COPiL CAMBRIDGE OCCASIONAL PAPERS *in* LINGUISTICS
Volume **12**, Article 1: pp. 1–30, 2020 | ISSN: 2050-5949

A Third-Factor Account of Locality: Explaining impenetrability and intervention effects with minimal search*

Seth Aycock
University of Cambridge

Abstract Current approaches to locality focus on two central ideas, intervention and impenetrability, attempting to explain the former with Relativized Minimality (Rizzi 1990, 2001, 2011, Starke 2001) and the latter with the Phase Impenetrability Condition from Chomsky's Phase Theory (Chomsky 2000, 2001, 2008). Some redundancy exists between the two constraints. In this paper, I propose an account deriving both intervention and impenetrability effects from minimal search. Minimal search is an optimal reflex of the third-factor notion of minimal computation. This paper details an account which derives an intervention principle of Probe Closest Goal, the Phase Impenetrability Condition and Antilocality from narrow syntax while removing several stipulated principles from the grammar. The approach also provides accounts for empirical data in a parsimonious manner. This paper thus presents a theoretically and empirically motivated third-factor account of locality, offering a useful foundation for future work.

1 Introduction

Locality, i.e. constraints on how local a syntactic relation can be, or the domain over which structure building functions operate (Stabler 2011), is central to syntax. Intervention and impenetrability are the two primary types of locality constraints discussed in the literature. Relativized Minimality (RM) (Rizzi 1990) purports to explain intervention effects, while the Phase Impenetrability Condition (PIC) in Chomsky's Phases framework (Chomsky 2000, 2001, 2008) is claimed to explain impenetrability effects, particularly Ross' (1967) island constraints. There is some conceptual and empirical redundancy between the two effects, as both reduce the search space of probes and enforce successive cyclic movement. Under the Minimalist Program (MP), redundancy should be reduced, either by reducing one type of locality to the other via a more powerful principle (cf. Abels 2003, Chomsky 2000, Rizzi 2009, Starke 2001, Torr 2012); or by deriving both locality effects from one source, as is attempted in this paper.

* I would like to thank Ian Roberts for his wise and invaluable supervision; Theresa Biberauer for her enthusiastic and meticulous assistance; Željko Bošković for his insightful technical advice; and attendees of Cambridge SyntaxLab for the constructive discussion and feedback. Any errors are, of course, the author's own.

This proposal attempts to unify impenetrability and intervention constraints under Minimal Search (MS) in a principled explanation that is: reducible to legibility conditions of the sensorimotor (SM) and conceptual-intentional (C-I) interfaces; an optimal solution i.e. maximally computationally efficient (Chomsky 2005b: 10); and the reflex of a third factor notion (as per Chomsky's (2005b: 6) hypothesis that I-language is shaped by what is genetically endowed (UG), the primary linguistic data, and third factor (F3) domain-general cognitive principles of data analysis and efficiency). MS relies on insights from computer science, being a maximally efficient and optimal tree search method. Let us briefly review the method by which MS can derive RM and the PIC, the details of which are discussed in due course.

Search is widely held to be a prerequisite of Merge (Rizzi 2013: 179); MS is a constrained version of Search, constrained by efficiency and optimality rather than stipulation. MS limits what Merge can 'see', for instance deriving the PIC as a lower bound past which Search cannot occur. Here, the PIC is a reflex of MS, and island constraints are the outcome of the PIC (combined with Chomsky's (2013) Labelling Algorithm (LA)). Intervention effects are due to MS constraining probe-goal relations in Agree; RM is recast as Probe Closest Goal, casting a widely held insight as F3 driven. The deep consequences of this proposal include a parsimonious view of the derivation as three operations (Merge, Labelling and Agree) operating cyclically, all involving the operation of MS. In this way, MS explains and unifies both impenetrability and intervention effects.

Below, I explore the details of the current proposal of MS-locality. In section 2, I consider theoretical motivations and a definition of MS. Section 3 is the implementation of MS deriving intervention effects, and its empirical testing. Section 4 derives the PIC from MS, explaining island effects with MS-locality. Section 5 offers a conclusion.

2 Minimal Search

2.1 Theoretical Motivation and Third Factors

In the MP, the derivation's syntactic component involves three operations: Merge, Agree and Labelling (Chomsky 1995, 2000, 2008, 2013). Lexical items are feature bundles, which are selected from the lexicon then built into sets by Merge; Agree and Labelling may further modify the output of Merge.[1] The syntactic output is then transferred to (or spelled out at) the SM and C-I interfaces.

The current proposal extends ideas raised by Chomsky (2013, 2015a), proposing that the Labelling Algorithm (LA) is "just minimal search, presumably appropriating a third factor principle, as in Agree and other operations" (2013: 43); in a later discussion, Chomsky (2015a) says LA is a special case of MS, like Agree (as they both involve Probe-Goal MS), which in turn falls under Minimal Computation (MC). Minimal Search is mentioned primarily by Chomsky (2013, 2015a,b), but is raised by others: Roberts (2019: 605) notes locality could be reduced to MS; Rizzi (2013: 179)

[1] Ke (2019: 17) notes there is no consensus on where Agree and Labelling occur in this model; I leave this debate aside here.

refers to MS as a restriction on Search for Merge; Larson (2015) characterises MS as a restriction on Merge. Ke (2019) provided the first formal definition of MS, in an attempt to unify Agree and LA. Ke's hypothesis is that an optimal system would have Agree and LA derived from an independently motivated F3, i.e. MC, unifying LA and Agree via MS and, as F3s are freely available, meaning LA and Agree would 'come for free' (Chomsky 2005b). Ke concludes the unification is only partial; some aspects (e.g. valuation) are first-factor and not reducible to MS. However, MS is still key to both operations. Here, I propose Merge, like Agree and LA, requires MS. The natural consequence is the reduction of locality to the derivation. Locality is a reflex of MS constraining Merge and Agree. Therefore, locality constraints can be unified while 'coming for free' in the derivation.

Approaches to both intervention and impenetrability have appealed to processing and efficiency. Phase Theory's (PT) alleged reduced memory load is easily criticised (Boeckx & Grohmann 2007), given the derivation's status as an abstract bottom-up proof, not a processing model relating to performance (see section 4.1.2 for discussion). RM is argued to have a natural evolutionary link with processing (Ortega-Santos 2011), but this is likely too simple, given our lack of knowledge about language evolution (Chomsky 2011, Berwick & Chomsky 2016). Any connection between locality and language use, i.e. processing or production, is likely highly indirect. Further, RM's status as a representational principle is awkward in the MP and is preferably explained derivationally[2] while maintaining its true empirical generalisations and coverage, as per the current F3-based MS-locality. A processing connection is thus equally likely for impenetrability and intervention. Maintaining the derivation as a model of competence means the only valid appeal to efficiency and optimality is through F3s. This proposal makes such an appeal in deriving intervention and impenetrability effects, without stipulation, from the F3 notion MS. The appeal to F3s is MS-locality's main motivation. In sum, locality constraints come for free and are a natural reflex of F3-driven optimisation, simplifying the notion of locality.

2.2 Formalising Minimal Search

2.2.1 Definition

For the formal definition of MS, I follow (Ke 2019: 44). MS is minimal in that search is terminated as soon as the first fitting target is returned.

(1) MS = ⟨ SA, SD, ST ⟩

[2] The derivational versus representational question is complex; e.g. multiple spell-out fits a derivational approach better but allows for a combined approach (Bošković p.c.); Chomsky's (2001) proposal evaluates locality phase by phase, which is neither purely derivational or representational. Brody (2002) argues no pure derivational theory of narrow syntax exists, but that mixed theories introduce redundancy and are less restrictive than pure varieties. The question remains immune to empirical evidence. It seems unimportant whether syntax is construed representationally or derivationally, thus it becomes a purely theoretical decision. Under the MP's goal to remove conditions on representations, I recast RM derivationally, allowing MS-locality to unify intervention and impenetrability effects.

Where MS is Minimal Search, SA is the search algorithm, SD is the search domain in which SA operates, and ST is the search target (i.e. the features SA is searching for).

(2) SA:

a. Given ST and SD, match against every head member of SD to find ST.

b. If ST is found, return the heads bearing ST and go to c. Otherwise, gget the set members of SD and store them as list L:

   i. If L is empty, search fails and go to c. Otherwise,

   ii. Assign each set in L as a new SD and go to a. for all these new SDs.

c. Terminate Search.

2.2.2 Example

Given the above definition of MS, let us consider an example of its operation. In Figure 1, MS is initiated for ST = [F], in SD = set α. Search runs iteratively three times to find the target. First, MS searches inside α, finding the head A, not bearing [F], and the set β. β is stored in list L_1 and assigned as the new SD, from which the second iteration of MS occurs. This run finds sets γ and δ, neither bearing [F], and both of which are stored in list L_2, then assigned as new SDs. The third iteration initiates two parallel minimal searches. In γ, no heads with the ST are found, and search terminates. In δ, set ϵ is added to list L_3, and the head $D_{[F]}$ is found bearing the ST [F], terminating search. MS returns $D_{[F]}$, which enters some syntactic relation (via Merge, Agree or LA).

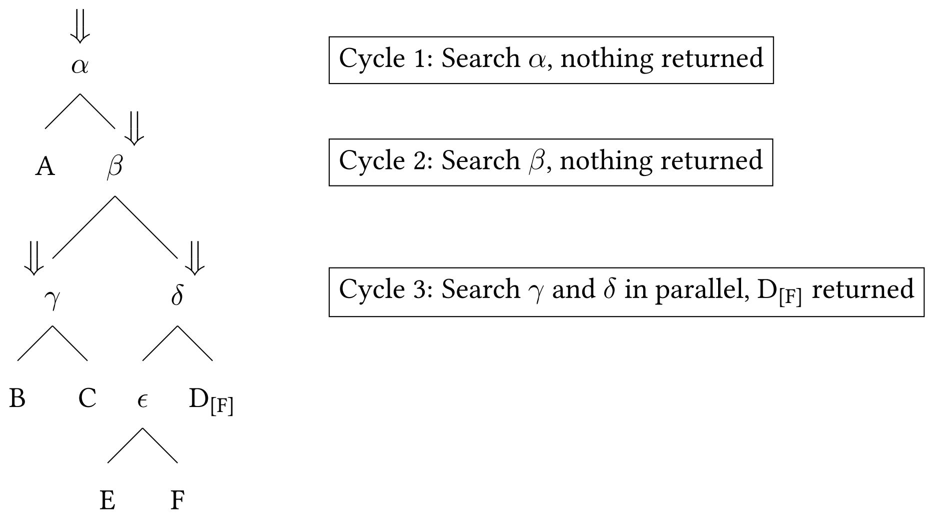

Figure 1

As Ke notes, the "minimality in the breadth-first search is… captured by storing sets as a list L", not by counting the levels of sets that MS looks into (2019: 47). All sets at the same level are stored in the same list, and new minimal searches are initiated to search these sets in parallel.

2.2.3 Points of Contention

There are two points of contention in Ke's algorithm: first, whether MS is breadth- or depth-first; and whether MS occurs in parallel.

In computer science, there are two major search algorithms, breadth-first search (BFS) and depth-first search (DFS). BFS algorithms search the root node, then the root node's successors, then their successors, and so on. All nodes are searched at a given depth in the search tree before any nodes at the next level are searched (Russell & Norvig 2010: 81). BFS is complete, in that in a binary tree, it will find a solution if it exists. BFS is not optimal, meaning that the number of steps taken in reaching the solution is high, and it suffers from a high memory demand (as each searched node must be kept in memory to return to), mirrored by Ke's List in which sets are stored then searched. Finally, BFS has a relatively high execution time (i.e. it searches a large number of nodes to reach the solution). In Figure 1, the search order for BFS is: α, A, β, γ, δ, B, C, ϵ, $D_{[F]}$, E, F.

DFS on the other hand exhaustively searches down a node before backtracking to higher nodes, then searching down these nodes, and so on. DFS is not complete nor optimal, because if there were an infinitely deep node, search may continue without finding a goal. For Figure 1, the search order for DFS is: α, A, β, γ, B, C, δ, ϵ, E, F, $D_{[F]}$. Due to exploring the depth of a node, here DFS returns the target last, less efficient than BFS. Thus neither method is maximally efficient. A more profitable strategy is iterative deepening depth-first search (IDDFS) (Korf 1985), which gradually increases the depth limit, with initial depth limit=1, until the goal is found; each iteration uses DFS. IDDFS combines the benefits of DFS and BFS, being complete, optimal, and having modest memory requirements. IDDFS is the preferred *uninformed* (or "brute-force") search method[3] when the depth of the solution is unknown (Russell & Norvig 2010: 90), and is useful in the context when we cannot explore below a given depth (Konar 2000: 155); the latter point is effectively the PIC, which imposes a lower bound on MS. Further, IDDFS might mirror syntactic locality in another way, with a preference for going down the clausal/nominal spine (being depth-first). CED effects (Huang 1982) arise from the cost of going out of the spine, i.e. extracting from adjuncts and specifiers is worse than extraction of complements;[4] island constraints are depth restrictions on IDDFS. If, under MP

[3] These search strategies are uninformed (Russell & Norvig 2010: 92); MS is "blind" and does not know how far to look before a solution is found. MS might be informed, with problem-specific knowledge beyond the basic search problem, to find solutions more efficiently. This involves either a heuristic function to evaluate nodes, or searching nodes closer to the solution. Syntactic MS is informed about intervention and impenetrability to the extent that it does not search past the first instance of a feature (intervention) and it does not search previously searched structure (impenetrability).

[4] This suggests MS might prefer searching the clausal spine before the specifier/adjunct position, potentially instantiating one way MS is informed. Müller (2010) derives CED effects from the PIC (derived by MS-locality), so CED effects may follow from MS-locality (section 4.3.4 discusses the Adjunct Condition).

assumptions of optimality, MS is maximally computationally efficient, we should discard Ke's List-based BFS MS, and adopt a restricted version of IDDFS with no list/memory requirements, instead increasing the depth limit each iteration, as in (3):

(3) SA:

a. Given ST and SD, match against every head member of SD to find ST [initial depth-limit of SD = 1; search depth-first].

b. If ST is found, return the head(s) bearing ST and go to d. Otherwise, go to c.

c. Increase the depth-limit of SD by 1 level; return to a.

d. Terminate Search.

For Figure 1, cycle one searches α, finding A and β; the second iteration searches α again, then searches inside set β, finding sets γ and δ; the third iteration repeats these searches then searches within γ finding B and C, before searching δ and returning $D_{[F]}$ as the target.[5] An IDDFS version of MS is thus more economical than BFS, and is preferred. Further, IDDFS can, like BFS (Ke 2019: 49), capture c-command relations, always reaching a c-commander before its c-commandee.

The second point of contention is whether MS occurs in parallel or not. Ke argues for parallel BFS because it searches both set members in parallel without distinguishing between them as there is no principled argument for this. Although Ke's argument for treating both set members equally is intuitive, there may be reason to suppose that syntactic structure might have different properties from a computational tree. That is, complements have a privileged status compared to specifiers and adjuncts (cf. CED effects, Huang (1982), captured by L-marking in *Barriers* (Chomsky 1986)). Perhaps features in syntactic structure can be evaluated and inform MS; but this is speculation. Adopting IDDFS makes parallel MS more difficult, which may or not be a welcome consequence. I leave this matter open, considering parallel MS further in section 4.3.4 regarding Across-the-Board movement.

2.2.4 Psychological Reality

Finally I consider, in its capacity as the reflex of a third factor constraint, whether MS has extra-linguistic cognitive analogues. Rieman (1994) suggests that, in the context of task-oriented exploratory behaviour, humans *might* use IDDFS as a search method, but the hypothesis is unconfirmed, and also involves the use of a label following heuristic to limit search to items semantically relevant to the task (Rieman, Young & Howes 1996: 747). Fujita (2017), relates MS to actions, especially in the capacity of labelling; he argues that object-specific attention (akin to headedness) correlates with MS in its application to general cognition. This is a

[5] See footnote 4. If MS privileged the clausal spine over specifiers/adjuncts, δ may be searched before γ, returning $D_{[F]}$ two steps earlier. I treat MS as uninformed; informed MS requires appealing to cognition, which lies beyond the scope of this paper.

direct analogue of MS in a different cognitive domain to language, supporting its F3 status. Cassimatis, Bello & Langley (2008: 1308), studying chess playing by humans and computers, claim humans do not perform "brute-force" (i.e. uninformed) search, but there is significant evidence (cf. Dingeman 1978) that humans do perform some search, generally using task-specific heuristics. They stress computational search mechanisms quite different to human cognition can yield human-like patterns, claiming this as evidence that humans use these representations. This both helps and hinders the current account; models whose mechanisms which are unfaithful to human cognition can nevertheless help explain aspects of it. This supports the MP as a model of competence; but it weakens appeals relating MS directly to cognition. However, MS appropriates MC, an independently motivated third-factor strategy, using it in a specific way for language, applying to Merge's output. It is domain-general in that it is an instantiation of domain-general MC; the process of MS itself is domain-specific.

In this section, I have explained the motivation of the proposal, given a definition of MS that is maximally computationally efficient, and shown its grounding in computer science, cognition and F3s. Below I apply the definition of MS to intervention constraints.

3 Intervention Effects

3.1 Theoretical Background

In this section, I detail previous approaches to intervention, argue for a derivational approach derived by MS, and show how this approach fares empirically.

Intervention effects are universally observed phenomena, involving constraints relativized to the identity of the elements at issue. To account for these effects, Rizzi (1990) introduced Relativized Minimality (RM), a representational constraint on certain configurations:

(4) In the configuration …X…Z…Y… a local relation cannot connect X and Y if Z intervenes and Z is of the same structural type as X.

   a. Intervention: Z intervenes between X and Y when X c-commands Z and Y, Z c-commands Y and Z does not c-command C.

   b. Structural types:

      i. A' positions

      ii. A positions

      iii. Heads

In effect, RM prohibits an intervener between elements engaging in a local relation, e.g. (5) where head could intervenes between have and its trace:

(5) *Have$_i$ they could t$_i$ left?

Starke (2001) and Rizzi (2001, 2004) updated the framework to give a deeper characterisation of the coarse notion of structural type in terms of features, which accounts neatly for argument-adjunct asymmetries. For Starke, Featural RM (FRM) is:

(6) In the configuration X … Z … Y, a local relation cannot connect X and Y if Z intervenes and Z fully matches the specification of X and Y in terms of the relevant features.

For example below, argument movement is acceptable but adjunct movement is banned:

(7) ? [[Which problem$_{[+Q, +N]}$] [C$_{[+Q]}$ [do you wonder [[how$_{+Q}$] [C$_{+Q}$ [PRO to solve ~~which problem~~$_{[+Q, +N]}$]]]]]]

(8) * [[How$_{[+Q]}$] [C$_{[+Q]}$ [do you wonder [[which problem$_{[+Q, +N]}$] [C$_{[+Q]}$ [PRO to solve ~~how~~$_{[+Q]}$]]]]]]

In (7), *which problem*$_{[+Q,+N]}$ (an argument) can enter a relation with its trace because *how*$_{[+Q]}$ (an adjunct) does not fully match the featural specification of *which problem*; however in (8), *which problem* intervenes as it has a superset of *how*'s features, and thus fully matches its specification, meaning no local relation can hold. The marginal unacceptability of (7) is due to a partial featural match with *how*.

Chomsky (1995) proposed the Minimal Link Condition (MLC), a derivational condition on movement, where the closest relevant element in terms of featural identity must be attracted. Chomsky (2000) revises MLC as a locality condition on Agree, which he sometimes calls MS. The current proposal takes this general direction.

3.2 *Current Proposal*

For locality constraints to apply, some relation must exist. In the MP, UG consists of Merge, Agree and LA; the Government-Binding theory concept of movement is reduced to Merge. Chomsky (2001: 8) distinguishes External Merge (EM) i.e. merger of two elements external to each other; and Internal Merge (IM) i.e. the merge of two elements, one being internal to the other, yielding displacement. Movement is motivated by feature-checking Chomsky (1995) of uninterpretable, movement-inducing (edge) features.[6]

MS applies in two cases in the derivation:[7] there is MS for Merge, whereby a movement-inducing feature initiates search for an element to merge (i.e. move).

[6] The exact nature of movement-inducing features is controversial; I avoid entering the debate as it lies beyond the scope of this paper.

[7] Several authors argue search for Probe-Goal and Merge are different. Ouali (2010) explains Agree applies upon establishing a c-command Probe-Goal relation via MS, independently of IM, in which MS is initiated by an edge feature; Chomsky (2015a) suggests while Merge is a different system to MS, simple search (i.e. MS) is required to select an item for Merge, a separate search to Probe-Goal MS.

Second, MS applies in the Probe-Goal relation case (Chomsky 2000, 2001), encapsulating Agree and LA. Chomsky (2015a: 80-81) notes Probe-Goal is simply a search procedure,[8] initiated by an unvalued feature [uF], then once in the relation, valuation occurs; Probe-Goal is thus a relation established by MC, which facilitates feature valuation. Importantly, Merge is not reduced to MS, but involves its own MS procedure. There are thus two cycles of MS which induce locality constraints. I argue that MS for Probe-Goal derives intervention effects (subsuming RM and A-over-A effects), while impenetrability constraints are derived from MS for Merge (cf. section 4).

In Probe-Goal, a [uF] probes the structure for a valued [vF]. The probing is MS, and thus finds the highest probe for the goal (as IDDFS MS does) (Chomsky 2011: 27); this is akin to the MLC. MS for Probe-Goal is a restraint on Agree, then; in other words, Agree Closest (Bošković 2007a) is MS. MS is a reflex of MC acting on the derivation, meaning the configuration is ungenerable (as opposed to a condition on representations e.g. RM which overgenerates then filters out at a later step). For Agree, ST = [uF] on the probe that initiates the search; SD = the sister of the head with [uF], and the output of MS is the head bearing ST, with which valuation occurs. Valuation involves copying [vF] values to corresponding [uF]s, and is not reducible to MS (*contra* Chomsky 2015a).

We can summarise the proposal as follows, formalising Agree Closest as below:

(9) Probe Closest Goal (PCG)
A probe for feature [F] returns the structurally closest goal with feature [F] in its search space, enforced by MS.

3.3 Examples

Consider a simple example:

(10) *Have$_i$ they could t$_i$ left?

(10) is ungrammatical because *have* cannot enter a relation with its trace due to the intervening head *could*, which has the same featural specification. This is because MS forces PCG (i.e. Agree Closest) with the nearest head, *could*, causing a crash.

Consider now an example of pure Agree with unaccusatives in Italian:

(11)

Sono	*arrivati*	*gli*	*studenti.*
be.3.PL	arrived.M.PL	the	students.M.PL

'The students have arrived.'

[8] Chomsky argues Probe-Goal should not be taken too literally: e.g. tense 'searches' for the subject, but it is just a minimal computational relation that exists, allowing feature valuation both ways.

The standard view on Italian is that subject-verb agreement is the result of T probing a DP. While for transitives and unergatives (13-15) agreement is with the external argument (EA), for unaccusatives agreement occurs with the internal argument (IA), which remains in-situ (unlike English). In (11), *gli studenti* is the internal argument position as the partitive clitic *ne* can be associated with it, and this cannot associate with subjects Belletti & Rizzi (1981) e.g. in (13, 14):

(12) *Ne sono arrivati molti.*
PRT be.3.PL arrived.M.PL many.PL
'Many of them have arrived.'

(13) * *Ne hanno telefonato molti.*
PRT have.3.PL phoned many.PL
'Many of them have phoned.'

(14) * *Molti ne hanno mangiato le mele.*
many.M.PL PRT have.3.PL eaten the apples.F.PL
'Many of them have eaten the apples.'

(15) *Ne hanno mangiato molte.*
PRT have.3.PL eaten many.F.PL
'They (null) have eaten many of them.'

In (12), T is probing *molti (ne)*, which is in the direct-object position. If v is absent in unaccusatives, then it does not intervene, and the sole argument is the closest even in object position, thus agrees with T. What this shows is that PCG holds; in the absence of an EA, T agrees with IA. If EA is present, PCG via MS means T cannot agree with IA, cf. (13, 14), but must agree with EA, (15).

Tsez provides another PCG example (Polinsky & Potsdam 2001 in Bošković 2007a: 7; elements in the agreement relation are bold-faced):

(16) *eni-r [už-ā* ***magalu*** *bāc'rułi]* ***b****-iyxo.*
mother-DAT boy-ERG bread.**III**.ABS ate **III**-know
'The mother knows the boy ate the bread.'

(17) * *eni-r [už-ā* ***magalu*** *bac'si-ƛin]* ***b****-iyxo.*
mother-DAT boy-ERG bread.**III**.ABS ate-**COMP** **III**-know.
'The mother knows that the boy ate bread.'

In (17), the overt complementizer CP has ϕ-features and is closer to the agreement target head than an NP embedded within the CP, so PCG via MS means the CP is returned first and blocks agreement with the NP. In (16), the embedded clause is suggested to be a TP (Bošković 1997), which does not block long-distance agreement. This is because Tsez TPs have no ϕ-specification (Bošković 2007a), meaning MS does not return TP as an intervener, so long-distance agreement can occur. This neatly exemplifies PCG.

Consider now an example of intervention in Korean wh-in-situ licensing:

(18) ?* *Mira-man nwukwu-lul manna-ss-ni?*
Mira-only who-ACC meet-PST-Q
'Who did only Mira invite?' (Kim 2006: 28)

(19) *Nwukwu-lul$_i$ Mira-man t_i manna-ss-ni?*
who-ACC Mira-only meet-PAST-Q
'Who did only Mira invite?'

In (18), the Focus operator *-man* (only) intervenes between the interrogative C and the wh-in-situ object *nwukwu-lul*. The wh-phrase must be licensed by interrogative C to be interpreted but in (18), the focus-sensitive operator *-man* intervenes, blocking C licensing *nwukwu-lul*, because PCG via MS finds Focus$_{[+\text{Foc}]}$ before *wh*$_{[\text{uFoc}]}$, seen below:

(20) $[_{\text{CP}}\ \text{C}_{\text{+Q, +Foc}}\ [\ \ldots \text{Focus}_{[+\text{Foc}]}\ \ldots [\ \ldots \text{wh}_{[\text{uQ, uFoc}]}\ \ldots]\ ]\ ]$

Here, PCG forces a relation between C and Focus, the closest element bearing ST [+Foc], meaning Agree cannot relate C and *wh* so *wh* goes unlicensed, causing a crash. The (partial) ungrammaticality of (18) forces scrambling, (19), to derive a grammatical output (Ko 2018) and avoid intervention, because now *wh* c-commands Focus. The marginal unacceptability of (18) is due to Focus' partial featural match with C$_{[+\text{Q, +Foc}]}$ as both Focus and *wh* are [Foc]-specified, but *wh* is also [Q]-specified. This FRM-type analysis also explains why a [+Foc]-specified wh-element can move over [+Foc]-specified Focus in (19): *wh*$_{[\text{uQ, uFoc}]}$ is more richly specified than Focus$_{[+\text{Foc}]}$ so MS continues past the partially matching Focus to the fully matching *wh*, then *wh* can move above Focus and avoid intervention. Once in the configuration in (19), Agree via PCG can then relate C and *wh*. The MS-locality analysis thus explains (19)'s acceptability, and gives a satisfactory account of intervention effects.

PCG, a condition on Agree, thus derives intervention effects as involving impossible relations in a derivation. Such effects fall out naturally from MS; they are essentially the same thing. The force of this section is that, for the basic paradigm of intervention effects, a derivational, F3-based MS-locality account proves capable of capturing a wide array of empirical data.

4 Impenetrability Effects

4.1 Theoretical Background

4.1.1 Subjacency and Barriers

Since Ross' (1967) recognition of island constraints, an approach to successfully unify them has remained notoriously elusive. Ross described island constraints in terms of movement bans, but gave little deeper characterisation; Chomsky's (1973, 1977) Subjacency principle attempted such a characterisation, stating that movement cannot occur over more than one bounding node at a time, where bounding nodes are TP and DP. In (21), while the first movement step only crosses one bounding node, the second step is banned because two bounding nodes are crossed:

(21) *[$_{\text{CP1}}$ Which person$_i$ did [$_{\textbf{TP1}}$ you believe [$_{\textbf{DP1}}$ the allegation [$_{\text{CP2}}$ t_i' that [$_{\textbf{TP2}}$ we had seen t_i]]]]] ?

However, Subjacency could not account for *all* cases of the Coordinate Structure Constraint (CSC), the Adjunct Condition (AC), and the Subject Condition (SC). Also, in (22), movement only crosses one bounding node in each step, so should not be banned:

(22) *[$_{\text{CP}}$ Who$_i$ was$_j$ [$_{\textbf{TP1}}$ [$_{\text{CP}}$ t_i that [$_{\textbf{TP2}}$ you met t_i]] t_j unexpected]] ?

In a revised approach, Chomsky (1986) proposed the *Barriers* framework, which explains why constructions are barriers based on structural context. Complements of lexical heads have a special property of being L-marked, preventing barrierhood. A category is a blocking category if it is not L-marked; a barrier is a blocking category except TP; and a construction immediately dominating a blocking category inherits barrier status. Movement can then only cross one barrier. This approach provided a deeper, *contextual* characterisation of blocking categories, and L-marking captured the fact that it is easier to move out of complements (Cook & Newson 2007). However, *Barriers* has issues unifying island constraints, and it is a complex, highly stipulative framework. Both Barriers and Subjacency have been all but abandoned in favour of a different theory of impenetrability locality.

4.1.2 Phase Theory

Chomsky's (2000, 2001, 2008) Phase Theory (PT) represents an important development in impenetrability locality. PT attempts to capture the recurrent notion that only a subset of structure is available at any time for further operations, and that this window is relative, moving as structure is built. This is captured via the notion of phases, i.e. that at certain points in the derivation, part of the structure (specifically, the complement of the phase head) becomes fixed and cannot be manipulated further. This is because once a phasal complement has been built, it is sent to spell-out; this occurs cyclically (Chomsky 2000, Uriagereka 1999).

However, cyclic spell-out is problematic when considering the recombination[9] that must occur at the C-I interface. Recombination is not difficult to formalise (Chomsky 2005a), but Boeckx & Grohmann (2007: 209) note recombination cannot amount to the cumulative outcome of cyclic spell-out because derivations proceed in parallel, and that there is no optimal justification of recombination. Recombination suggests the interfaces mirror the work in the syntactic component, effectively carrying out the derivation twice – just as LF cycled over the same chunks of structure as S-structure did in Government-Binding theory. This redundancy cannot easily be overcome. The current approach is agnostic towards implications for spell-out, and thus makes no comment on recombination – a point which I take to be an advantage of this proposal.

To be available for further movement, elements must move to phase edges, i.e. the specifier of the phase head; otherwise they would be inaccessible because the entire complement has been transferred to spell-out. This is Chomsky's Phase Impenetrability Condition (PIC_1). Chomsky argued vP and CP are phases, because vP is thematically complete and CP is propositionally complete, containing the proposition and force markers. However, such a non-contextual characterisation seems inadequate, and echoes Subjacency's fixed bounding nodes; e.g. DP and NP also act as phases in DP and non-DP languages respectively (Bošković 2005).

Bošković (2013, 2014) offers a solution with a contextual approach to phases, which I adopt here. Under this approach, a phase is the highest projection in the extended domain of a lexical head (2014: 2). Bošković collapses the three domains of the clause into two, with the discourse (C) and inflection (T) collapsed together. This gives C and v as phase heads in the usual case, but can change depending on context. The second aspect of the approach is to reformulate Chomsky's PIC_1 as follows (Bošković 2015: 9):

(23) PIC_2
In a phase α with head H, only the immediate domain of H is accessible to operations outside of α, where K is in the immediate domain of H if the first node that dominates K is a projection of H.

Essentially, YP, the complement of H, is accessible but nothing *inside* YP is visible, indicated by the dashed line in Figure 2.

For reference, Chomsky's PIC_1 sets the boundary for visibility *above* YP, because for Chomsky the entire complement is sent to the interfaces, indicated by the dotted

[9] Binding Theory exemplifies issues of recombination. Principle C appears not to respect locality conditions, e.g. (a) shows the interpretation of a pronoun as a bound variable when c-commanded by a quantifier:

(a) Every politician always betrays the people who vote for him.

Here, the only condition for binding is c-command. *Him* can be bound by *every*, i.e. for every politician x, x betrays the people y who vote for x. *Him* is inside a CNPC island, from which extraction is banned, but variable binding is acceptable. At C-I, multiple phases have been transferred, but the structure must be recombined so that *every* c-commands *him*, meaning *him* is interpreted as a variable. Despić (2011) shows cases where even c-command is unnecessary, suggesting non-syntactic factors at work.

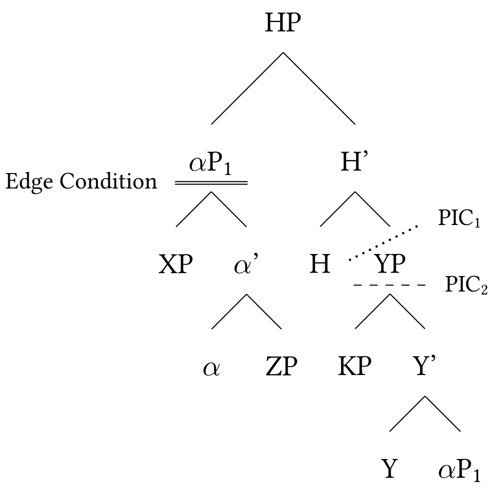

Figure 2

line. Bošković's PIC_2 follows Uriagereka's (1999) original conception of multiple spell-out, arguing that when an element A is sent to spell-out, nothing within the phrase is available for further syntactic operations but the phrase itself is available. Sending A to spell-out results in the establishing of word order within A, turning A into a lexical item whose internal structure is inaccessible to syntax (N.B. PIC_1 does not capture this). Bošković's PIC_2 also induces an Edge Condition (the double line), capturing Hiraiwa's (2005) observation that what is located at the edge of the edge of phase HP is not at the edge of HP for PIC_2. In Figure 2 above, anything in spec,αP (here XP) or adjoined to αP is not located at the edge of HP, and thus is not accessible to operations outside HP. This follows from the PIC_2.

4.2 Current Proposal

In this section I derive the PIC_2 from MS for Merge without stipulating it as an iron curtain. The PIC is a lower-bound on MS. Phase heads are exceptional, as they cause MS to terminate. I assume phase heads introduce an uninterpretable 'edge' feature [F_E] which induces movement[10] (cf. Gallego's (2010) phase condition, Chomsky (2000, 2015a),[11] Larson (2015)[12]). Phases, and thus the PIC, are intimately connected with the valuation of unvalued features.

[10] That phase heads trigger movement is not universally accepted. Bošković (2007b, 2011) argues for movement features to be specified on moving-elements, in a Greed approach, with compelling motivations; however, this conflicts with the assumption that phase heads introduce movement features, core to MS-locality. One issue that would need to be overcome is how island effects are induced if an element can move for its own reasons. Nunes' (2014, 2016) hybrid approach to edge features offers a possible foundation for reconciling MS-locality and Bošković's approach.

[11] Chomsky (2015a) suggests uninterpretable features identify phases, i.e. points where strict cyclicity applies.

[12] Larson (2015: 61) argues we can at least assume that phase heads always introduce movement features; Richards (2007) and Chomsky (2013) may be incorrect and movement-inducing (edge) features may

Unvalued features trigger movement. For movement, i.e. IM, to occur, an appropriate target T must be found, via MS for Merge; T is appropriate if it has a feature [F] that can check, i.e. value, [F_E] via movement into the phase edge (i.e. the immediate structural domain of phase head H) (Bošković 2015: 9). Consider Figure 2 above.

(24) $[_{H'}$ H $[_{YP}$ **KP** $[_{\mathbf{Y'}}$ **Y** $[\boldsymbol{\alpha}\mathbf{P_1}]$ **]**]]]

In (24), αP is being searched for to merge in spec,HP. MS must search the structure top-down to find αP, carrying feature [F] to check H's [F_E]. MS searches the bold material; this already has parallels with PIC_2. To capture PIC_2 I propose MS for Merge is minimal to the extent that:

(25) MS can only search a structure once.

(26) MS must begin in the set below the set containing phase head H.

These minimality conditions mean any further operation in Figure 2 attempting to move an element cannot search past YP, capturing the PIC_2 as a lower bound on MS.

(25) is a natural minimality condition, as searching the same structure for the same reason (i.e. MS for Merge) would introduce redundancy into the derivation. (25) makes MS maximally minimal, a reflex of MC for brute-force search.

(26) states that MS is irreflexive. MS searches merged pairs; given MS for Merge is initiated by H, MS cannot search its immediate merged pair, {H, YP}, as otherwise it would find H as the goal to its own probe, causing a crash. Only searching YP would be asymmetrical and under current assumptions of minimality, MS is symmetrical. MS must then start in the embedded merged pair, {KP, Y'}, to avoid failure.[13] The combination of (25) and (26) thus derives the PIC_2.

(26) is reminiscent of antilocality (AL) (Abels 2003, Bošković 1994, 2015, Grohmann 2003, Ticio 2005), a derivational ban on movement that is too short.[14] Indeed, (26) provides a natural characterisation of Bošković's AL (2015: 11):

(27) Movement of A targeting B must cross a projection distinct from B (where unlabelled projections are not distinct from labelled projections).

In other words, movement must cross a phrase. (26) forces this, given that search must begin in the next phrase, i.e. the set immediately below H's set. Whereas in Bošković's approach, AL is stipulated, a major advantage of the current proposal is that I derive both the PIC_2 and AL from the same source, MS.

---

arise with any head. In fact IM must be forced at non-phase positions as well, e.g. for A movement, which is XP movement to non-phase edges, canonically spec,TP.

[13] MS' irreflexiveness represents another way MS is informed.

[14] Although not an issue specific to MS-locality, AL conflicts with current accounts of roll-up in deriving word order. It is possible that a return to Kayne's (1994) proposals, where roll-up movement is higher than complement to specifier, might reconcile AL and roll-up, although the exact details remain open.

Finally, (25) and (26) account for the Edge Condition (Bošković 2014, 2016a). When a phrase moves, MS must search down to the head to find the relevant feature [F] on α, returning α as the target; the entire phrase αP is merged with H', forming HP, with αP in its edge. Given that the entire phrase αP was searched to find [F], (25) applies for further operations, meaning the internal structure of αP becomes fixed and inaccessible, with αP itself remaining visible.[15] The Edge Condition is thus also accounted for.

4.3 Island Constraints

4.3.1 Complex XP Constraints

I now test this derivation of the PIC_2 empirically across island effects, beginning with Bošković's (2015) Complex XP Constraint (CXPC). Ross' (1967) Complex NP Constraint (CNPC) states that:

(28) Extraction from complex NPs is disallowed.

For example, (29) shows banned extraction from a complex NP:

(29) ?? Who did you hear [$_{NP}$ claims [$_{CP}$ that [$_{TP}$ a bird [$_{vP}$ pecked t_i]]]] ?

To clarify, NP, not DP, is the issue, based on evidence from non-DP languages e.g. Serbo-Croatian (Bošković 2015), where the CNPC holds, and DP locality issues i.e. the Left Branch Constraint, generally do not arise. Moving to Bošković's (2015) phasal explanation of the CNPC, consider (30), with an accurately labelled structure:

(30) ?? Who$_i$ did you hear [$_{DP}$ [$_?$ t_i'' [$_{NP}$ claims [$_?$ t_i' [$_{CP}$ that [$_{TP}$ a bird [$_{vP}$ pecked t_i]]]]]]]] ?

Here, NP and CP are phases, being the highest projections in the domains of lexical heads; I assume N and C introduce edge features, so movement targets CP and NP. Given Chomsky's (2013) Labelling Algorithm (LA), movement must occur through unlabelled projections (with labels resolved later via traces). Movement is forced to phase edges, either by adjoining to the phase projection HP or by creating unlabelled projections (Bošković 2015: 16). (30) is underivable because movement from t' to t'', forced by N's edge feature, violates AL, not crossing a labelled projection (see (27)). The PIC_2 is derived here because N's edge feature initiates MS for Merge which searches down to t', finding *who*; MS then cannot search within NP again. If MS searched down to t', failed to select *who* for movement, then later attempted to search NP again to find *who*, search would be prohibited by (25). The ungrammaticality of (30) is derived by the conspiring of AL and PIC_2: the short movement step is banned by antilocality, and the long movement step would

[15] There remain questions regarding how MS finds an element within a phrase if e.g. the features project to give TP as ϕP, and MS can only see the features. It is expected that the answer relates to the timing of labelling and of A movement, but this question remains open.

involve two violations – searching and finding *who* but not moving it, and violating (25) by searching NP twice.

This proposal, instead of the PIC_2 being a boundary *over* which movement cannot occur, forces movement to a phase edge when possible. If movement occurs after an edge feature is introduced and initiates MS, the PIC_2 is violated. In MS-locality then, the PIC_2 is more related to timing than distance of movement. Non-CP N complements are also accounted for, e.g. in Greek (Horrocks & Stavrou 1987 in Bošković 2015: 11):

(31) * *Tu* *vivliu*$_i$ *mu* *ipes* *pos* *dhiavases* *[*$_{DP}$ *[*$_?$ *t*$_i$*"* *[*$_{NP}$ *entasi*
the-GEN book-GEN$_i$ me said.2.SG that read.2.SG objection
[$_?$ *t*$_i$*'* *[*$_{DP}$ *tis* *[*$_{NP}$ *kritikis* *t*$_i$*]]]]]]*
the review-GEN

'You told me you read the objection to the review of the book.'

Movement must target NP and the embedded DP and NP, but movement from t' to t" violates AL, and movement from t' to the matrix clause violates the PIC_2. The ban on extraction is, by extension, from all nominal complements. MS-locality thus captures the CNPC.

The CXPC captures the fact that APs pattern with NPs, as extraction from clausal complements of APs is disallowed, as below:

(32) *How$_i$ are you [$_?$ t$_i$" [$_{AP}$ happy [$_?$ t$_i$' [$_{CP}$ that Peter hired Alex t$_i$]]]] ?

As for NPs, non-CP complement extraction is also banned, shown below in a Greek example (from Bošković 2015: 6):

(33) * *Tu* *ktiriu*$_i$ *ipefthinos* *[gia* *to* *fotismo* *t*$_i$*]*
the-GEN building-GEN is-responsible for the lighting

'The building he is responsible for the lighting of.'

The analysis for the Complex AP Constraint mirrors the CNPC: in (32), the second step of movement from t' to t" violates AL, not crossing a labelled projection; and given AP is a phase (being the highest projection in its domain), long distance movement from t to the matrix CP is banned by the PIC_2 given MS is initiated and has searched AP.

Finally, PPs pattern with NPs and APs in extraction bans (Landau 2009), leading to the positing of the Complex PP Constraint. Interestingly, the same constraint holds for pied-piping:

(34) ??Who$_i$ did you read [$_?$ t$_i$" [$_{PP}$ about [$_?$ t$_i$' [$_{DP}$ friends of t$_i$]]]] ?

(35) *Of who(m)$_i$ did you read about friends t$_i$?

The analysis above holds: in (34), the second step of movement from t' to t" violates antilocality, but long-distance movement to the matrix CP violates the PIC$_2$. As with NPs and APs, extraction from other P-complements is also banned, e.g. Dutch (van Riemsdijk 1997):

(36) *Hij kan zich niet [in [de bibliografie [van dat boek]]] vinden.*
he can himself not in the bibliography of that book find

'He cannot find himself in the bibliography of that book.'

(37) **[Van dat boek]$_i$ kan hij zich niet [in de bibliografie t$_i$] vinden.*

This leads Bošković to posit the CXPC, where a lexical head excludes V:

(38) Extraction out of a lexical head is disallowed.

The current MS-locality approach successfully derives this constraint.

4.3.2 Left Branch Constraint

The Left Branch Constraint (LBC) bans Left Branch Extraction (LBE), i.e. movement of the leftmost constituent in an NP (Ross 1967), blocking extraction of determiners, possessors and adjectives from NP, as well as blocking adjunct extraction from NP. For example:

(39) *Beautiful, he saw [$_{DP}$ [$_{NP}$ t$_i$ houses]].

Some languages, e.g. Serbo-Croatian, allow LBE, because they lack DP (Bošković 2005: 2).

(40) *Lijepe$_i$ je video [$_{NP}$ t$_i$ kuće]*
beautiful is seen houses

'Beautiful houses, he saw.'

Here I adopt Bošković's (2005) phasal account of LBC. The main point is that DP is a phase, which follows from assuming contextual phases (Bošković 2013), being the highest projection in the domain of a lexical head, meaning D introduces an edge feature. Assuming AL, we can derive the ungrammaticality of AP LBE in (39):

(41) *[$_{DP}$ AP$_i$ [$_{D'}$ D [$_{NP}$ t$_i$ [$_{NP}$ …

(42) * AP_i $[_{DP}$ $[_{D'}$ D $[_{NP}$ t_i $[_{NP}$ …

(41) is ruled out by AL, as AP does not cross a distinct projection. (42) is ruled out by the PIC_2. D introduces an edge feature, initiating MS when merging for spec,DP; MS searches NP and finds AP *beautiful*. Either AP moves to spec,DP as in (41) and violates AL; or AP is found but does not move, and later MS is initiated again, violating the PIC_2 by searching inside NP twice to find AP again to move it. The MS-locality analysis of the PIC_2 and AL thus derives the LBC.

An example of successful extraction out of DP is (43), where *who* moves through spec,DP:

(43) Who_i do you hate $[_{DP}$ $[_{NP}$ friends of t_i]]?

This has the below configuration, where AP can move far enough not to violate AL, while still obeying PIC_2 by moving to spec,DP:

(44) $[_{DP}$ AP_i $[_{D'}$ D $[_{NP}$ $[_{N'}$ $[_{PP}$ t_i …

Non-DP languages exhibit different LBE patterns, e.g. Serbo-Croatian (all examples from Bošković (2005)). This is explained by assuming contextual phases, where NP is the highest projection in its domain so is a phase. This rules out the structure below:

(45) * *Čije$_i$ je on vidio $[_{NP1}$ prijatelia $[_{NP2}$ t_i $[_{NP3}$ majke]]]?*
whose is he seen friend mother

'Whose mother did he see a friend of?'

NP_1 is a phase, introducing an edge feature; movement from the position adjoined to its complement is ruled out by the PIC_2. When merging for spec,NP_1, MS is initiated, finding *Čije*; movement to spec,NP_1 violates AL, and later movement past spec,NP_1 would violate the induced PIC_2, as NP_2 has already been searched. Bošković (2005) notes the improved status of (46) is accounted for by Chomsky's (2001) proposal that locality and the PIC are evaluated at the next phase level, which involves look-ahead. Given this assumption, movement of NP_3 out of the object position NP_2 is acceptable since at the point of evaluation, NP_2 takes a trace complement, so its maximal projection is not a phase (Bošković 2005).

(46) (?)? *Čije$_i$ je on $[_{NP3}$ t_i majke$]_j$ vidio $[_{NP1}$ prijatelja $[_{NP2}$ t_j]]?*

MS-locality explains Chomsky's proposal because the PIC_2 is not visible from below, when movement to the phase edge occurs. MS forces phase escape to the phase edge, searching the phase complement; the PIC_2 is only visible from 'above', when MS for Merge reaches the boundary of what has already been searched (the internal structure of the phase complement and edge), essentially the PIC_2. This proposal reduces Chomsky's stipulation involving look-ahead to the natural action of MS. Finally, (47), involving double AP LBE, followed by remnant AP fronting, is accounted for under the current proposal:

(47) *[$_{AP}$ *Visoke*]$_i$ [$_{AP}$ *lijepe*]$_j$ *on gleda* [$_{NP}$ t_i [t_j *[djevojke]]]*
tall beautiful he watches girls

'He is watching tall beautiful girls.' (Bošković 2005: 12)

The APs cannot be moved together since they do not form a constituent, being adjoined to NP. If APs undergo separate LBE, (47) is ruled out as a PCG-intervention violation, since an AP would move over an AP, meaning no local relation can hold between the APs and their traces. This is interesting, because this constraint can be explained by intervention and MS for Probe-Goal. Under MS-locality, both intervention and impenetrability effects are derived from MS, so the similarity and occasional redundancy between accounts of the two effects are expected, supporting unifying the two effects. MS-locality thus accounts for the full paradigm of LBE in DP and non-DP languages.

4.3.3 Coordinate Structure Constraint

The Coordinate Structure Constraint (CSC) is notoriously resistant to a satisfactory account. Below, I detail Oda (2018) and Bošković's (2018) accounts and tentatively consider how MS-locality can account for them. The CSC in fact is two island constraints, CSC-1 and CSC-2:

(48) Extraction *of* conjuncts is banned.

(49) Extraction *from* conjuncts is banned.

A CSC-1 example is:

(50) *[Which table]$_i$ will he buy [t_i and the chair]?

In this paper, I assume conjuncts can be analysed as (51), with a Spec-Comp relation (Zhang 2010):

(51) [$_{ConjP}$ XP [$_{Conj'}$ Conj^0 YP]]

CSC-1 violations pattern with LBE in non-DP languages, and receive the same explanation, e.g. LBE and CSC-1 are allowed in Serbo-Croatian (Stjepanović 2014):

(52) *Kakvom*$_i$ *ga je pretnja* [t_i *smrću*] *uplašila?* (LBE)
what.kind.GEN him is threat death.INS scared

'The threat of what kind of death scared him?'

(53) *Zatvorom ga je pretnja [t_i i ubistvom] uplašila.* (CSC-1 violation)
prison.INS him is threat and murder.INS scared

'The threat of prison and murder scared him.'

Both LBE and CSC-1 are allowed from inherently case-marked noun (and adjunct) complements (52, 53) (Stjepanović 2014). In non-DP languages, with the structure of (51) ConjP is a phase because it is the highest projection in the domain, so conjunct 1 (C1) XP can move out of ConjP after spell-out without violating the PIC or AL, in the same way as adjective LBE. C2 cannot be extracted because either it violates the PIC by moving past (i.e. later than) the spec,ConjP, or it violates AL by moving to spec,ConjP.

In non-DP languages, e.g. Serbo-Croatian, CSC-1 can be accounted for by adopting Talić's (2015) Structural Parallelism (SP): if a language always requires a functional layer in the domain of one lexical category, it has a functional layer in the domain of all lexical categories. In DP languages, NP and AP have functional projections above them (DP and F_{Adj}P). ConjP is a functional projection, but it is standardly assumed that ConjP inherits the nature of lexical projections from its conjuncts (Oda 2018).[16] Oda suggests ConjP has an unspecified categorial feature determined by feature sharing (Chomsky 2013) with its conjuncts. We can then, under SP for DP languages, posit a functional category, F_{Conj}P above ConjP. The relevant structure is then:

(54) $[_{FConjP}\ [_{FConj'}\ F_{Conj}{}^0\ [_{ConjP}\ XP\ [_{Conj'}\ Conj^0\ YP]]]]$

F_{Conj}P is now the highest projection in the domain and thus is a phase, so the $F_{Conj}{}^0$ head[17] is assumed to introduce an edge feature. Extraction of C1 in (50) is banned because either C1 moves to Spec,F_{Conj}P and violates AL, or MS searches and finds *which table*, does not move it, then MS is initiated later, searching ConjP again to find *which table*, violating the PIC_2.

Movement of C2 is banned by intervention and PCG; C1 has the same specification as C2, thus no local relation can hold between moved C2 and its trace because C1 intervenes. Again, the conspiring of intervention and impenetrability effects supports the current proposal to unify the two; what in previous analyses looks like the combination of two separate constraints is here explained by the action of the same syntactic process MS acting on Agree and Merge. CSC-1 thus receives a natural characterisation under MS-locality, explained by the PIC_2, AL and PCG.

[16] It is debated whether conjuncts have a unique *Conj* label (cf. Kayne 1994, Stjepanović 2014) or whether ConjP lacks inherent categorial features, inheriting relevant feature specifications from its conjuncts (Zoerner 1995). The intuition of the latter stance is that when NPs, APs or VPs are conjoined, the whole structure continues to function as (a larger) NP, AP or VP. Assuming inheritance occurs, the ConjP label should not be taken literally; I use it here abstractly, assuming the categorially unspecified ConjP inherits the categorial status and label of the specific lexical projection of its conjuncts. Importantly this means in non-DP languages, the ConjP structure is the highest projection in the domain of a lexical head, so is a phase.

[17] Oda (2018) suggests *both* is a possible $F_{Conj}{}^0$ head; when present, *both* blocks extraction of C1.

The CSC-2, (55), is a more complex island, primarily due to the Across-The-Board (ATB) movement exception, (56).

(55) *Who$_i$ did you see [[enemies of t$_i$] and John]?

(56) Who$_i$ did you see [[enemies of t$_i$] and friends of t$_i$]?

Bošković (2018) provides an account of CSC-2 with LA, PIC$_2$, and the Coordination of Likes[18] (CL) principle (Chomsky 1957, Williams 1978), that conjuncts must be parallel in categorial status. If movement occurs to the edge of only one conjunct, this delabels it according to Bošković's (2016b) Timing of Labelling account, meaning the labels of C1 and C2 are mismatched, violating CL and causing a crash. CSC-2 thus follows from CL, an interface condition required for interpretation (Oda 2018).

I now turn to the ATB exception. Under Bošković's account, movement of A to the edge of both conjuncts[19] gives a {XP, YP} structure (because neither A is a trace with no chain being formed), with no feature sharing, making each conjunct unlabelled, meaning they are parallel in their categorial status, and CL is not violated. According to Oda (2018: 39), if C1 and C2 are unlabelled, then ConjP is unlabelled, and F_{Conj}P is unlabelled because it is an extended functional projection of the ConjP domain, meaning it shares the same categorial status with the lowest phrase in that domain. Bošković (2018) suggests unlabelled syntactic objects cannot be phases; there is no way to determine whether the unlabelled node is the highest phrase of an extended projection since categorial information for determining an extended projection is not provided (Oda 2018) – so F_{Conj}P is not the highest head of an extended projection and cannot be a phase, and thus ATB movement is not blocked by the PIC$_2$.

MS-locality derives ATB movement as follows. If an element is a phase, it introduces an edge feature. MS derives the PIC$_2$ because it searches feature bundles, and cannot search the same bundles twice. According to Bošković (2018), projecting features requires projecting a label, so unlabelled elements do not project features, and MS cannot return unlabelled elements. I tentatively suggest a link between labelling and the edge feature[20] – F_{Conj}P is unlabelled and no PIC$_2$ is induced, suggesting MS is not initiated, and elements can move beyond spec,F_{Conj}P. MS-locality thus gives a preliminary account for the availability of the ATB exception, but a satisfactory account of ATB movement remains elusive.[21]

[18] See Bayer (1996) and Weisser (2015) for discussion of CL-violating asymmetric coordination.

[19] This is a potential issue with the account. It is unclear how *who* is motivated to move to the specifier of both conjuncts. A Greed-approach (see footnote 10) might explain this, but conflicts with the potential labelling-based approach to edge features.

[20] This may be wrong: if LA drives successive cyclic movement (Chomsky 2013), the resulting labels would often be incorrect, and undoing such labels would leave unlabelled structure at C-I. This issue is left open and requires further work.

[21] Multiple conflicting accounts of ATB movement exist; Bošković & Franks (2002) suggest a null operator approach; Gazdar, Pullum, Sag & Wasow (1982) explain ATB via a categorial matching requirement (cf. CL); Zhang (2010) suggests ATB movement does not exist, but is instead movement of X from C1, with a pro-ϕP argument in C2, linked by identity to X. However, if null resumptives are allowed

Bošković (2018), following Nunes (2004), suggests ATB movement of X occurs from C2 to C1 via sideward movement, followed by wh-insertion of another copy of X in the matrix CP, which forms two separate chains with each lower X, linearizing them. This is undesirable because wh-insertion is awkward if movement is IM (operating by MS for Merge); and sideward movement (cf. multidominance, Citko 2005) is not an instantiation of binary Merge, "violat[ing] the minimal search requirement for Merge itself" (Chomsky 2015a: 82). In a ternary relation, an element A is found in the workspace, an element B is found inside this, and a third element C is found to which B is attached. It is not clear what constitutes a phase under this approach, and the general issue of distinguishing copies from repetitions stands.[22] It is undesirable then to posit sideward movement for ATB phenomena.[23] Bošković & Franks (2002: 123) suggest the two elements move to the same position, with PF-deletion under identity,[24] also assumed by Oda; this seems a less bad approach but is difficult to instantiate formally (Bošković & Franks 2002). What is clear is that no account of ATB movement implies parallel MS, so positing parallel MS seems unnecessary. In sum, the details of ATB movement remain open. The current approach does, however, give a neat characterisation of CSC-1, CSC-2 and the ATB exception via the PIC_2 and labelling.

4.3.4 Adjunct Condition

The Adjunct Condition (AC) is another difficult constraint. Bošković (to appear) takes adjunction structures to involve coordination (cf. Higginbotham 1985), with the same structure as (51), making possible a unification of AC and CSC. Below is an AC-violating example:

(57) ?*What$_i$ did you [$_{VP}$ [$_{VP}$ fall asleep] [after John had fixed t$_i$]]?

The AC involves the same derivation of ungrammaticality as the CSC-2, and thus I will not repeat it here. What is important is that because adjuncts and coordination structures are the same structure, they share several similarities. This is evident because extraction is exceptionally possible out of both, in parasitic gaps (PGs) (58) and ATB movement respectively:

(58) What$_i$ did you file t$_i$ without reading PG$_i$?

---

in islands, it is unclear how any island effects would be observed; and Zhang (undesirably) suggests ATB structures are quite different to CSC-2 violating structures. The details of ATB movement remain unclear.

[22] Chomsky's (2001) occurrences delineate copies from repetitions; see Collins & Groat (2018) for further discussion.

[23] Cf. Larson (2015), Adger (2017: 12) for further arguments against parallel Merge/sideward movement.

[24] Bošković & Franks (2002: 123) give the following example:

(b) I wonder [who$_i$ [~~who$_i$~~]] [Jane detests t$_i$] and [Harry adores t$_i$].

This derives the identity requirement on ATB dependencies, but technical details are left aside.

(59) *What$_i$ did you file [the book]$_i$ without reading t$_i$?

(59) is the AC equivalent of a CSC-2 violation. Given CED effects are an abstraction over the AC and Subject Condition, we should expect a similar exception. This is observed in (marginally acceptable) subject parasitic gaps:

(60) (?)I discovered which song$_i$ [everyone who listened to PG$_i$] failed to enjoy t$_i$.

Presumably, given the AC and CSC-2 are analysed in the same way, so should PGs and ATB. In both cases, extraction out of both conjuncts/parts of the adjunction structure saves the derivation. There are several attempts to unify PGs and ATB, including Nunes (2004) using sideward movement, but as discussed, sideward movement is undesirable. Bošković & Franks's (2002) proposal of ATB movement might be extended to PGs, but without clear formalisation, this remains open. Importantly, PGs appear not to involve parallel MS.

The AC and CSC-2 also have a similar semantic exception (Bošković to appear):

(61) What$_i$ did Christ die [to save us from t$_i$]? (AC; Truswell 2011: 131)

(62) The stuff which$_i$ Arthur sneaked in and [stole t$_i$] (CSC-2; Postal 1998: 53)

There is, according to Bošković, no good explanation for why the semantic condition voids the AC and CSC-2, but the crux is that AC and CSC-2 act in the same way, supporting their unification.

The AC and CSC-2 are different in that CSC involves conjoined [DP & DP], while adjunction involves [VP & Adjunct], giving an asymmetry where extraction is not banned from C1 (VP) in adjunction structures, cf. (63),[25] because V-to-v movement allows movement out of VP, obeying AL and PIC$_2$, rescuing its islandhood. As with the CXPC, the presence of v above V means VP behaves exceptionally.

(63) What$_i$ did you [buy t$_i$] slowly? (Bošković to appear: 10)

Further, many languages only have ATB, not PGs; and Bošković (p.c.) notes a PG-ATB asymmetry in fronted homophonous wh-phrases, which forces lower-copy pronunciation to avoid a PF-violation (cf. Bošković 2002):

(64) What conditions what without influencing PG$_i$. [PGs]

[25] Pseudo-coordination is relevant here, where *and* patterns with infinitival-*to*:

(c) What did Alan try and buy t$_i$?

If coordination is adjunction, it becomes difficult to distinguish post-coordination and regular coordination; cf. de Vos (2005), Weisser (2015), Zhang (2010) for discussion of possible coordination structures.

(65) *What conditions t_i without influencing what.

(66) What conditions t_i and influences what. (ATB)

(67) *What conditions what and influences t_i.

Here, PGs (64-65) prefer pronouncing the middle copy, the opposite of ATB (66-67), which requires pronouncing the lowest copy. Thus the proposed unification is complex and remains open.

The exact details of AC and PGs are not conclusive; this section intends to show their potential for unification, and given that MS-locality provides a generally satisfactory account of CSC-2 and ATB, AC and PGs are likely to also receive a satisfactory analysis with MS-locality.

4.4 Other Concerns

Agree and IM have different locality constraints. Bošković (2007a) argues Agree is not subject to the PIC_2. This follows in MS-locality, as MS for Probe-Goal and MS for Merge are different cycles, another advantage of the proposal. Consider (68), an example from Chukchee (Inénlikéj & Nedjalkov 1973, in Bošković 2007a; English translation by Mel'čuk 1988).

(68) ənan qəlγiļu ləŋərkə-nin-et [iŋqun ∅-rətəmŋəv-nen-at qora-t]
he regrets-3-PL that 3.SG-lost-3.PL reindeer-PL

'He regrets that he lost that reindeer.'

Here, the matrix v agrees with the embedded clause object, an Agree relation violating (i.e. crossing) the PIC_2. This agreement reaches into a finite CP and is evidence of Agree being unconstrained by the PIC_2. Chomsky (2008) notes redundancy between the PIC_2 and intervention effects (PCG); Bošković suggests this is another argument for Agree not being constrained by the PIC_2. Essentially, the delineation between Agree-based intervention and IM-based impenetrability is predicted by MS-locality.

MS-locality is agnostic toward implications for spell-out. Uriagereka's (1999) multiple spell-out entails a phasal view of the derivation, but the reverse is not true: the PIC_2 does not entail cyclic spell-out. As above, cyclic spell-out is problematic (Boeckx & Grohmann 2007). The agnosticism regarding spell-out is a further advantage.

Finally, Chomsky (2015a: 89) suggests representational principles such as Binding Theory (BT) ought to be completely reduced to MS. Chomsky argues if they can be reduced to MS, then they are F3 (rather than UG) principles (e.g. MS-locality reduces UG-based RM to F3-based PCG), meaning the core of language is derivational with representational principles arising from F3s. This suggests BT, like Agree, can be reduced to MS. Given Principles A and B are local,[26] this is a desirable conclusion, providing a direction for future work.[27]

In this chapter, I derived the PIC_2 and AL from MS for Merge. This approach covers the CXPC, LBC, CSC and AC (plus ATB and PG exceptions). The discussion has shown the current proposal survives empirical testing, suggesting it is an important step forward.

5 Conclusion

In this paper, I have proposed a third-factor account of locality effects. This account goes a long way toward unifying intervention and impenetrability effects. We have seen a feature-based MS, a language-specific reflex of MC, naturally derives intervention effects via an RM-like principle of PCG; MS also derives the PIC_2, and provides a basis for AL, thus accounting for impenetrability (i.e. island) effects. Further, MS-locality resolves redundancy between intervention and impenetrability principles, deriving them from the same source. Locality is derived from the derivation, i.e. narrow syntax, without stipulation. The three elements of UG, Merge, Agree and LA, all involve the operation of MS, giving a parsimonious view of the derivation based on cycles of MS. MS-locality is agnostic towards multiple spell-out, a further theoretical advantage. In sum, this paper provides a conceptually appealing unification of locality with profound consequences for syntactic theory, offering promising avenues for future work.

References

Abels, K. 2003. *Successive cyclicity, anti-locality and adposition stranding*: University of Connecticut dissertation.

Adger, D. 2017. A memory architecture for merge. Unpublished ms. Queen Mary University of London.

Bayer, S. 1996. The coordination of unlike categories. *Language* 72. 579–616.

Belletti, A. & L. Rizzi. 1981. The syntax of "ne": some theoretical implications. *The Linguistic Review* 1(2). 117–154.

Berwick, R. & N. Chomsky. 2016. *Why only us: Language and evolution*. Cambridge, MA: MIT Press.

Boeckx, C. & K. Grohmann. 2007. Remark: Putting phases in perspective. *Syntax* 10(2). 204–222.

Bošković, Ž. 1994. D-structure, theta-criterion, and movement into theta-positions. *Linguistic Analysis* 24(3-4). 247–286.

Bošković, Ž. 1997. *The Syntax of Nonfinite Complementation: An Economy Approach*, vol. 43. Cambridge, MA: MIT Press.

Bošković, Ž. 2002. On multiple wh-fronting. *Linguistic Inquiry* 33(3). 351–383.

[26] See footnote 9 for discussion of Principle C.

[27] Reuland (2011) derives binding conditions from independent properties of the grammar, offering an appropriate starting point for an MS approach to BT.

Bošković, Ž. 2005. On the locality of left branch extraction and the structure of NP. *Studia Linguistica* 59. 1–45.

Bošković, Ž. 2007a. Agree, Phases and Intervention Effects. *Linguistic Analysis* 33. 54–96.

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

Bošković, Ž. 2011. Last resort with Move and Agree in derivations and representaions. In C. Boeckx (ed.), *The Oxford Handbook of Linguistic Minimalism*, 327–353. Oxford: Oxford University Press.

Bošković, Ž. 2013. Phases beyond clauses. In L. Schürks, A. Giannakidou, U. Etxebarria & P. Kosta (eds.), *Nominal Constructions in Slavic and Beyond*, 75–128. Berlin: Mouton de Gruyter.

Bošković, Ž. 2014. Now I'm a phase, now I'm not a phase: On the variability of phases with extraction and ellipsis. *Linguistic Inquiry* 45(1). 27–89.

Bošković, Ž. 2015. From the complex NP constraint to everything: On deep extractions across categories. *The Linguistic Review (GLOW issue)* 32. 603–669.

Bošković, Ž. 2016a. Getting really edgy: On the edge of the edge. *Linguistic Inquiry* 47. 1–33.

Bošković, Ž. 2016b. On the timing of labelling; Deducing comp-trace effects, the subject condition, the adjunct condition and tucking in from labelling. *The Linguistic Review: Special Issue on Labels* 33. 17–66.

Bošković, Ž. 2018. On the Coordinate Structure Constraint, Across-the-Board-Mpvement, Phases and Labeling. Unpublished ms. University of Connecticut.

Bošković, Ž. to appear. On Unifying the Coordinate Structure Constraint and the Adjunct Condition. In A. Bárány, T. Biberauer, J. Douglas & S. Vikner (eds.), *Rethinking Grammar (Festschrift for Ian Roberts)*, Berlin: Language Sciences Press.

Bošković, Ž. & S. Franks. 2002. Across-the-Board Movement and LF. *Syntax* 3. 107–128.

Brody, M. 2002. On the status of representations and derivations. In S. Epstein & T. Seely (eds.), *Derivation and explanation in the Minimalist Program*, 19–41. Oxford: Blackwell.

Cassimatis, N., P. Bello & P. Langley. 2008. Ability, breadth, and parsimony in computational models of higher-order cognition. *Cognitive Science* 32(8). 1304–1322.

Chomsky, N. 1957. *Syntactic Structure*. The Hague: Mouton.

Chomsky, N. 1973. Conditions on Transformations. In S. Anderson & P. Kiparsky (eds.), *A Festschrift for Morris Halle*, 232–286. New York: Holt, Rinehart and Winston.

Chomsky, N. 1977. On wh-movement. In P. Culicover, T. Wasow & A. Akmajian (eds.), *Formal Syntax*, 71–132. New York: Academic Press.

Chomsky, N. 1986. *Barriers*, vol. 13. Cambridge, MA: MIT Press.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. 2000. Minimalist inquiries: The framework. In R. Martin, D. Michaels & J. Uriagereka (eds.), *Step by step; Essays on minimalist syntax in honor of Howard*

Lasnik, 89–155. Cambridge, MA: MIT Press.

Chomsky, N. 2001. Derivation by phase. In M. Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.

Chomsky, N. 2005a. Problems and Prospects of the Minimalist Program. Lectures, MIT, 25-27 July 2005.

Chomsky, N. 2005b. Three factors in language design. *Linguistic Inquiry* 36(1). 1–22.

Chomsky, N. 2008. On phases. In R. Freidin, C. Otero & M. Zubizarretta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–166. Cambridge, MA: MIT Press.

Chomsky, N. 2011. *Language and other cognitive systems. What is special about language?*, vol. 7.

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

Chomsky, N. 2015a. A discussion with Naoki Fukui and Mihoko Zushi. *Sophia Linguistica: Working Papers in Linguistics* 64. 69–97.

Chomsky, N. 2015b. Problems of projection: Extensions. In E. Di Domenico, C. Hamann & S. Matteini (eds.), *Structures, strategies and beyond: Studies in honour of Adriana Belletti*, vol. 223, 3–16. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Citko, B. 2005. On the nature of merge: External merge, internal merge, and parallel merge. *Linguistic Inquiry* 36(4). 475–496.

Collins, C. & E. Groat. 2018. Distinguishing copies and repetitions. Unpublished ms. http://ling.auf.net/lingbuzz/003809.

Cook, V. & M. Newson. 2007. *Chomsky's Universal Grammar*. Oxford: Blackwell.

Despić, M. 2011. *Syntax in the absence of the determiner phrase*: University of Connecticut dissertation.

Dingeman, A. 1978. *Thought and choice in chess*. Berlin: de Gruyter.

Fujita, K. 2017. On the parallel evolution of syntax and lexicon: A Merge-only view. *Journal of Neurolinguistics* 43. 178–192.

Gallego, Á. 2010. *Phase theory*, vol. 152. Amsterdam: John Benjamins Publishing.

Gazdar, G., G. Pullum, I. Sag & T. Wasow. 1982. Coordination and transformational grammar. *Linguistic Inquiry* 13. 663–676.

Grohmann, K. 2003. *Prolific domains: On the antilocality of movement dependencies*. Amsterdam: John Benjamins Publishing.

Higginbotham, J. 1985. On semantics. *Linguistic Inquiry* 16. 547–593.

Hiraiwa, K. 2005. *Dimensions of symmetry in syntax*: MIT dissertation.

Horrocks, G. & M. Stavrou. 1987. Bounding theory and Greek syntax: Evidence for wh-movement in NP. *Journal of Linguistics* 23. 79–108.

Huang, C.-T. 1982. *Logical relations in Chinese and the theory of grammar*: MIT dissertation.

Inénlikéj, P. & V. Nedjalkov. 1973. Glagolv čuvstva v čukotskom jazyke. In S. Kacnelson (ed.), *Lingvističeskie issledonanija 1972*, 175–203. Moscow: Institut jazykoznanija.

Kayne, R. 1994. *The antisymmetry of syntax*, vol. 25. Cambridge, MA: MIT Press.

Ke, H. 2019. *The Syntax, Semantics and Processing of Agreement and Binding Grammatical Illusions*: University of Michigan dissertation.

Kim, S. 2006. *Intervention effects in questions*: University of Frankfurt dissertation.

Ko, H. 2018. Scrambling in Korean Syntax. In *Oxford Research Encyclopedia of Linguistics*, Oxford: Oxford University Press.

Konar, A. 2000. *Artificial intelligence and soft computing: Behavioral and cognitive modelling of the human brain*. Boca Raton: CRC Press.

Korf, R. 1985. Depth-first iterative-deepening: An optimal admissable tree search. *Artificial Intelligence* 27(1). 97–109.

Landau, I. 2009. *The locative syntax of experiencers*. Cambridge, MA: MIT Press.

Larson, B. 2015. Minimal search as a restriction on merge. *Lingua* 156. 57–69.

Mel'čuk, I. 1988. *Dependency Syntax: Theory and Practice*. Albany: SUNY Press.

Müller, G. 2010. On deriving CED effects from the PIC. *Linguistic Inquiry* 41(1). 35–82.

Nunes, J. 2004. *Linearisation of chains and sideward movement*. Cambridge, MA: MIT Press.

Nunes, J. 2014. Adjunct control and edge features. In P. Kosta, S. Franks, T. Radeva-Bork & L. Schürcks (eds.), *Minimalism and beyond: Radicalizing the interfaces*, 79–108. Amsterdam: John Benjamins Publishing.

Nunes, J. 2016. On the locus and licensing of edge features. Unpublished ms. University of São Paulo. http://www.researchgate.net/publication/310156524_On_the_locus_and_licensing_of_edge_features.

Oda, H. 2018. Decomposing and deducing the Coordinate Structure Constraint. Unpublished ms. University of Connecticut.

Ortega-Santos, I. 2011. On Relativized Minimality, memory and cue-based parsing. *Iberia* 3(1). 35–64.

Ouali, H. 2010. Computation efficiency and feature inheritance in crash proof syntax. In M. Putnam (ed.), *Exploring crash-proof grammars*, 15–30. Amsterdam: John Benjamins Publishing.

Polinsky, M. & E. Potsdam. 2001. Long-Distance Agreement and Topic in Tsez. *Natural Language & Linguistic Theory* 19. 583–646.

Postal, P. 1998. *Three investigations of extraction*. Cambridge, MA: MIT Press.

Reuland, E. 2011. *Anaphora and language design*, vol. 62. Cambridge, MA: MIT Press.

Richards, M. 2007. On feature inheritance: An argument from the Phase Impenetrability Condition. *Linguistic Inquiry* 38(3). 563–572.

Rieman, J. 1994. *Learning strategies and exploratory behavior of interactive computer users*: University of Colorado dissertation.

Rieman, J., R. Young & A. Howes. 1996. A dual-space model of iteratively deepening exploratory learning. *International Journal of Human-Computer Studies* 44(6). 743–775.

van Riemsdijk, H. 1997. Push chains and drag chains: Complex predicate split in Dutch. In S. Tonoike (ed.), *Scrambling*, 7–33. Tokyo: Kurosio Publishers.

Rizzi, L. 1990. *Relativized Minimality*, vol. 16. Cambridge, MA: MIT Press.

Rizzi, L. 2001. Relativized Minimality effects. In M. Baltin & C. Collins (eds.), *Handbook of Contemporary Syntactic Theory*, 89–110. Oxford: Blackwell.

Rizzi, L. 2004. On the cartography of syntactic structures. In L. Rizzi (ed.), *The structure of the CP and IP: The cartography of syntactic structures*, vol. 2, 3–15. Oxford: Oxford University Press.

Rizzi, L. 2009. Movement and concepts of locality. In M. Piattelli-Palmarini, J. Uriagereka & P. Salaburu (eds.), *Of Minds and Language: The Basque Country encounter with Noam Chomsky*, 155–168. New York: Oxford University Press.

Rizzi, L. 2011. Minimality. In C. Boeckx (ed.), *The Oxford Handbook of Linguistic Minimalism*, 220–238. Oxford: Oxford University Press.

Rizzi, L. 2013. Locality. *Lingua* 130. 169–186.

Roberts, I. 2019. *Parameter Hierarchies and Universal Grammar*. Oxford: Oxford University Press.

Ross, J. 1967. *Constraints on Variables in Syntax*: MIT dissertation.

Russell, S. & P. Norvig. 2010. *Artificial Intelligence: A Modern Approach*. Upper Saddle River, NJ: Pearson Education 3rd edn.

Stabler, E. 2011. Computational perspectives on minimalism. In C. Boeckx (ed.), *The Oxford Handbook of Linguistic Minimalism*, 617–643. Oxford: Oxford University Press.

Starke, M. 2001. *Move dissolves into merge*: University of Geneva dissertation.

Stjepanović, S. 2014. Left branch extraction and the Coordinate Structure Constraint. In J. Iyer & L. Kusmer (eds.), *Proceedings of the North East Linguistics Society 44*, 157–170. Amherst: GLSA.

Talić, A. 2015. Adverb extraction, specificity, and structural parallelism. *The Canadian Journal of Linguistics* 60. 417–454.

Ticio, E. 2005. Locality and anti-locality in Spanish DPs. *Syntax* 8(3). 229–286.

Torr, J. 2012. Phasing out phases and re-relativizing relativized minimality. *Cambridge Occasional Papers in Linguistics* 6. 95–129.

Truswell, R. 2011. *Events, phrases, and questions*. Oxford: Oxford University Press.

Uriagereka, J. 1999. Multiple spell-out. In S. Epstein & N. Hornstein (eds.), *Working minimalism*, 251–282. Cambridge, MA: MIT Press.

de Vos, M. 2005. *The syntax of pseudo-coordination in English and Afrikaans*: Leiden University dissertation.

Weisser, P. 2015. *Derived coordination: A minimalist perspective on clause chains, converbs and asymmetric coordination*, vol. 561. Berlin: De Gruyter.

Williams, E. 1978. Across-the-board rule application. *Linguistic Inquiry* 9. 31–43.

Zhang, N. 2010. *Coordination in syntax*, vol. 123. Cambridge: Cambridge University Press.

Zoerner, E. 1995. *Coordination: The syntax of &p*: University of California, Irvine dissertation.

Seth Aycock
University of Cambridge
sjsa2@cam.ac.uk