedural explanation for relative locality effects as in (3) above. Probing with the "minimal" characteristic described above is then guaranteed only when the interaction and satisfaction conditions are equal: a probe of the form [INT:F, SAT:F] will terminate immediately upon matching with a [F] goal, if any, or else terminates after exhausting the search space.

In the case of Nez Perce complementizer agreement above, the satisfaction feature of the probe will match a subset of the nodes that interact with the probe.[15] But following Keine 2016, we also imagine there to be cases where a probe can be "prematurely satisfied," i.e. terminating without any identified goal. In Branan and Erlewine to appear, we discuss the fact that many languages exhibit Ā-movement processes that necessarily target the closest DP, often resulting in a descriptively subject-only extraction restriction. On the approach developed there, a probe may interact with nodes that bear both a [D] feature and a relevant Ā-feature, with [D] alone being a satisfaction feature. We notate such a probe as [INT:Ā+D, SAT:D]. Search by a probe of this form will necessarily terminate upon encountering the closest nominal.[16] If that closest nominal bears the requisite Ā-feature, the probe will interact with it, as reflected by its successful movement. However, if that closest nominal lacks the relevant Ā-feature, probing will terminate without successfully matching with any goal. The end result is a Ā-extraction process that cannot skip the closest DP.[17] See also Keine 2016 for additional discussion of satisfaction conditions of probes that lead to premature satisfaction. So-called *defective interveners* (Chomsky, 2000; see also McGinnis, 1998) may also be modeled in a similar way.

By separating the trigger of search termination (*satisfaction*) from the process of matching itself (*interaction*), Deal's interaction-satisfaction theory of probe specifications allows for the description of a wide range of attested probing interactions. This includes cases where the probe may interact with multiple goals before terminating (11a–b), as well as cases where search may terminate prematurely

---

[15] Deal proposes that this may be a general constraint on probing: "satisfaction features must be a subset of interaction features" (Deal, 2015: 3). Our discussion here of Ā-probing for the closest DP (Branan and Erlewine, to appear) and the examples of "horizons" in Keine 2016 serve as arguments against this particular detail in Deal's discussion of the interaction-satisfaction theory, which we otherwise adopt and advocate for.

[16] See also related discussion at the end of section 4.2.

[17] Precursors to this approach to such subject-only extraction asymmetries include Aldridge 2004, as we discuss in Branan and Erlewine to appear, as well as Erlewine 2018: 686–687 and Coon, Baier, and Levin 2021.

before interacting with a potential goal within the search space (11c).

(11) **Some probe specifications in interaction-satisfaction theory:**

a. Nez Perce C:
[INT:φ, SAT:ADDR]

b. Japanese T for nominative assignment:
[INT:uCase[18], SAT:−]

c. Ā-probing for the closest DP:
[INT:Ā+D, SAT:D]

d. Minimal search for [F]:
[INT:F, SAT:F] = [PROBE:F]

This empirical coverage provided by this interaction-satisfaction framework argues against the hypothesis that all forms of probing necessarily reflect "minimal search," where "minimal" means that the search will halt immediately upon encountering one matching goal. However, such a "minimal" probe may still be described in this theory, as a probe with identical interaction and satisfaction conditions (11d). We furthermore might consider specifications of this form to be some sort of default, explaining the prevalence of the widely attested Closest constraint on locality (2) which may be derived by probing involving the Chow 2022 search algorithm or breadth-first search (see section 2 above).[19] Finally, we note that Deal's theory is a *procedural* theory, or at least most naturally described in procedural terms. This demonstrates the strength of this procedural mode of description, even for describing interactions which do not obey the Closest constraint on locality.

4 Defining the search space

Our discussion thus far has concentrated on the mechanics of the search procedure associated with probing, given a particular search space, i.e. a syntactic structure which the search procedure traverses to find its goal(s). A procedural theory of probing must also describe the search space itself. If a particular bit of structure is not included in the search space, this guarantees that its contents will not be found by the search procedure.[20] This offers a means for describing *absolute locality* conditions on syntactic dependencies, where probes are simply unable to find potential goals in certain structural positions, in contrast to the relative locality effects described above, which are triggered by the presence of other potential goals in the structure.

The default assumption is for the search space for probe P to be the entire sister of P. In this section, we discuss two classes of potential revisions to this assumption, which lead to different sorts of absolute

[18] We use the notation [INT:uCase] here to specify targets with *unvalued* Case features. See also note 3 above.

[19] On the other hand, evidence from child language acquisition discussed in Friedmann, Belletti, and Rizzi 2009: 82–85 and Rizzi 2013: 180–182 suggests that probes which are unable to skip partially matching interveners and thus more susceptible to premature satisfaction, such as (11c), may in fact be the "default" in the course of acquisition.

[20] As noted in footnote 11 above, some works such as Preminger 2019 conceptualize search space restrictions — i.e. which nodes are accessible for probing — into the description of the search procedure itself. We believe it is productive to distinguish these two aspects in the description of probing procedures.

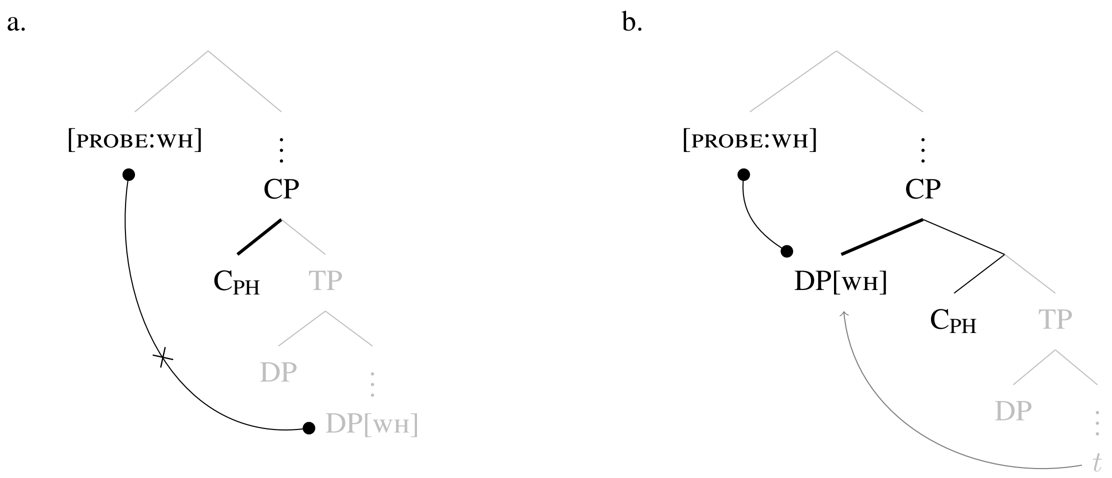
locality restrictions: first, the idea that some subparts of the search space may be made inaccessible for probing, and second, modifications to the shape and size of the search space.

4.1 Opacity in the search space

A recurrant idea in syntax is that operations can apply only within certain domains, ruling out true long-distance dependencies. Islands (Ross, 1967), the strict cycle condition (Chomsky, 1973), and barriers (Chomsky, 1986) all have described some such restrictions. In contemporary Minimalist work, the notion of *phase* has been developed (Chomsky, 2000, 2001). On this approach, certain heads are specified as *phase heads*, which delimit domains for probing.

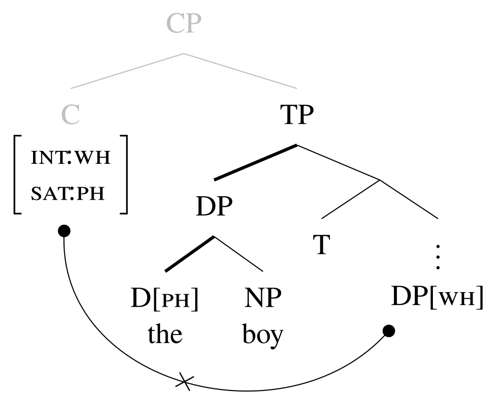
Consider the structures in (12), which reflect one portion of the derivation of a long-distance *wh*-question such as *Which book did Joey say that Meghan bought?*. Suppose that the embedded complementizer C is a *phase head* (PH), with the contents of its complement invisible for probing from above, indicated by gray coloring below. Search initiated by [PROBE:WH] in a higher phase cannot consider those gray nodes in the complement of C, and thus a *wh*-phrase within the lower phase will not be found in (12a), predicting the unavailability of direct, "one fell swoop" movement out of the CP. However, if the *wh*-phrase can first move to the "edge" of the phase as in (12b) — itself possibly triggered by an appropriate probe on the intermediate phase head itself — it escapes the domain which cannot be considered for probing from above, feeding long-distance movement in a *successive-cyclic* fashion (see e.g. Chomsky, 1977). In many works, phase head categories include C as well as *v* and D, but this is an area of active research; see e.g. Van Urk 2020. Phases which do not support such movement to the edge will be strictly opaque, offering a possible description for the strong islands of Ross 1967.

(12) **Phase impenetrability and intermediate movement:**



Following our earlier discussion of probe specification in terms of interaction and satisfaction conditions (Deal, 2015, to appear), we might be tempted to describe phase head categories as default satisfaction features of probes. For example, a *wh*-probe might have a specification such as [INT:WH, SAT:PH], with phase heads such as D, C, and *v* bearing the [PH] feature.[21] We note that this is not a satisfactory approach to implementing and parameterizing phasehood. A probe of this form indeed will be unable to probe into the complement of a [PH]-marked phase head, but we furthermore predict that the search will then terminate and not be able to consider any further structure elsewhere in the search space as well. For example, a left-to-right depth-first search initiated by this probe in order to initiate object *wh*-movement in a sentence such as *What did the boy read?* in (13) will be unable to probe into the sister of the subject DP, in addition to blocking search into the complement of the phase head D.[22] Phase impenetrability thus cannot be implemented in terms of the specification of the probe.

(13) **Specifying phase heads as satisfaction features (a bad idea):**



The inaccessibility of certain portions of the search space is therefore commonly modeled via modification of the search space's structure itself. Chomsky (2000, 2001) proposes that the opacity of phasal complements is the result of a process of *Spell-Out* which applies to the complements of phase heads at certain points in the derivation. Specifically, Chomsky (2001: pg. 5) proposes that Spell-Out will remove syntactic features from the portion of the clause targeted for Spell-Out, making it effectively invisible for subsequent probing. Nunes and Uriagereka 2000 develop a similar proposal, but where the result of Spell-Out is an atomic element, akin to a word, without its internal structure.[23] Regardless of the precise characterization of this mechanism, we can think of phases as categories that make their

[21] See Rackowski and Richards 2005 for a proposal along these lines, but also combined with a proposal where the phase "unlocks" and the search continues into the phase following Agree with the phase itself. The horizons of Keine 2016 are also specifications on probes that lead to premature satisfaction, but which are argued to be distinct from phases.

[22] The reader can verify that a breadth-first or Chow 2022 search procedure will not fare much better for the situation at hand.

[23] Nunes and Uriagereka's approach additionally forces moved phrases (and specifiers more generally) to undergo Spell-Out, removing them from the search space, bearing on the discussion of similarities and differences between different search algorithms in section 2 above.

complements invisible for all forms of probing from above — as reflected in (12) above, effectively removing the gray nodes from the search space — obviated only by prior movement of the potential goal to the phase edge.

An alternative to these approaches to absolute locality effects via modification of the search space structure is to view at least some absolute locality effects as the result of problems at the PF or LF interface. One promising strain of research along these lines involves the *Cyclic Linearization* framework developed in Fox and Pesetsky (2005) and developed in various ways in Bachrach and Katzir (2009); Ko (2014); O'Brien (2017); Davis (2020). There, one of the functions of Spell-Out is to fix the linear order of portions of the clause, but without rendering the material inaccessible for later probing. Probing into portions of the clause that have undergone Spell-Out is allowed, provided that subsequent operations triggered by the probe do not alter the relative linear order of those elements.

Another set of facts motivating such an approach involve cases where probing seems to be able to cross the same phase boundary in some contexts, but not others. These cases include the "unlocking" effects discussed in Rackowski and Richards (2005); Branan (2018); Halpert (2019); Preminger (2019), as well as the bound pronoun effect described in Grano and Lasnik (2018), Huang to appear and references there; in both of these cases a phase becomes transparent for probing if and when some portion of the phase is part of an independently well-formed non-local dependency with a higher element.[24] An example of this sort of effect is shown in (14) below. Here we see that extraction from a finite adjunct is allowed just in case the subject of the island is a pronoun bound by an argument in the matrix clause.

(14) **Bound pronominal subject effect:** (Grano and Lasnik, 2018: 494)

a. *What did Ann go home after Mary read?

b. ?What did Ann go home after she$_{\text{Ann}}$ read?

Such selective opacity effects pose a look-ahead problem for the common approach to phasal impenetrability as reflecting an irreversible process that removes portions of the search space. For example, on a theory where adjunct islands arise from phasal impenetrability, the derivation in (14) would have to "know" whether or not the subject of the adjunct will eventually be bound, in order to determine whether or not the adjunct should undergo Spell-Out. One move — that made by Grano and Lasnik (2018) — would be to complicate the definition of phase along those lines. The alternative is to treat the ungrammaticality of examples such as (14a) as due to a problem at the interfaces. For example, Truswell (2007) describes an LF condition which, roughly speaking, rules out cases of movement that crosses structures that describe multiple events. In this case, the establishment of a binding relationship into the

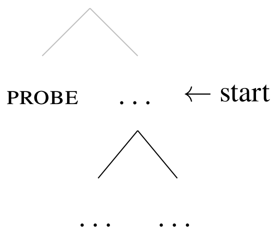
[24] Such effects possibly fall under the rubric of *Principle of Minimal Compliance* effects à la Richards 1998; see related discussion in Huang to appear.

adjunct would determine whether or not this condition is met. See also Keine 2016: §3.2.1 for further discussion of selective opacity phenomena and their challenges to phase theory.

4.2 Revising the shape of the search space

In all of the cases we have considered thus far, we have assumed the search space to be the sister of the probe. This possibility is illustrated in (15) below with the start node of the search labeled, and with higher structure that is not subject to search in gray. Two motivations are frequently given for this view. First, this choice ensures that the probe will c-command any goals (see e.g. Epstein, Groat, Kawashima, and Kitahara, 1998; Epstein, 1999), which is often assumed to be part of the desideratum for probe-goal relations. Second, this allows probing to be described as taking place as soon as possible, i.e. right after the probe is Merged into the structure, necessitated perhaps by a principle such as the Earliness Principle as stated by Pesetsky and Torrego (2001: 400).

(15) **Searching from the probe's sister down (most common assumption):**

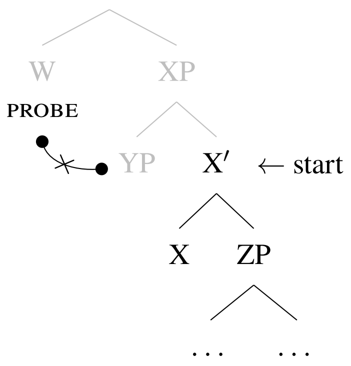


In this section, we discuss and consider various modifications to this assumption that the search space is the sister of the probe, together with its potential motivations above. We first briefly consider the possibility of the start node being lower than (dominated by) the probe's sister, and then consider start nodes higher than (dominating) the probe itself.

Let us consider the possibility of a probe's search space being properly contained within the probe's sister, i.e. with a start node lower than the default position, and what probing behavior we predict. Suppose for concreteness that the start node for probing takes the sister's lowest non-minimal projection, e.g. its bar-level projection, without its specifiers. This possibility is illustrated in (16), where XP is the probe's sister and the start node is labeled X′. A search of this form will effectively skip any potential goals that are specifiers of XP, such as YP, or contained within, as they are not dominated by the search's start node. If the probe seeks a goal to move to its specifier, this search space specification predicts an effect similar to that of *Spec-to-Spec Anti-Locality* (Bošković, 2016; Erlewine, 2016, 2020; Deal, 2019; Branan, to appear), a proposed constraint that bans movement of a specifier (e.g., Spec,XP) to the specifier position of the next projection up (Spec,WP).[25]

[25] The predicted effect is different, however, in that Spec-to-Spec Anti-Locality as described in these works does not ban

(16) **Searching from the sister's lowest non-minimal projection:**



More generally, for probes whose search space is specified to be a proper subpart of its sister, there will be a "gap" in the tree between the probe and the node from which search takes place. Nodes contained within this gap cannot be found by minimal search, and thus cannot be Agreed with or moved. Specifying the search space in this way makes a portion of the probe's sister "off limits," similar to the effect of structure removal as discussed in the previous section, resulting in a different kind of absolute locality effect, i.e. an *anti-locality* effect. Depending on how exactly the start node is specified, this approach may be able to derive other anti-locality effects such as those discussed in Ishii 1997; Bošković 1997; Saito and Murasugi 1999; Abels 2003, and Grohmann 2003, also reviewed in Grohmann 2011.[26]

Next we consider the possibility of searches that start at a node higher than (i.e. dominating) the probe itself. Specifically, in the remainder of this section, we explore the consequences of taking the start node of searches to always be the root of the tree; i.e. for the search space to always be the entire tree (modulo inaccessible subparts, discussed in the preceding section).[27] If combined with the assumption that probing takes place as soon as the probe is Merged into a structure (e.g. Pesetsky and Torrego's Earliness Principle), this yields an expectation that probing always takes the probe's mother as its start node, as in (17a) below. (We discuss (17b) in a moment.) The predictions of this formulation are however not drastically different from the sister-start formulation in (15); if we furthermore assume that a probe does not match itself nor its mother — at least in cases where the probe is a head which projects its features to its mother — we still maintain the expectation, above, that the probe should c-command its goal(s).

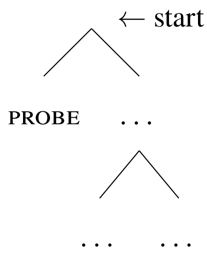
subextraction out of the specifier (movement of material within YP) to Spec,WP, whereas the restricted search space proposal in (16) additionally predicts an inability to interact with subparts of the specifier (YP) as well.

[26] But see also e.g. Zyman 2021 for alternative procedural explanations for a range of observed anti-locality constraints — including Spec-to-Spec, mentioned above, as well as Abels' Comp-to-Spec Anti-Locality and Zyman's own Phasal Anti-Locality — which do not involve shifting the start node of the search space.

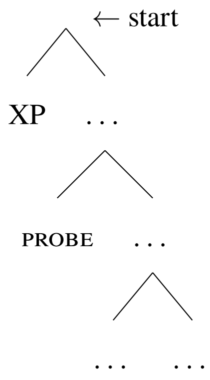
[27] This view may in fact be quite natural within a theory where syntactic operations apply to and modify the workspace, such as Chomsky 2019. In such a theory, we could describe search as always considering a member of the workspace (specifically, the complex object that contains the probe), rather than a subpart thereof. Note however that the possibility that we explore here is *not* that search considers the entire workspace containing the probe.

(17) **Searching from the root of the tree:**

a. Immediately after Merging the probe:



b. Following Merge of a specifier XP:



Now suppose we relax the expectation that probing takes place immediately after the probe is Merged in, allowing for further structure building to take place before we initiate a search from the root node. For example, this possibility is illustrated in (17b) for the case of search initiated after a phrase XP is Merged in as a specifier of the phrase whose head hosts the probe. By waiting to initiate the search associated with a probe, we effectively expand the search space of the probe, upwards.

Béjar and Rezac 2009 develops a proposal for various person hierarchy effects in φ-agreement which involves a process of search space expansion akin to that in (17b). More specifically, they discuss a complex φ-probe on v which probes once after Merging with its sister (as in (17a)) and then probes again after Merger of its specifier (as in (17b)), if a fully satisfactory goal was not identified earlier.[28] Note that the specifier XP and its contents (if accessible) are the only potential goals made available in this second cycle of probing (17b) which were not available in the first cycle (17a). Richards 2004 argues for a similar configuration of probing for *wh*-movement in Bulgarian.

For Béjar and Rezac 2009 as well as Carstens 2016, a probe's search can be delayed or retried in cases where an earlier search failed to yield a fully satisfactory goal. But suppose that the search — still necessarily starting at the tree's root node — can be further delayed. Facts from subject-oriented complementizer agreement — such as those described for Lubukusu in Diercks 2010 — could be captured straightforwardly under such a model. In Lubukusu, certain complementizers which head finite complement clauses agree in noun class with the subject of the next highest finite clause, as in (18). The agreement pattern truly seems to be subject-oriented, rather than controlled by the (intuitively) closest nominal, as internal arguments of the higher verb cannot control complementizer agreement:

[28] For Béjar and Rezac 2009, this derivation involves projection of the probe features to each successive projection of the head, with each step of probing then described as searching within the probe's daughters (described as its "sister" in bare phrase structure terms, in p. 48 fn. 7). See also Carstens 2016 (especially §1.4 and citations there), as well as Branan 2019 and Clem 2019 for further discussion of probes on projections — rather than heads — initiating searches.

(18) **Subject-oriented complementizer agreement in Lubukusu:**[29] (Diercks, 2010: 298)

Ewe	w-a-bol-el-a	Nelsoni	[$_{CP}$ {**o**-li	/ ***a**-li}	ba-keni	ba-rekukha	].
2sg	2sg-PST-say-AP-FV	Nelson(3sg)	2sg-that	*3sg-that	3pl-guests	3pl-left	

'You told Nelson that the guests left.'

If search always proceeds from the root of the tree, and may be delayed, then it may be that these factors conspire to allow the embedded complementizer to agree with an *unintuitively* closest nominal. In particular we could imagine that the pattern described arises as a result of search for class features by Lubukusu C being delayed until a significant portion of the higher clause has been constructed — in particular, until whatever functional scaffolding has been added to the tree that maps to the relevant notion of "subjecthood" that is apparently relevant for determining what may act as a controller of agreement. At that point in the derivation, search initiated by the embedded complementizer will take place from the root of the tree, and find the closest nominal to the root, which will then consistently be the higher subject. See Ke 2019: 80ff for a proposal along these lines.[30]

Finally, we note that in Branan and Erlewine to appear, we show that a phrase and its specifier may be equidistant to higher probes. Evidence for this form of equidistance comes from the behavior of Ā-probing for the closest DP, discussed in section 3: Ā-extraction with such probes is restricted to the highest nominal *or its specifier*. This form of equidistance is predicted by a declarative formulation of Closest as in (2), but is at odds with the procedural approaches discussed here: in particular, for all of the search algorithms that we have considered here, a phrase should always be found before its specifier. One way to describe such probing in procedural terms would be to allow the result of one search operation to serve as the starting node for a subsequent search operation under certain conditions. In particular, we propose that a partial match for a joint probe could then lead to a secondary search using the partial matching goal as its search domain. In this case, a partially matching DP that lacks a [Ā] feature would serve as the domain for a secondary [Ā+D] search. The possibility of such search procedures also provide a potential explanation for *unlocking* effects of the sort discussed in Rackowski and Richards 2005, Branan 2018, Halpert 2019 a.o., where extraction from certain constituents is contingent on Agree with said constituent.

5 Conclusion

In this chapter, we attempted to explicate the *procedural* approach to locality effects in Minimalist syntax. This procedural approach aims to characterize the workings of a foundational syntactic operation

[29] Following Bantuist convention, Diercks glosses the third-singular Nelson as noun class 1 and third-plural 'guests' as noun class 2, but here we have reglossed these values with their corresponding ϕ-features. AP = applicative, FV = final vowel.

[30] Furthermore, if delayed search can be retried over and over as we expand the search space step by step, generalizing the process in (17), until a suitable goal is found, we effectively derive a probing process similar to that of "upwards" Agree as in Bjorkman and Zeijlstra 2019 and citations there, using a downward search algorithm (section 2).

— probing — which underlies processes such as agreement, case-assignment, and movement, and to *derive* their observed locality restrictions as predictions of the probing procedure. This procedural approach contrasts with the *declarative* approach to locality effects — common in earlier work on syntactic locality but also still common in contemporary work — which puts forward constraints on licit and illicit dependency configurations, such as the widely adopted Closest c-command condition (2), with deeper motivations for the observed constraints only offered in some cases.

We began by discussing the precise search algorithm underlying probing in section 2, observing that the Closest c-command condition is naturally derived by a breadth-first search procedure which is "minimal," i.e. immediately terminates after one suitable goal is found (Ke, 2019). We however also noted there that, in various configurations, depth-first search may in fact make similar — or in some cases superior — predictions for relative locality effects as compared to breadth-first search. We then highlighted the novel search algorithm in Chow 2022, which combines the positive qualities of both breadth-first and depth-first algorithms, as a potential contender for adoption. In section 3 we discuss phenomena which empirically challenge the "minimal" property of probing as minimal search, motivating the *interaction-satisfaction* revision to the procedural theory of locality (Deal, 2015, to appear). We showed that this framework successfully allows for the description of interactions where a probe may probe with multiple goals, as well as those where a probe may prematurely terminate before finding a goal or exhausting its search space. From this discussion of the search procedure which allows for the derivation of various *relative locality* effects and their variation, in section 4 we move to different potential modifications to the shape and size of the search space, which allows for the derivation and explanation of various *absolute locality* effects.

The conceptual shift away from declarative constraints and towards procedural explanations in Minimalist theorizing reflects the pursuit of the hypothesis that many aspects of grammatical behavior "fall out in some natural way from the computational process" (Chomsky, 2000: 113), governed by "third factor" considerations of "efficient computation" (Chomsky, 2005: 6). In this procedural mode of explanation, it may therefore be tempting to to consider and appeal to the relative computational "costs" of particular proposals. For example, we might consider the time and space (memory) utilization of different search algorithms (see e.g. Korf, 1985), with an expectation that the language faculty must choose the more efficient option, or suggest that the reduction of search spaces e.g. via Spell-Out (section 4.1) has an efficiency motivation.

In this chapter, we have not emphasized such possible motivations for the nature of locality effects from considerations of computational efficiency, for two reasons. First, in the interest of space, we have concentrated on considering and illustrating how the probe-goal model allows us to adequately account for, and make sense of, attested patterns of locality effects.[31] Second, there is reason to believe that search procedures of the form described here are not reflected in real-time processing. As summarized in e.g. Kush 2013: ch. 2, online sentence processing measures suggest that dependency formation such as for agreement and movement take constant time, i.e. not taking proportionately more time with longer probe-goal paths or larger search spaces. This suggests that these mechanisms in online processing do *not* carry out a node-by-node search procedure of the type described here.[32] Given this disconnect between (our current understandings of) the procedural theory of locality and online processing behavior, appeals to computational efficiency in motivating particular analytical choices are at best premature, even when we adopt this (in our opinion successful) procedural approach to locality effects.

[31] Notably, doing so led us to abandon the strong view that probing is always minimal search (see section 3), despite the idea and claim that probing being minimal search is one of the "natural conditions of efficient computation" (Hauser, Chomsky, and Fitch, 2002: 1578) and the "principles so elementary that they would be incorporated in any serious analysis" (Chomsky, 2014: 97). In this sense, the empirical facts seem to force us into a decidedly unserious position.

[32] We do not however take this tension to suggest that the procedural theory described here is bankrupt. Within Marr's three-level model for the analysis of cognitive systems (Marr, 1982), Johnson (2016) notes that "linguistic theories are computational-level theories of language, while psycholinguistic theories of comprehension or production are algorithmic-level descriptions of how knowledge of language can be put to use" (p. 172) and furthermore emphasizes that complexity at one level of description does not necessarily correlate with complexity at another level.

References

Abels, Klaus. 2003. Successive cyclicity, anti-locality, and adposition stranding. Doctoral Dissertation, University of Connecticut.

Aldridge, Edith. 2004. Ergativity and word order in Austronesian languages. Doctoral Dissertation, Cornell University.

Atlamaz, Ümit. 2019. Agreement, case, and nominal licensing. Doctoral Dissertation, Rutgers.

Bachrach, Asaf, and Roni Katzir. 2009. Right-node raising and delayed spellout. In *Interphases: Phase-theoretic investigations of linguistic interfaces*, 283–316. Oxford University Press.

Belletti, Adriana, ed. 2004. *Structures and beyond: The cartography of syntactic structures, volume 3*. Oxford University Press.

Bhatt, Rajesh. 2005. Long distance agreement in Hindi-Urdu. *Natural Language & Linguistic Theory* 23:757–807.

Bjorkman, Bronwyn M., and Hedde Zeijlstra. 2019. Checking up on (φ-)Agree. *Linguistic Inquiry* 50:527–569.

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

Bošković, Željko. 1997. *The syntax of nonfinite complementation: An economy approach*. MIT Press.

Bošković, Željko. 2003. Agree, phases, and intervention effects. *Linguistic Analysis* 33:54–96.

Bošković, Željko. 2007. On the locality and motivation of Move and Agree: An even more minimal theory. *Linguistic Inquiry* 38:589–644.

Bošković, Željko. 2016. On the timing of labeling: Deducing comp-trace effects, the subject condition, the adjunct condition, and tucking in from labeling. *The Linguistic Review* 33:17–66.

Branan, Kenyon. 2018. Attraction at a distance: Ā-movement and case. *Linguistic Inquiry* 49:409–440.

Branan, Kenyon. 2019. Nominals Agree with each other in Coahuilteco. Manuscript, National University of Singapore.

Branan, Kenyon. to appear. Locality and anti-locality: The logic of conflicting requirements. *Linguistic Inquiry* .

Branan, Kenyon, and Michael Yoshitaka Erlewine. to appear. Ā-probing for the closest DP. *Linguistic Inquiry* .

Branan, Kenyon, and Abdul-Razak Sulemana. 2019. In Buli, covert movement licenses parasitic gaps. In *Proceedings of WCCFL 36*, ed. Richard Stockwell, Maura O'Leary, Zhongshi Xu, and Z.L. Zhou, 81–90.

Bruening, Benjamin. 2014. Precede-and-command revisited. *Language* 90:342–388.

Béjar, Susana, and Milan Rezac. 2009. Cyclic agree. *Linguistic Inquiry* 40:35–73.

Cao, Yu. 2017. Labeling for linearization. Manuscript, Rutgers University.

Carstens, Vicki. 2016. Delayed valuation: a reanalysis of goal features, "upward" complementizer agreement, and the mechanics of Case. *Syntax* 19:1–42.

Chomsky, Noam. 1973. Conditions on transformations. In *A festschrift for Morris Halle*, ed. Stephen R. Anderson and Paul Kiparsky. New York: Holt, Reinhart, and Winston.

Chomsky, Noam. 1977. On *wh*-movement. In *Formal syntax*, ed. Peter Culicover, Thomas Wasow, and Adrian Akmajian, 71–132. New York: Academic Press.

Chomsky, Noam. 1986. *Barriers*. MIT Press.

Chomsky, Noam. 1995. *The minimalist program*. MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: the framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89–156. MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In Kenstowicz (2001), 1–52.

Chomsky, Noam. 2004. Beyond explanatory adequacy. In Belletti (2004).

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

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130:33–49.

Chomsky, Noam. 2014. Some core contested concepts. *Journal of Psycholinguistic Research* 44:91–104.

Chomsky, Noam. 2019. The UCLA lectures. URL https://ling.auf.net/lingbuzz/005485, lecture notes by Robert Freidin.

Chow, Keng Ji. 2022. A novel algorithm for minimal search. *Snippets* 42:3–5.

Clem, Emily Catherine. 2019. Agreement, case, and switch-reference in Amahuaca. Doctoral Dissertation, University of California Berkeley.

Collins, Chris. 2005. A smuggling approach to the passive in English. *Syntax* 8:81–120.

Coon, Jessica, Nico Baier, and Theodore Levin. 2021. Mayan Agent Focus and the ergative extraction constraint: Facts and fictions revisited. *Language* 97:269–332.

Coon, Jessica, and Alan Bale. 2014. The interaction of person and number in Mi'gmaq. *Nordlyd* 40:85–101.

Coon, Jessica, and Stefan Keine. 2021. Feature gluttony. *Linguistic Inquiry* 52:655–710.

Davis, Colin. 2020. The linear limitations of syntactic derivations. Doctoral Dissertation, MIT.

Deal, Amy Rose. 2015. Interaction and satisfaction in φ-agreement. In *Proceedings of NELS 45*, 1, 179–192.

Deal, Amy Rose. 2019. Raising to ergative: Remarks on applicatives of unaccusatives. *Linguistic Inquiry* 50:388–415.

Deal, Amy Rose. to appear. Interaction, satisfaction, and the PCC. *Linguistic Inquiry* Manuscript, University of California at Berkeley.

Diercks, Michael James Kroening. 2010. Agreement with subjects in Lubukusu. Doctoral Dissertation, Georgetown University.

Epstein, Samuel David. 1999. Un-principled syntax: The derivation of syntactic relations. In *Working minimalism*, ed. Samuel David Epstein and Norbert Hornstein, 317–345. MIT Press.

Epstein, Samuel David, Erich M. Groat, Ruriko Kawashima, and Hisatsugu Kitahara. 1998. *A derivational approach to syntactic relations*. Oxford University Press.

Epstein, Samuel David, and T. Daniel Seely, ed. 2002. *Derivation and explanation in the minimalist program*. Blackwell.

Erlewine, Michael Yoshitaka. 2016. Anti-locality and optimality in Kaqchikel Agent Focus. *Natural Language & Linguistic Theory* 34:429–479.

Erlewine, Michael Yoshitaka. 2018. Extraction and licensing in Toba Batak. *Language* 94:662–697.

Erlewine, Michael Yoshitaka. 2020. Anti-locality and subject extraction. *Glossa* 5:1–38.

Fox, Danny, and David Pesetsky. 2005. Cyclic linearization of syntactic structure. *Theoretical Linguistics* 31:1–45.

Friedmann, Naama, Adriana Belletti, and Luigi Rizzi. 2009. Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua* 119:67–88.

Grano, Thomas, and Howard Lasnik. 2018. How to neutralize a finite clause boundary: Phase theory and the grammar of bound pronouns. *Linguistic Inquiry* 49:465–499.

Grohmann, Kleanthes K. 2003. *Prolific domains: On the anti-locality of movement dependencies*. John Benjamins.

Grohmann, Kleanthes K. 2011. Anti-locality: Too close relations in grammar. In *Oxford handbook of linguistic minimalism*, ed. Cedric Boeckx, 260–290. Oxford University Press.

Halpert, Claire. 2019. Raising, unphased. *Natural Language & Linguistic Theory* 37:123–165.

Hamann, Jakob. 2014. Argument encoding in direction systems and specificity-driven Agree. In *Scales and hierarchies*, ed. Ina Bornkessel-Schlesewsky, Andrej L. Malchukov, and Marc D. Richards. De Gruyter Mouton.

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

Hiraiwa, Ken. 2001. Multiple Agree and the defective intervention constraint in Japanese. In *Proceedings of HUMIT 2000*, 67–80. MIT Working Papers in Linguistics.

Huang, Cheng-Teh James. 1982. Logical relations in Chinese and the theory of grammar. Doctoral Dissertation, Massachusetts Institute of Technology.

Huang, Nick. to appear. How subjects and possessors can obviate phasehood. *Linguistic Inquiry* .

Ishii, Toru. 1997. An asymmetry in the composition of phrase structure and its consequences. phdthesis, University of California at Irvine.

Johnson, Mark. 2016. Marr's levels and the minimalist program. *Psychonomic Bulletin & Review* 24:171–174.

Kayne, Richard. 1994. *The antisymmetry of syntax*. MIT press.

Ke, Hezao. 2019. The syntax, semantics, and processing of agreement and binding grammatical illusions. Doctoral Dissertation, University of Michigan.

Keine, Stefan. 2016. Probes and their horizons. Doctoral Dissertation, University of Massachusetts Amherst.

Kenstowicz, Michael, ed. 2001. *Ken Hale: A life in language*. MIT Press.

Ko, Heejeong. 2014. *Edges in syntax: Scrambling and cyclic linearization*. Oxford University Press.

Korf, Richard E. 1985. Depth-first iterative-deepening: An optimal admissible tree search. *Artificial Intelligence* 27:97–109.

Kremers, Joost. 2003. The Arabic noun phrase: A minimalist approach. Doctoral Dissertation, University of Nijmegen.

Kremers, Joost. 2009. Recursive linearization. *The Linguistic Review* 26:135–166.

Krivochen, Diego Gabriel. 2021. The search for minimal search. Manuscript.

Kuno, Susumu, and Jane J. Robinson. 1972. Multiple wh questions. *Linguistic Inquiry* 3:463–487.

Kural, Murat. 2005. Tree traversal and word order. *Linguistic Inquiry* 36:367–387.

Kush, Dave W. 2013. Respecting relations: Memory access and antecedent retrieval in incremental sentence processing. Doctoral Dissertation, University of Maryland.

Lahne, Antje. 2012. Specificity-driven syntactic derivation. In *Ways of structure building*, ed. Myriam Uribe-Etxebarria and Vidal Valmala, 271–296. Oxford University Press.

Lasnik, Howard, and Mamoru Saito. 1992. *Move alpha: Conditions on its applications and output*. MIT Press.

Marr, David. 1982. *Vision*. MIT Press.

McGinnis, Martha. 1998. Locality in A-movement. Doctoral Dissertation, MIT Working Papers in Linguistics.

Nakamura, Masanori. 1998. Reference set, minimal link condition, and parameterization. In *Is the best good enough?*, ed. Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky. MIT Press and MIT Working Papers in Linguistics.

Nunes, Jairo, and Juan Uriagereka. 2000. Cyclicity and extraction domains. *Syntax* 3:20–43.

O'Brien, Chris Harris. 2017. Multiple dominance and interface operations. Doctoral Dissertation, Massachusetts Institute of Technology.

Pesetsky, David, and Esther Torrego. 2001. T-to-C movement: Causes and consequences. In Kenstowicz (2001), 355–425.

Pesetsky, David, and Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In *Clausal and phrasal architecture: Syntactic derivation and interpretation*, ed. Simin Karimi, Vida Samiian, and Wendy K. Wilkins, 262–294. John Benjamins.

Preminger, Omer. 2014. *Agreement and its failures*. MIT Press.

Preminger, Omer. 2019. What the PCC tells us about 'abstract' agreement, head movement, and locality. *Glossa* 4:1–42.

Rackowski, Andrea, and Norvin Richards. 2005. Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry* 36:565–599.

Richards, Marc. 2012. Probing the past: On reconciling long-distance agreement with the PIC. In *Local modelling of non-local dependencies in syntax*, ed. Artemis Alexiadou, Tibor Kiss, and Gereon Müller, 135–154. de Gruyter.

Richards, Norvin. 1998. The principle of minimal compliance. *Linguistic Inquiry* 29:599–629.

Richards, Norvin. 2004. Against bans on lowering. *Linguistic Inquiry* 35:453–463.

Rizzi, Luigi. 1990. *Relativized minimality*. The MIT Press.

Rizzi, Luigi. 2004. Locality and the left periphery. In Belletti (2004), 223–251.

Rizzi, Luigi. 2013. Locality. *Lingua* 130:169–186.

Ross, John Robert. 1967. Constraints on variables in syntax. Doctoral Dissertation, Massachusetts Institute of Technology.

Saito, Mamoru, and Keiko Murasugi. 1999. Subject predication within IP and DP. In *Beyond principles and parameters*, ed. Kyle Johnson and Ian Roberts, 167–188. Berlin: Kluwer Academic Publishers.

Truswell, Robert. 2007. Locality of wh-movement and the individuation of events. Doctoral Dissertation, University College London.

Van Urk, Coppe. 2020. Successive cyclicity and the syntax of long-distance dependencies. *Annual Review of Linguistics* 6:111–130.

Van Urk, Coppe, and Norvin Richards. 2015. Two components of long-distance extraction: Successive cyclicity in Dinka. *Linguistic Inquiry* 46:113–155.

de Vries, Mark. 2009. On multidominance and linearization. *Biolinguistics* 3:344–403.

Yasui, Miyoko. 2003. A graph-theoretic reanalysis of bare phrase structure theory and its implications on parametric variation. In *Proceedings of linguistics and phonetics 2002*. Meikai University.

Zyman, Erik. 2021. Antilocality at the phase edge. *Syntax* 24:510–556.