

Not Just Islands: Relativized Blocking Throughout the Grammar

Kenneth Hanson 

Department of Linguistics, Yale University, New Haven, CT 06511, USA; mail@kennethhanson.net

Abstract

Recent research in computational linguistics has found that most long-distance linguistic dependencies are TIER-BASED STRICTLY 2-LOCAL (TSL-2) over their respective representations, providing a precise characterization of blocking elements in phonology and syntax. A subset of the well-known syntactic island effects have been claimed to be one such class of blockers, explaining their superficial diversity via a common computational basis. This article revisits the TSL-2 analysis of islands in light of subsequent work on other types of syntactic locality constraints and on the treatment of gradient blocking. It finds that the TSL-2 islands display variation and gradience of the sort found in other syntactic blockers, while those islands which are not TSL-2 can be distinguished in this regard. Several directions are suggested for future research on islands and other syntactic locality phenomena.

Keywords: locality; islands; horizons; relativized minimality; gradience; grammatical variation; Minimalist Grammars; computational complexity; tier-based strictly local languages

1. Introduction

Long-distance dependencies in language from syntactic islands to vowel harmony are subject to blocking effects, which have a peculiar recurring form: the presence of a single intervening element of the right type is enough to trigger the effect. This applies whether the dependency in question is categorical or gradient, phonological or syntactic. But why should this be so?

Recent work on the subregular complexity of natural language has provided an answer. Long-distance dependencies across modules of grammar, including many instances of phonological assimilation and dissimilation (McMullin & Hansson, 2016), syntactic movement (Graf, 2022b), agreement (Hanson, 2025c), and case assignment (Hanson, 2025b), predominantly fall within a class of formal patterns known as TIER-BASED STRICTLY 2-LOCAL, or TSL-2. Such patterns are defined by a class of visible elements, which constitute a TIER, together with a set of constraints on elements which are adjacent when relativized to the tier. From this perspective, a blocking element is simply one which is visible to the computations underlying a given dependency but which is also inert to said dependency. Even though it is inert, the blocker disrupts tier adjacency between dependent elements, causing an otherwise grammatical structure to become ungrammatical, or vice versa.

Many of the classic island constraints (Ross, 1967) can be characterized in this way, as shown by Graf (2022b). Therefore, if we are to understand islands, we must understand how they are like (and how they are unlike) other types of blocking elements, and the TSL-2 model provides the tools needed to do this. In addition to being computationally simple (i.e., restrictive), the class TSL-2 has multiple well-understood mathematical characterizations (Heinz et al., 2011; Lambert, 2023; Lambert & Rogers, 2020). As such, although it was originally defined for string sets in order to model long-distance phonotactics, it has since



Academic Editors: Thomas Graf and Aniello De Santo

Received: 16 February 2026

Revised: 6 June 2026

Accepted: 7 August 2026

Published: 29 September 2026

Copyright: © 2026 by the author.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

been extended to tree sets (Graf, 2018, et seq.), weighted and stochastic string/tree sets (Mayer, 2021, 2025; Torres et al., 2023), and string/tree transductions (Burness et al., 2021; Graf, 2023). By capitalizing on these results, we can potentially understand (i) why islands exist at all, (ii) why they are so diverse and variable, and (iii) how they fit into the theory of linguistic locality more broadly.

This article is organized into two parts. The first part surveys recent work on the TSL-2 model of syntactic locality and related computational issues. Section 2 introduces the concept of a relativized blocker, with several examples of locality phenomena which have the same abstract form, including HORIZONS (Keine, 2019) and RELATIVIZED MINIMALITY (Rizzi, 1990). Next, Section 3 reviews the TSL-2 analysis of syntactic islands and the formal system in which it is situated, focusing on the TREE TIER analysis of Graf (2022b), and provides a demonstration of how gradience can be incorporated via probabilistic tier projection (Torres et al., 2023). It also introduces an alternative model based on TIERS OVER PATHS through the structure (Graf & Shafiei, 2019; Hanson, 2024, 2025c; Shafiei & Graf, 2020), which is then used in Section 4 in order to provide a precise comparison between islands and other syntactic locality constraints.

The second part of the article extends the TSL-2 analysis of islands and suggests additional directions for future work. Section 5 examines several island types which appear to defy the TSL-2 model and argues that there is a principled divide between those islands which look like other syntactic locality conditions and those which do not. Returning to the TSL-2 islands, Section 6 sketches how several known complications involving variation across items and constructions fit into the analysis. Finally, Section 7 revisits locality phenomena in the domains of case and agreement, exploring how optionality which arises in the presence of a blocking element might be understood as the analog of ungrammaticality caused by blockers for movement. Section 8 concludes.

In order to keep this article accessible to a wide audience, technical concepts are introduced incrementally and with the minimum amount of detail necessary to tackle the goals outlined above. As such, it may also serve as a general introduction to the literature on the subregular complexity of syntax.

2. The Pervasiveness of Relativized Locality

The computational complexity of a class of patterns is a measure of the kind of abstract computational abilities needed to recognize or produce structures that match a arbitrary pattern in that class. Different classes of patterns are organized into a nested hierarchy, with more complex (i.e., expressive) classes subsuming the simpler (more restrictive) classes. Recent research has discovered that among these, just a few very simple classes, known collectively as SUBREGULAR, subsume most types of linguistic patterns (Graf, 2022a; Heinz, 2018). Of these, the subregular class which best characterizes syntactic movement and many other long-distance dependencies is TSL-2. In this article, I focus exclusively on locality effects, but it should be kept in mind that the TSL-2 generalization also sheds light on many specific aspects of grammatical variation, which are discussed extensively in many of the works cited, but particularly Graf (2022b) and Hanson (2025b, 2025c).

Since this topic may be unfamiliar to many readers, I begin with an example of relativized blocking from phonology, which is simpler to understand while illustrating several key points (Section 2.1). I then informally sketch how the TSL-2 characterization of a blocking element extends from phonology to syntax. In addition to islands (Section 2.2), I discuss two other types of syntactic locality constraints: HORIZONS (Section 2.3) and RELATIVIZED MINIMALITY (Section 2.4). The strong parallel that emerges motivates a unified computational explanation (Section 2.5). The presentation in this section is deliberately non-

technical; the formal syntactic models underpinning the generalization will be introduced in Sections 3 and 4.

2.1. TSL-2 String Patterns

In the computational complexity literature, phonological patterns are conventionally described in terms of `STRING LANGUAGES`, or `STRING SETS`. The subregular class of string languages which encompasses most long-distance dependencies is `TIER-BASED STRICTLY LOCAL`, or `TSL` (Heinz et al., 2011; Lambert, 2023; Lambert & Rogers, 2020).¹ A `TSL` computation selectively ignores items according to their label, producing a `TIER PROJECTION`; items which appear on the tier projection are subject to local constraints within some fixed window. Empirically, the window size is never larger than two elements, hence we say that long-distance linguistic dependencies are `TSL-2` (Heinz, 2018; McMullin & Hansson, 2016).

To illustrate, some languages have a rule of sibilant harmony, such that sibilants must match in anteriority even if they are extremely distant. Samala (Chumash family) is one such language. Following Heinz (2018), the tier for sibilant harmony contains all sibilant consonants, which for this language are just ‘s’ and ‘ʃ’, and the local constraints on the tier ban the substrings `*sʃ` and `*ʃs`.²

(1) Unbounded harmony (Samala)

	Word	Tier Projection
✓	pisotonosikiwat	ʃʃ
✗	pisotonofikiwat	sʃ

Importantly, it is possible for other items to appear on the tier; if present, such items could prevent the constraints from applying, allowing some words which appear to violate the harmony rule. This is precisely what happens in certain other languages with sibilant harmony, such as Slovenian (McMullin & Hansson, 2016). Here, ‘t’ is also visible, so when it intervenes, the mismatched sibilants are not adjacent on the tier, so no constraints are violated.

(2) Long-distance harmony with blocking (Slovenian)

	Word	Tier Projection
✓	ʃpiʃ	ss
✗	sʃpiʃ	sʃ
✓	nasitiʃ	stʃ

In this way, a `TSL-2` string language encodes the visibility conditions for long-distance dependencies together with constraints which are maximally local in character. This specific combination of computational ingredients has several important consequences. First, it places a firm bound on expressivity, allowing the types of long-distance dependencies which are characteristic of natural language while excluding many strange patterns (see Section 2.5). Second, it allows us to account for the relevant points of variation between languages, such as the fact that ‘t’ blocks harmony in Slovenian, but not in Samala. Third, it predicts that the tier for a certain dependency may contain extra elements which in a sense “should not be there”. This is particularly important in syntax, where blocking elements can be highly idiosyncratic.

Before continuing, I should briefly clarify some important technical points. While above I stated the visibility conditions in terms of atomic elements, we can just as easily pick out entire classes of elements based on their subatomic features; this will be the norm going forward, though the possibility for exceptional behavior of specific elements remains relevant. Additionally, the tier projection itself is best through of as a visual metaphor

which need not be literally constructed, as TSL computations can be implemented in many ways (cf. Lambert & Rogers, 2020). Another metaphor which linguists should find intuitive is a moving window of attention (i.e., a memory buffer) which remembers only the most recent elements which are members of the tier; this corresponds roughly to the finite-state machine implementation.

2.2. Islands as Blockers on a Tier

Now, we extend the ideas just introduced to syntax, focusing on the basic intuition for now. In anticipation of the MG-based formal system to be developed in Section 3, we will consider syntactic movement as a dependency between MOVERS and their LANDING SITES. We take *wh*-movement as representative, so our movers are syntactic elements such as *who/what/which/etc.* which head a moving phrase, and our landing sites are just those complementizers which attract a *wh*-mover.

Like phonological harmony, this dependency is UNBOUNDED in that there is no maximum distance as measured by the number of nodes crossed. It may potentially cross one or more clause boundaries and is in fact unbounded within a single clause, as the large number of potentially intervening functional categories (Voice, Aspect, Tense, etc.) and adjunction sites means that there is no principled finite bound even for clause-internal movement. As such, for present purposes we may abstract away from the possibility of successive cyclic movement, which in any case is not conventionally represented in MG analyses, nor in most non-transformational frameworks.³

(3) Unboundedness of *wh*-movement

- a. **What** did she see ____ ?
- b. **What** did you say [(that) she saw ____]?
- c. **What** did you think [(that) I said [(that) she saw ____]]?

Also like phonological harmony, even though movement is unbounded, this does not mean that it is unrestricted. In particular, there are several types of elements which prevent dependency formation when they intervene between a mover and landing site. Many islands fall into this category, as shown by Graf (2022b). For example, the *wh*-island constraint says that *wh*-movement may not cross the boundary of an interrogative embedded clause, even one in which the interrogative element is plausibly base-generated.

(4) The *wh*-island for *wh*-movement

- a. * **What** did you wonder [if she saw ____]?
- b. * **What** did you wonder [whether she saw ____]?

Thus, even though *wh*-movement can cross a long distance, just a single element of the right type is enough to block it, as in the phonological example. This is the principal reason why such blocking effects are TSL-2, not TSL-*k* for some larger *k*, nor some other more complex class. Note that while the complementizer which heads the embedded clause is the most likely candidate for the blocker here, we might also assign blame to the verb in some cases; I discuss this possibility in Section 6.

Leaving aside the technical details, we would like to say that the tier specifying the constraints on *wh*-movement in English—henceforth, the *wh*-tier—includes the interrogative complementizers *if* and *whether* but not declarative *that*. Naturally, it must also include all *wh*-movers and *wh*-landing sites, since otherwise the constraints cannot be stated. The relevant constraint on the tier is that a landing site must be adjacent to a suitable mover and vice versa, where tier adjacency is defined by relativizing an appropriate structural relation (dominance, c-command, etc.). This holds in the case of movement across a declarative clause boundary (3b), but is broken in the case of movement out of an interrogative clause (4b). The latter is therefore ruled out. This is schematized below.

(5) *Wh*-island as tier blocker

	Sentence	Tier Projection
✓	$C_{wh} \dots$ [that $\dots$ <i>what</i>]	C_{wh} <i>what</i>
✗	$C_{wh} \dots$ [<i>whether</i> $\dots$ <i>what</i>]	C_{wh} <i>whether what</i>

Here, the structural relation that matters is DOMINANCE, which means that the relevant intervener is actually the maximal projection (interrogative CP). This may not be obvious when thinking in terms of the phrase structure tree representation, but it is highly intuitive in the context of the formal model introduced in Section 3. It can also be motivated by the fact that certain island structures can intervene via c-command without inducing the corresponding island effect; for example, a VP or TP adjunct does not block *wh*-movement of the complement or something contained within. That said, there are blocking effects which clearly require some sort of command relation; these will be discussed shortly.

The above characterizes not only the *wh*-island, but several others as well, to be discussed in Section 3. It also applies to other locality constraints which are even more pervasive in syntax, far beyond just the A'-movement types which are the locus of island effects. Let us examine these now.

2.3. Horizons

It is well-known that different syntactic dependencies are subject to different locality restrictions; conversely, the same types of restrictions may apply to a wide range of dependencies, such as movement, agreement, and case. When we look at these restrictions through a formal lens, we discover that they have the same abstract form as an island.

For example, consider English-style movement of the subject to Spec-TP, which I will refer to as EPP-movement.⁴ Like *wh*-movement, EPP-movement can cross an unbounded distance, even assuming successive cyclic movement. However, it is blocked by any clause boundary which includes a CP layer.

- (6) a. **This theory** ____ has some problems.
 b. **This theory** seems [_{TP} to ____ have some problems].
 c. **This theory** seems [_{TP} to be likely [_{TP} to ____ have some problems]].
 d. * **This theory** seems [_{CP} (*that*) ____ has some problems].

A similar effect can be observed with verbal agreement in existential clauses, where existential *there* in the clausal subject position may be skipped in favor of the logical subject, even when the latter is embedded in a non-finite TP, though not a CP. (We assume that the expletive subject is invisible for agreement, if it is represented at all.)

- (7) a. There **are** problems (with this theory).
 b. There **seem** [_{TP} to be some problems (with this theory)].
 c. There **seem** [_{TP} to be likely [_{TP} to be some problems]].
 d. * It **seem** likely [_{CP} *for* there to be some problems].

Keine (2019, 2020) calls such blocking categories HORIZONS. In the case of English, C is a horizon for A-dependencies such as EPP-movement and verbal agreement, but not for A'-dependencies such as *wh*-movement. This is not the only possible analysis of such facts, but it is simple, uniform, and extends to many other data points which are less amenable to other explanations. Crucial, this restriction or some analog must be stated even in a system which assumes intermediate movement to Spec-CP, as in many incarnations of transformational-generative theory, for otherwise we have no explanation for why A'-movement may escape CP while A-movement cannot.⁵ Furthermore, some

languages do seem to allow A-raising past a finite clause boundary, a phenomenon known as hyperraising; we would say that C is not a horizon for A-raising in such languages.

Horizons are conceptually different from islands as they include all members of a given category in the functional spine (V, T, C, etc.); additionally, they apply to both agreement and movement dependencies, whereas islands are usually thought to be specific to A'-movement. However, they are formally similar in that they define the kinds of phrases which a given class of movers can escape, or out of which morphological information can be transmitted, and they do so from the perspective of dominance/containment (again, this will become more clear in the formal implementation). It is also easy to see how they fit into the TSL-2 model of syntactic locality: as described in Hanson (2024, 2025c), all members of the relevant categories appear on the tier regulating the movement or agreement dependency in question. We revisit horizons in more detail in Section 4.

2.4. Relativized Minimality

The final class of syntactic locality effects I will discuss fall under the principle of RELATIVIZED MINIMALITY (Rizzi, 1990, 2001, 2013).⁶ The principle, stated informally, holds that a syntactic operation (movement, agreement, etc.) involving two elements X and Y cannot hold if there is another element Z which intervenes and which could also participate in said operation, where intervention is usually defined in terms of c-command. This encompasses, among other things, the classic superiority constraint: when forming a *wh*-question with multiple *wh*-phrases, only the structurally highest one can move to Spec-CP.

- (8) a. **Who** did you say [___ saw what]?
- b. * **What** did you say [**who** saw ___]?

Relativized minimality is also the reason why verbal agreement tends to target the structural subject in preference to the object or some other nominal. This is understood under the assumption that the T(ense) node is the locus of agreement in English and many other languages, and the subject is closer to T than the object is.

- (9) a. The cat_{SG} **chases**_{SG} the rats.
- b. * **The cat** **chase**_{PL} the rats_{PL}.

As shown in Hanson (2025c), relativized minimality follows from the TSL-2 model under the assumption that both actual *and* potential participants for a given dependency are always visible on the relevant tier, combined with an appropriate formalization of the c-command relation. The model also accommodates well-attested complications such as case-sensitivity (e.g., where agreement targets the closest *nominative* nominal, as in Icelandic or Hindi), which impede a strict interpretation of relativized minimality. I cover command-based intervention and the basic minimality constraint in Section 4.

2.5. The Power of the TSL-2 Hypothesis

We have previewed several types of syntactic locality constraints which have the same abstract form as those in phonology, of which islands are just one. Once we factor out differences in the geometric relation (string precedence, dominance, or command) and the substance of the constraint (phonetic content, clause type, functional category, feature similarity, etc.), the form of the constraint is the same: the dependent elements must be adjacent on their tier.

This is not a trivial result. As discussed by McMullin and Hansson (2016), blocking effects in long-distance phonological processes are always triggered in the presence of *one or more* blockers, whereas other hypothetical blocking effects—e.g., assimilation that occurs only when there is *at most one* blocker or *exactly one* blocker—are unattested. This is

explained by the assumption that such dependencies are restricted to the class TSL-2, as other types of locality conditions cannot be expressed.

The same seems to be true of syntax. The most robustly attested locality conditions are triggered whenever there is at least one blocker between the dependent elements, as emphasized by the works cited above (Graf, 2022b; Hanson, 2025c). Indeed, it is very easy to imagine other kinds of blocking rules which are computationally simple yet linguistically preposterous. Consider for example, “XP is an island iff there are no other XPs in the same clause” or “XP is an island only for movers that contain a PP”, which are just two of the non-existent island types pointed out by Graf. There are many strange linguistic phenomena, but patterns like these seem not to be among them. If syntactic and phonological dependencies are united by virtue of being TSL-2 over their respective structures, then the parallel is explained.

In contrast, while there have been claims in the literature that certain syntactic dependencies can cross one blocker but not two, notably the theory of subjacency (Chomsky, 1973, 1977), there is usually an alternative explanation available; indeed, as we will soon see, all of the island types which admit a subjacency-based account also have a simple TSL-2 analysis. This would be an extraordinary coincidence if subjacency (which requires at least TSL-3) was the true explanation. Furthermore, given that the TSL-2 hypothesis accounts for other types of locality conditions, including those in phonology, and many other aspects of syntax besides (Graf, 2022b; Hanson, 2025b, 2025c), we would like to push this single explanatory factor as far as it can go.

3. The TSL-2 Analysis of Syntactic Islands

Having established the intuition for the TSL-2 analysis of islands, we now formalize it in order to provide a unified account for many of the classic island constraints. Along the way, we will touch on several past studies of islands in the subregular syntax literature (Graf, 2022b; Hanson, 2024; Shafiei & Graf, 2020; Torres et al., 2023), all of which explore variations of the same basic idea—that islands represent an interruption to adjacency as defined by RELATIVIZED DOMINANCE. This sets the stage for the subsequent discussion of other syntactic locality conditions (Section 4) and the handful of islands which are not TSL-2 (Section 5).

This section begins with an overview of the Minimalist Grammar (MG) dependency tree representation used in the studies cited (Section 3.1), then explains how constraints on movement are implemented using TREE TIERS (Section 3.2). Next, it provides a survey of the classic syntactic islands which can be understood as TSL-2 locality constraints (Section 3.3). From there, it extends the analysis by showing how gradience can be added to the model (Section 3.4), and then considers alternative approaches which treat islands in terms of PATHS though the tree defined by the dominance relation, which are called ANCESTOR STRINGS (Section 3.5). The section concludes with a short discussion of some formal issues which come up along the way (Section 3.6).

3.1. MG Dependency Trees

To begin, I briefly describe the Minimalist Grammar (MG) dependency tree representation used in Graf (2022b) and other recent work in subregular syntax. While another syntactic representation could in principle be used, the MG view of movement lends itself well to this kind of analysis, and the use of dependency trees eliminates some mathematical warts that obscure the key observations. For brevity, I omit some technical details which do not matter to the focus of this article. Also see Graf (2026) for a modern introduction to MGs for syntacticians and Graf and Kostyszyn (2021) for a first-order axiomatization of MG dependency trees as used here.

MINIMALIST GRAMMARS (MGs, [Stabler 1997, 2011](#)) are a mildly context-sensitive syntactic formalism which assemble sentences from their atoms using a small number of primitive operations. In the standard version, there are just two operations: Merge, which combines elements into phrases, and Move, which displaces an existing phrase to the root of the structure. The atoms are “lexical items” which pair phonological (and semantic) information with a string of FEATURE DIACRITICS which describe the operations they undergo.⁷ For Merge, we have selector features (F^+) and category features (F^-), whose names are straightforward. For example, a transitive verb bears the feature D^+ , and a determiner bears the corresponding D^- . For Move we likewise have two types of features: licensor features (f^+) for landing sites of movement and licensee features (f^-) for the movers themselves. This means that the complementizer which attracts a *wh*-mover bears wh^+ , and the head of the moving phrase (who/what/which/etc.) bears wh^- .⁸ Importantly, only elements which actually move in the present derivation bear a licensee feature. Thus, in the sentence *Who saw what?* only *who* bears the wh^- feature.

MG DEPENDENCY TREES represent the sequence of Merge and Move steps in an MG derivation; that is, they are a kind of DERIVATION TREE.⁹ In comparison, a conventional phrase structure tree represents the output of a syntactic derivation (up to some point), and can therefore be called the DERIVED tree. For illustration, both the dependency tree and phrase structure tree for (3b) *What did you say that she saw?* are depicted in Figure 1. For simplicity, movement of the subject from within VP to Spec-TP not represented. Additionally, I omit selector/category features to reduce clutter.

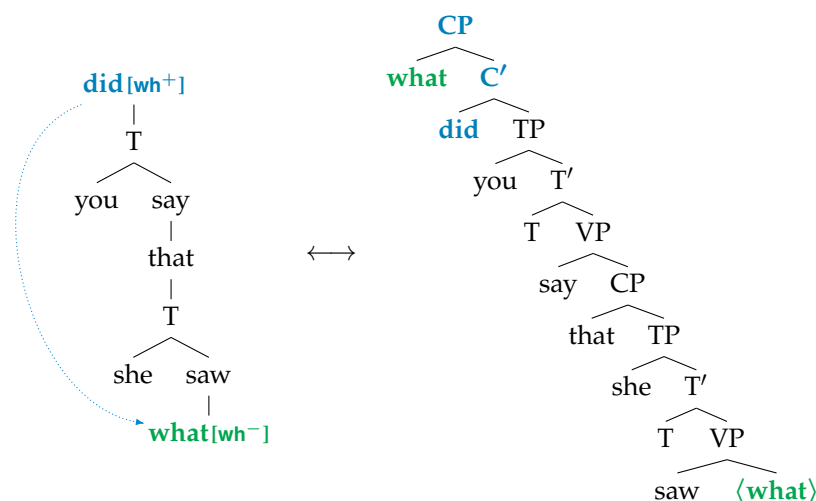


Figure 1. Dependency tree (left) and phrase structure tree (right) for *What did you say that she saw?* In the dependency tree, the projections each head are collapsed, movement is undone, and MG features indicate movement operations. The landing site and mover for the *wh*-movement dependency are highlighted and linked by an arrow for visual emphasis.

For present purposes, the reader needs to understand three properties of the dependency tree representation. First, every node represents a lexical item in its base position, where it is first introduced by Merge. Second, the mother-of relation represents selection of the daughter by the mother; when a single head selects multiple arguments, they are ordered in reverse order of selection, such that structurally superior co-arguments appear to the left. Thus, the left-sister relation corresponds to local c-command (see Section 4.2). Also, every subtree represents the phrase headed by the node at its root. Third, long-distance dependencies such as movement are indicated only via the feature diacritics on the participating heads. Dotted lines linking mover and landing site are provided for visual emphasis, and are not part of the representation.

A few technical notes follow. First, as mentioned earlier, successive cyclic movement of the *wh*-mover to embedded CP is not conventionally represented in MG analyses. Such movement is not required for the TSL-2 analysis and so I omit it. This equates the view that such movement is not feature driven; (Graf, 2023) shows how copies or traces can be inserted in the relevant locations in the mapping from the dependency tree to the corresponding phrase structure tree, if desired. Additionally, even though the features of a lexical item are normally ordered in MGs, we assume here that all licensee features are unordered with respect to each other. This ensures that the dependency tree language is TSL without affecting the weak generative capacity of the system; see Graf and Kostyszyn (2021) for details.

3.2. Tree Tiers and the *wh*-Island

We now have a compact representation which indicates the basic syntactic structure created by local selection and the positions of movers and landing sites within that structure. This is enough to formulate various constraints on the locality of movement, starting with the classic islands.

In the TSL-2 model of movement, originating in Graf (2018) and further developed in Graf (2022b), it is assumed that each movement operation pairs the mover with the closest appropriate landing site. This kind of relativized locality is represented in the form of a TREE TIER PROJECTION (OR TREE TIER) containing all movers and landing sites of the relevant type, as shown in Figure 2. Notice that the landing site immediately dominates its corresponding mover on the tier. Assuming that movement obeys the extension condition (no sideways movement, etc.), this will always be the case. This is why it is most natural to model islands in terms of RELATIVIZED DOMINANCE. As discussed in the previous section, it is also possible for other types of elements to appear on the tier regulating a particular long-distance dependency. As such, the island structure in (4b) *What did you wonder whether she saw?*, illustrated on the right of Figure 2, is just one in which adjacency between mover and landing site is broken on the relevant tier.

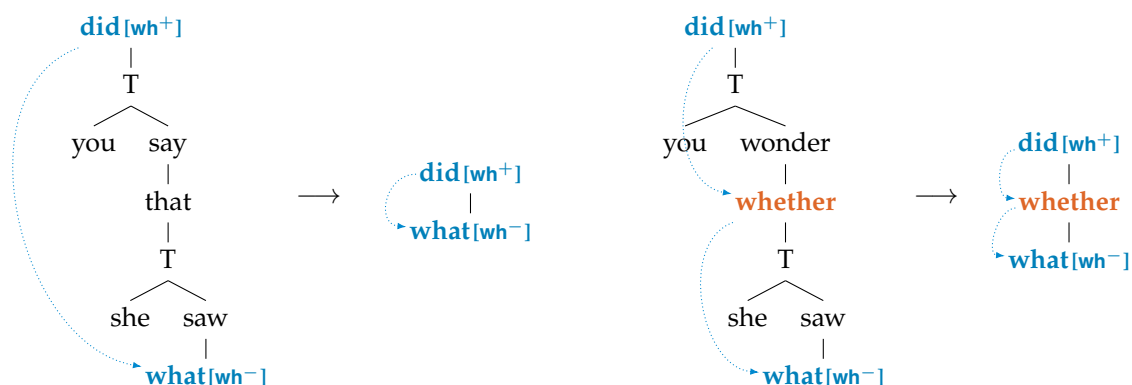


Figure 2. (Left) Dependency tree and *wh*-tier for cross-clausal *wh*-movement in *What did you say that she saw?* The mover is local to its landing site on the tier, as required. (Right) Illicit movement due to *wh*-island structure in **What did you wonder whether she saw?* The island node interrupts adjacency on the tier, making the structure ill-formed.

Let us make this constraint more precise. Following Graf (2022b), we define a tree tier which contains all elements bearing either the *wh*⁺ or *wh*⁻ diacritics, plus any blocking elements, in this case the complementizers *if* and *whether*. On this tier, every item bearing *wh*⁺ must have exactly one marked by *wh*⁻ among its daughters; conversely, every item with *wh*⁻ must have one with *wh*⁺ as its mother. This constraint is satisfied in our running example of long-distance *wh*-movement, as illustrated on the left side of Figure 2, but rules

out illicit structures such as those with a landing site but no mover, or multiple movers sharing the same landing site, among many others. Crucially, it is also violated in the *wh*-island structure on the right side of Figure 2: not only does the wh^+ landing lack a wh^- mover among its daughters, but the mover also has something other than a landing site as its mother.

This brings us to an important technical point. It is possible to have many visible nodes which are all dominated by one node on the same tier, none of which dominate the other. This means that there is no limit to the number of daughters of a node on a tree tier, even if there is such a limit in the original tree language. Because of this, informal constraints such as “a node marked wh^+ has exactly one node marked wh^- among its daughters” are encoded in the DAUGHTER STRING LANGUAGE of each node that appears in the tier, defined only by its label. If each daughter string matches the language specified by the mother, then the tier is licit, and so is the original tree, modulo independent constraints. Crucially, each daughter string language may be at most TSL-2, just like the tree tier itself. This is true of the constraints as just stated.¹⁰

Graf (2022b) makes use of both the vertical and horizontal dimensions of the tree tier model in order to analyze several other well-known movement phenomena, including multiple *wh*-movement and extraction morphology. Here, I focus on the relativized dominance relations represented by the tree tier, which is where island effects arise. Later, we will remove the tier sisters entirely and focus on the paths created by the dominance relation. This approach then generalizes to intervention by specifiers, which is defined in terms of command rather than dominance, as explained in Section 4.

3.3. Other TSL-2 Island Constraints

To briefly review, we have a mathematically rigorous model of long-distance dependencies in syntax where islands arise when some node which is neither a mover or landing site is nonetheless visible on the tree tier regulating a particular movement type. For *wh*-islands, these nodes are just the interrogative complementizers. Now, let us consider some other islands with the same form, which include the adjunct island, relative clause island, complex NP island, and perhaps also the subject island. Below are examples illustrating these for *wh*-movement, replicated from Graf (2022b):

- (10) a. NON-ISLAND STRUCTURE
 What did John complain [that Mary brought ____ to the party]?
 b. WH-ISLAND
 * **What** did John wonder [whether Mary brought ____ to the party]?
 c. ADJUNCT ISLAND
 * **What** did John complain [because Mary brought ____ to the party]?
 d. RELATIVE CLAUSE ISLAND
 * **What** did John complain about the person [that brought ____ to the party]?
 e. COMPLEX NP ISLAND
 * **What** did John complain about [the fact that Mary brought ____ to the party]?
 f. SUBJECT ISLAND
 * [**Which student**] did [the advisor of ____] study island constraints?

Once again, in each of the illicit cases it is fairly intuitive that there is a particular phrase containing the mover which interferes with movement to a position outside, which for us means that there is a particular head which intervenes in the dependency tree. Which head is relevant depends somewhat on one’s analytical assumptions, so I will provide just one or two for illustration.

First, the `ADJUNCT ISLAND` constraint is straightforward; we just need to be able to identify adjunct clauses. Assuming adjunction as a primitive operation as in [Frey and Gärtner \(2002\)](#), the head of the adjunct clause bears an adjunction feature such as $V^\approx$ or $T^\approx$. Or, we could posit adjunctizer heads which wrap the adjunct before combining with the adjunction site, as suggested by [Graf \(2022b\)](#); in this case the adjunctizer head is the blocker.¹¹

Similarly, `RELATIVE CLAUSES` can be identified as C heads which adjoin to NPs using either analysis of the adjunct island. Alternatively, we can identify them as bearing the licenser feature rel^+ for A' operator movement in theories that assume this; this might be preferable since it would also cover free relatives.

Next, the `COMPLEX NP ISLAND` would appear at first glance to defy this approach, as neither NPs or declarative CPs on their own block movement, but in fact the relevant heads are just those nouns which select a CP complement, that is, those with the features $C^+ N^-$.

By the same token, Graf claims that `SUBJECT ISLANDS` can be identified as subjects are just those nominals which bear a licensee feature indicating movement to Spec-TP, which I notate epp^- . Note that this implementation of the subject island does not cover unmoved subjects, like the logical subject of an existential clause. Graf suggests that if such subjects undergo covert movement then they also bear a feature. If this is not tenable then the subject island might instead be treated with those islands falling under the purview of the model discussed in Section 5. I retain it for illustration of the case where a licensee feature identifies the island.

There are yet other islands not discussed by [Graf \(2022b\)](#) which also appear amenable to a TSL-2 analysis. These include the second part of [Ross's \(1967\)](#) `FROZEN STRUCTURE CONSTRAINT`, which bans movement out of an extraposed clause, as in the example below. Such clauses bear a licensee feature ext^- , making their analysis directly analogous to the above analysis of the subject island.

- (11) Frozen structure constraint ([Ross, 1967](#))
- a. [A girl $___i$] came in [who had worn this coat] $_i$.

b. *The coat [which $_j$ [a girl $___i$] came in [who had worn $___j$] $_i$] was torn.

Note that this analysis does not extend to the more general class of “freezing effects” (cf. [Wexler & Culicover, 1980](#)), which state that once a certain XP has moved, a YP contained within XP may not be extracted. (The opposite situation, in which YP moves out of XP, then XP moves to yet a higher position, is known as remnant movement, and is generally thought to be possible.) Such a constraint is not TSL-2 in the MG dependency tree model as the mere fact that one mover is contained within in another is not enough to know which moved first. I set freezing effects aside as they are not as clear cut either empirically or theoretically as classic islands, but see [Shafiei and Graf \(2020\)](#) for a subregular analysis.¹²

All of the islands we have discussed so far are summarized in Table 1.

Table 1. Various island constraints and their blocking items.

Island Type	Tier Blockers
<i>wh</i> -island	Interrogative C heads
Adjunct island	Items bearing any adjunction feature $X^\approx$
Relative clause island	Items bearing licenser feature rel^+
Complex NP island	Heads bearing features $C^+ N^-$
Subject island	Heads bearing licensee feature epp^-
Frozen structure	Heads bearing licensee feature ext^-

It is notable that in spite of the apparent disunity, all admit one or more very simple TSL-2 analyses. Indeed, as emphasized by [Graf \(2022b\)](#), that islands should exist at all

is a prediction of the model, as there is no *a priori* reason why the tier projection cannot contain such miscellaneous elements. To this I would add that it is especially interesting that almost every element of the simple MG-based model assumed here appears to have a corresponding island type, including arbitrary lexical items as well as entire classes of items which share certain MG features. The model also explains why such islands have the same form as other locality constraints in language, including those discussed in Section 2, while other types of locality constraints are so unusual (recall Graf's non-existent islands defined by arbitrary logical connectives).

To be clear, why these particular groups of items happen to be islands and others not requires a separate explanation. What is crucial is that whatever the origin of the constraint might be, the form is always the same. One way of interpreting such facts—espoused by Rogers and Pullum (2011) as well as Graf (2022a)—is that the relevant cognitive system is only sensitive to certain types of structural relations, TSL-2 in the case of long-distance dependencies. In this case, we might posit that only external factors which can be correlated with such structural relations have the opportunity to be encoded as syntactic constraints. And we do need to encode at least some of them since, as discussed in Section 6, there is also a degree of arbitrary grammatical variation.

In summary, the TSL-2 model describes most island types as simple locality conditions of the same sort found elsewhere in the grammar. That they should be so varied is unsurprising, since linguistic blocking elements are in general quite heterogeneous. At the same time, some of the classic island constraints cannot be described in this manner. In particular, the left branch condition and the coordinate structure constraint do not seem very amenable to a TSL-2 analysis, but might be better served by a different sort of computational explanation. I defer discussion of these until Section 5. For the remainder of this section, I address some formal questions related to the TSL-2 analysis of islands.

3.4. Incorporating Gradience

The relation between gradient acceptability and the syntactic grammar is a contentious issue. I do not take a strong stance on the matter here, however, there are two important considerations that I wish to point out. First, if one wishes to incorporate gradience into the grammar formalism—whether for modeling experimental data, training a predictive parser, or otherwise—then it is very straightforward to do so. We simply replace yes/no decisions with multiple discrete values or real numbers, such that every word or sentence is assigned a weight or probability; classic examples include weighted finite-state automata and probabilistic context-free grammars. Second, TSL-2 provides a very natural place to do this to model gradient island effects, which is the tier projection function. No change to the core structural analysis is required.

I bring this up in part because one occasionally encounters the claim that formal theories cannot account for gradience. For example, Chaves and Putnam (2022) criticize the TSL-2 model of islands for being intrinsically categorical, and therefore incompatible with the empirical facts. On the contrary, gradient acceptability makes eminent sense from this perspective, as do some of the other empirical complications that they cite. While I cannot address all of their criticisms in this article, we tackle gradience now, and other types of variation are covered in Section 6.

The specific model I describe here USES PROBABILISTIC TIER PROJECTION, first proposed by Mayer (2021) and generalized to trees by Torres et al. (2023). This is not the only way of incorporating gradience into a TSL-2 grammar (cf. Mayer, 2025), but it is fairly intuitive. Each node label is associated with a projection probability between 0 and 1. Those with a fractional value may or may not appear on the corresponding tier, such that each tree may have multiple tier projections corresponding to a given long-distance dependency. The tier

constraints determine which of these are well formed, as in the categorical version. Finally, a score between 0 and 1 is assigned to the original structure by adding up the probabilities of all licit tier projections.

An example derivation is shown in Figure 3 for our recurring *wh*-island sentence, *What do you wonder whether she saw?*, which has fairly low acceptability. In order to keep things simple, we assign a projection probability of 1 to all landing sites and movers, which by hypothesis are always visible, a probability of 0.7 to the island node *whether*, indicating that it is a moderately strong blocker, and a probability of 0 to all others, which are presumably irrelevant. There are now two *wh*-tier projections: one with the illicit island structure, and the other well formed. The licit structure has a projection probability of 0.3, so this is the score assigned to the derivation, as outlined by the calculation below.

$$\begin{aligned} \text{Score}(\text{tree}) &= p(\text{tier1}) * \text{licit}(\text{tier1}) + p(\text{tier2}) * \text{licit}(\text{tier2}) \\ &= 0.7 * 0 + 0.3 * 1 \\ &= 0.3 \end{aligned}$$

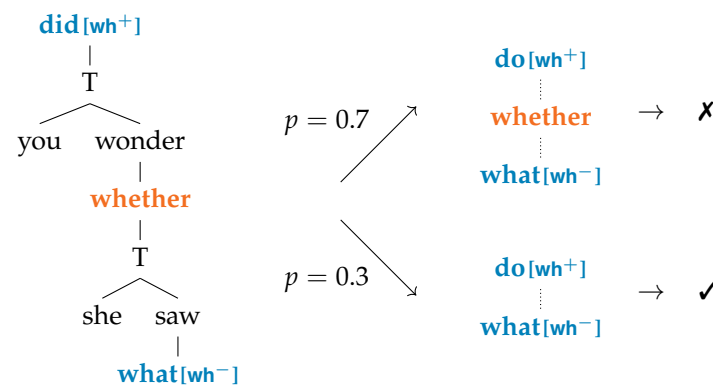


Figure 3. Probabilistic tier projection with a *wh*-island. The lexical item *whether* is assigned a projection probability of 0.7, leading to two distinct *wh*-tier projections, only one of which is licit.

To reiterate what I said earlier regarding categorical tier projection, this does not imply that speakers literally construct and evaluate many projections for the *wh*-tier, and likewise for every other tier in the syntactic grammar. Once again, probabilistic tier projection is a mathematical abstraction which might be implemented in any number of ways. Furthermore, there are multiple potential interpretations for the resulting weights. Here, we focus on just one, which is as a correlate of gradient acceptability. In Section 7, I introduce a different interpretation which results in optionality rather than gradience.

Torres et al. (2023) fit the projection probabilities of several island nodes (*whether*, adjuncts, and complex NPs) plus the finite declarative complementizer to a set of MG dependency trees paired with normalized mean Likert score ratings derived from Sprouse et al. (2016). The resulting projection probabilities correlate with the relative badness of the different island violations under consideration while replicating the so-called superadditivity effect, which is that extraction from an island is worse than the combined effects of extraction from a declarative embedded clause and the mere presence of an island elsewhere in the sentence.

This study has a number of limitations. For example, there are not many data points, so the projection probabilities of most nodes are fixed to 0 or 1 (as in the above example) in order to avoid overtaxing the model fitting procedure. In addition, the projection probability of each island node is shared across movement types, but they may be independent according to the definition of the formalism, and as discussed in Section 6, there do seem to be empirical differences.¹³ In light of these considerations, I suggest

a follow-up study which probes this issue directly. We would construct several models with different parameter spaces: the first with shared projection probabilities for island nodes across movement types, as in Torres et al. (2023), and the second with fully separate probabilities, which is more faithful to the TSL-2 model as a whole. The latter model should provide a better fit, especially if appropriately discriminating sentence types are tested.

There are many other possible directions besides this one, and it may even turn out that probabilistic tier projection as currently defined does not make quite the right cut. Perhaps there should be structured relations between tiers, which equate to an interaction between parameters in the statistical model, or we might decide that gradience in the tier constraints (i.e., the daughter string languages) is also crucial.¹⁴ What I hope is clear by now is that the present computational approach is not only compatible with studies of gradient acceptability, it naturally suggests several productive directions for future research.

3.5. A Path-Based Alternative

Above, we saw that islands block adjacency in the dimension of dominance. This would seem to suggest that we can do away with tier sisters and just utilize the PATHS defined by relativized dominance. The resulting string constraints would look nearly identical to those from phonology: we simply ban any configuration in which a mover and its landing site are not adjacent in the tier projection. Since path-based approaches to long-distance syntactic dependencies have extensive precedent (Bouma et al., 2001; Kaplan & Zaenen, 1989; Kayne, 1984, a.o.), it should be little surprise that they have been explored in subregular syntax as well (Graf & Shafiei, 2019; Hanson, 2024, 2025c; Shafiei & Graf, 2020). Let us therefore try to reformulate the island constraints in this manner.

To begin, consider the dominance path from a given node to the root of the tree, which Shafiei and Graf (2020) call an ANCESTOR STRING, or A-STRING. For illustration, the same structures from Section 3.2 are repeated in Figure 4 with the a-strings of the movers superimposed. The a-strings themselves are written out below along with their *wh*-tiers. Once arranged this way, string precedence in the a-string represents reversed dominance.

- (12) A-strings for cross-clausal movement and *wh*-island
- What did you say that she saw?
A-string: what[wh⁻] · saw · T · that · say · T · did[wh⁺]
wh-tier: what[wh⁻] · did[wh⁺]
 - *What did you wonder whether she saw?
A-string: what[wh⁻] · saw · T · whether · wonder · T · did[wh⁺]
wh-tier: what[wh⁻] · whether · did[wh⁺]

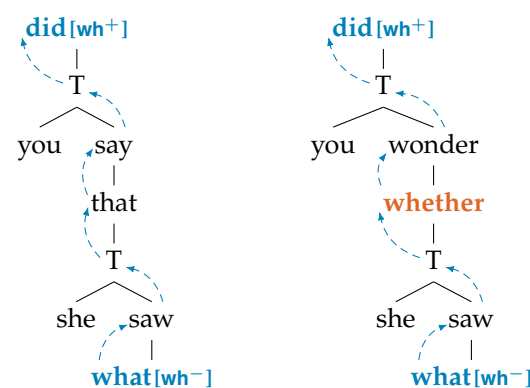


Figure 4. A-string of a *wh*-mover for cross-clausal movement (left) and island structure (right). Examining the a-string of a mover allows locality violations to be detected.

In an a-string grammar, every a-string must be well formed for the tree to be well formed, similar to how in a tree tier grammar every daughter string must be well formed. Suppose that our TSL-2 a-string constraints require every element bearing wh^- to be immediately followed by a wh^+ and vice versa—a maximally literal reinterpretation of the analysis in Graf (2022b). At first glance, this appears to work, since the mover and landing site are adjacent in the licit example, but not in the island structure. In fact the latter is doubly ruled out as the tier substrings $what[wh^-] \cdot wonder$ and $wonder \cdot did[wh^+]$ are both illicit. Such redundancies are usually harmless, and every tier substring must be classified as either licit or illicit in any case.

However, there is a catch. If we take an arbitrary a-string and examine its wh -tier, we might find that a constraint is violated even though the full tree is clearly licit. For example, the a-string of *she* in the licit structure in Figure 4 is $she \cdot T \cdot that \cdot say \cdot T \cdot did[wh^+]$, which contains a lone landing site. Similarly, we can construct examples that appear to contain an island violation from the perspective of an arbitrary a-string. These include sentences like *Who did you ask whether she saw anything?*, illustrated in Figure 5. The a-string for any node in the embedded clause will contain the forbidden substring $whether \cdot did[wh^+]$, but this is irrelevant because the island node does not dominate the mover. This shows that we cannot translate the tree tier analysis to a-strings without modification.

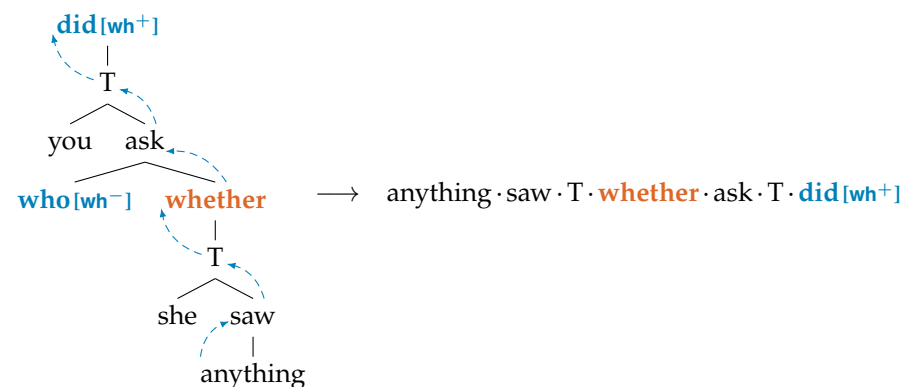


Figure 5. An a-string of a non-moving node may contain a landing site without a mover. The highlighted a-string appears to contain an island violation, but in fact the structure is fine. This is an important limitation of a-strings.

This situation is a direct consequence of the asymmetry of the dominance relation, which branches in the downward (mother to daughter) direction. Every path-based analysis of movement must deal with this fact, TSL-2 or otherwise. There are a variety of potential solutions, of which I will present just one, based on Hanson (2024). (See below for a note on some alternatives.) The essential idea is that of two parts of Graf’s constraint—every landing side must tier-dominate its mover, and every mover must be tier-dominated by its landing site—the second is the actual locus of island effects. Since including the first part in our a-string grammar is not merely redundant, but harmful, the solution is to omit it.

The final analysis is minimally different from the above sketch. As before, the wh -tier contains all landing sites and movers for wh -movement plus any blockers, including island nodes as discussed in Section 3.3. The constraints, on the other hand, simply mandate that no wh -tier may contain a mover followed by anything *other than* a landing site. This is summarized in (13). Note that $\times$ marks the end of the a-string, which covers the case where there is no landing site at all.

(13) A-string grammar for *wh*-movementTier: all $D_{[wh^-]}$, all $C_{[wh^+]}$, all island nodes

$$\text{Constraints: } \left\{ \begin{array}{l} *D_{[wh^-]} \cdot \times \\ *D_{[wh^-]} \cdot \text{whether} \\ *D_{[wh^-]} \cdot \text{because}[\text{T}^\approx] \\ *D_{[wh^-]} \cdot C_{[\text{rel}^+]} \\ *D_{[wh^-]} \cdot \text{fact}[C^+ N^-] \\ *D_{[wh^-]} \cdot D_{[\text{epp}^-]} \\ *D_{[wh^-]} \cdot C_{[\text{ext}^-]} \\ \dots \end{array} \right\}$$

This a-string grammar can be seen as a simplified version of the tree tier analysis of movement. Indeed, tree tier grammars are strictly more powerful than a-string grammars, as the former include all information in the latter, plus constraints involving the tier sister relation. While we cannot handle all syntactic constraints with a-strings, dominance-based locality can be expressed in a simple and transparent manner. Better yet, the approach generalizes easily to paths which follow a command relation, as shown in Section 4.2.

I now briefly describe two additional path-based alternatives to the present analysis. First, the original a-string-based analysis of islands presented in Shafiei and Graf (2020) uses a more expressive subregular class known as IBSP (cf. Graf, 2017).¹⁵ Conceptually, we first demarcate a potential movement path within each a-string (e.g., from a wh^- to a wh^+), then ask whether any forbidden substructures occur within this restricted domain; if no movement path is found, then there is no need to check for locality violations. This represents a somewhat different way of thinking about islands, which is also similar to that found in Heuser et al. (2024), discussed in passing in Section 6. Note that it is still not possible to detect a landing site without a mover; this requires tree tiers or some comparable extension.

The other alternative, based on the model in Graf and Hanson (2025), would combine elements of both of the above analyses. The core idea is that a node can select its a-string language, just as a node on a tree tier determines its daughter string language. In this case, we can apply a TSL-2 grammar only to those a-strings which begin with a mover of the relevant type. This solves the problem of filtering out partial movement paths without increasing the expressivity of the a-string language, though tree tiers still cannot be replaced completely. A proper exploration of this idea is left to future work.

To summarize, since islands require only relativized dominance, we can step down from tree tiers to a-strings, which provide just enough structure to facilitate a TSL-2 analysis. Under this system, all island constraints take the form of a banned substring consisting of a mover followed by anything other than a landing site. While other path-based analyses are also possible, the basic a-string model captures the essence of intervention via containment. As we will soon see, horizons take exactly the same form; the same is true of relativized minimality, modulo the difference between dominance and command.

3.6. Summary and Discussion

Many but not all of the classic islands can be understood as simple blockers on a tier, interrupting relativized dominance between a mover and its landing site. We operationalized this idea in two ways: tree tiers and tiers over a-strings. The latter are a weaker formal model, but they are sufficient to express locality constraints such as these. We also saw how a categorical TSL-2 tree grammar can be converted into a weighted grammar, modeling gradient acceptability of island violations via probabilistic tier projection. The same idea might be further generalized to create a probabilistic grammar, or indeed, applied to a-strings.

Although our focus is on the computations of syntax, these are inherently entangled with other formal issues such as the choice of grammatical formalism, the choice of representation, and the inherent properties of the tree geometry. While I cannot do justice to any of these here, I wish to at least address them briefly.

First, as mentioned at the beginning of this section, subregular analysis need not be limited to MGs, but can also be applied to other grammatical formalisms. In particular, the derivation tree languages for other mildly context-sensitive formalisms such as TAG (Tree Adjoining Grammar) and CCG (Combinatory Categorical Grammar) would be a natural starting place, though to my knowledge nothing like the present analysis has been attempted. Additionally, different syntactic formalisms may suggest other types of computational explanations for locality restrictions; for example, see Frank (2002, 2004, 2006) regarding islands in TAG.¹⁶

Second, as far as I can tell, the basic insights of the TSL-2 analysis of islands in the MG dependency tree model translate to classical derivation trees (and also to phrase structure trees), except that input structure-sensitive tier projection is required, marking a small but notable increase in computational complexity. That is to say, the analysis would not be TSL-2, but ITSL-2 (De Santo & Graf, 2019). Graf (2018) presents an interesting argument regarding the uniform complexity of selection and movement in that system. Conversely, the typological predictions of the model become weaker due to increased expressivity, particularly for selection and adjunction, as discussed in Hanson (2025a).

Last, regarding the tree geometry, I noted earlier that the asymmetric branching of the dominance relation prevents a complete implementation of Graf's movement rule in the a-string model. This turned out not to be a problem since islands can be described solely from the perspective of the mover; as we will soon see, the same is true of horizons and minimality. Not all syntactic constraints can be described in this manner, however. The tier sister relation which is included in the tree tier model is one solution to the problem; the complement spine restriction in Hanson (2025c), also discussed below, represents another. At the same time, the asymmetry of the tree structure yields an interesting automata-theoretic explanation for the coordinate structure constraint (Graf & De Santo, 2019), which is described in Section 5.

All considered, there are many interesting formal issues which interact with the TSL-2 hypothesis for syntax, more than enough to merit a dedicated study. For the remainder of this article, we will focus on the empirical side as we extend the analysis a wider array of locality phenomena. That said, there is one formal problem in particular we must still tackle, which is the definition of a suitable command relation to complement dominance.

4. Other Syntactic Locality Constraints

In the previous section, we examined the TSL-2 model of the classic island constraints in considerable detail. In this section, we do the same, albeit more briefly, for the other syntactic constraints introduced earlier: horizons and relativized minimality. As we will soon see, the form of each constraint is identical, though relativized minimality requires the introduction of a new structural relation. In the MG dependency model, the relevant relation is known as DERIVATIONAL COMMAND, or d-command, and corresponds roughly to asymmetric c-command in the phrase structure tree.

Path-based models (Graf & Shafiei, 2019; Hanson, 2024, 2025c) are employed throughout this section. For horizons, I use the a-string model introduced previously (Section 4.1). For minimality, I introduce COMMAND STRINGS, which is defined in terms of d-command instead of dominance (Section 4.2). Since we already cast islands in terms of a-strings, this allows us to make the comparison maximally explicit. Finally, for completeness, I also

cover the variant of the command string model introduced in Hanson (2025c) for modeling syntactic agreement (Section 4.3).

4.1. Horizons with Ancestor Strings

Horizons as TSL-2 constraints are discussed in Hanson (2024, 2025c). The former work is chiefly concerned with learning which elements are blockers, and uses a-strings, while the latter focuses on parametric variation in agreement dependencies, couched in a variant of the command string model. Here, I follow the presentation in Hanson (2024) since as discussed earlier, horizons really should be defined in terms of dominance.

Recall the example from Section 2 in which EPP-movement is blocked by a complementizer, though not by a tense node. The key sentences which illustrate this, first shown in (6), are repeated below.

- (14) a. **This theory** seems [TP to ____ have some problems].
 b. * **This theory** seems [CP (**that**) ____ has some problems].

As mentioned previously, this phenomenon has been given various explanations, but treating the complementizer as a horizon for EPP-movement situates it within a unified framework for understanding variation between different dependencies and across languages, as outlined by Keine (2019, 2020). Additionally, it is the form of the constraint which is of interest to us, and so I focus on EPP-movement only as an exemplar of the general template.

Following standard MG practice, finite T and the moving subject bear a feature $\text{epp}^+/\text{epp}^-$, just as interrogative C and a *wh*-mover bear wh^+/wh^- . I also assume that the subject originates within the VP (more properly, *v*P) and moves to Spec-TP even in simplex clauses. This established, the dependency trees for the above examples are shown in Figure 6 with the a-string of the mover overlain, as in the previous section. Already, the parallel with the *wh*-island example should be clear, but let us make it fully explicit. We assume that the EPP-tier contains all EPP-movers and landing sites plus any blockers, the latter including all C heads. Both a-strings and their EPP-tiers are given below.¹⁷

- (15) A-strings for raising and hyperraising
- a. This theory seems to have some problems.
 A-string: $\text{this}[\text{epp}^-] \cdot \text{have} \cdot \text{to} \cdot \text{seems} \cdot \text{T}[\text{epp}^+]$
 EPP-tier: $\text{this}[\text{epp}^-] \cdot \text{T}[\text{epp}^+]$
- b. * This theory seems that has some problems.
 A-string: $\text{this}[\text{epp}^-] \cdot \text{have} \cdot \text{T} \cdot \text{that} \cdot \text{seems} \cdot \text{T}[\text{epp}^+]$
 EPP-tier: $\text{this}[\text{epp}^-] \cdot \text{that} \cdot \text{T}[\text{epp}^+]$

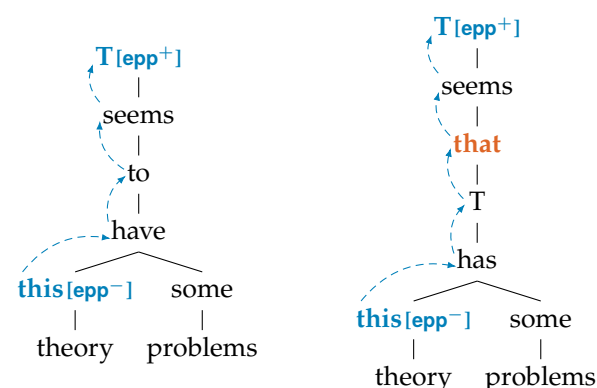


Figure 6. A-string for EPP-mover in licit raising structure (left) and illicit hyperraising structure (right). The complementizer is analyzed as a horizon.

As before, the relevant constraint is that the mover must be immediately followed on the tier by its landing site. The resulting a-string grammar is shown below. As with the grammar for *wh*-movement, I leave an ellipsis in the constraint list to emphasize that there could in principle be other constraints.

(16) A-string grammar for EPP-movement

Tier: all D_[epp⁻], all T_[epp⁺], all C nodes

$$\text{Constraints: } \left\{ \begin{array}{l} *D_{[epp^-]} \cdot \times \\ *D_{[epp^-]} \cdot \text{that} \\ *D_{[epp^-]} \cdot \text{for} \\ \dots \end{array} \right\}$$

Crucially, this assumes that the embedded T node does not bear the *epp⁺* diacritic. One alternative analysis, in which finite T is the relevant blocker, says that it must bear the critic and is therefore guaranteed to be on the tier, in contrast to non-finite T, which need not bear the diacritic. Another example which does not readily admit such an explanation is blocking of agreement by the complementizer first shown in (7). The crucial contrast is repeated below.

- (17) a. There **seem** _{[TP} to be some problems].
 b. *It **seem** likely _{[CP} **for** there to be some problems].

Here, we hold finiteness constant, so non-finite T cannot be the blocker. Admittedly, there is still a confound, which is that the expletive subject in (17b) is *it* rather than *there*. But notice that while there are various analyses for the locality restrictions of each dependency in isolation, horizons provide a single solution that works for both of them, and many others besides. In the end, the locality conditions in these examples are TSL-2 under a variety of analyses, which is ultimately what matters.

In the interest of brevity, I omit the full derivations for the agreement examples. A complete example of minimality for agreement is provided below.

4.2. Command Strings, Relativized Minimality

The dominance relation needs little explanation, but command is more complicated. Although it is often taken for granted that c-command (Reinhart, 1981) is the other chief geometric relation in syntax, many other versions have also been explored (cf. Barker & Pullum, 1990). Furthermore, all of these are designed around the shape of syntactic dependencies in the phrase structure tree. To achieve the same effect in the dependency tree, a new command relation is required. This is DERIVATIONAL COMMAND, or D-COMMAND.

As first described in Graf and Shafiei (2019), d-command is a partial order over the nodes of the dependency tree which encodes their structural prominence, reflecting the order of feature checking by Merge. Nodes which had their category feature checked later d-command those whose feature was checked earlier. So, a verb that takes both a specifier and complement d-commands each of them, and the specifier also d-commands the complement. This is illustrated on the left side of Figure 7. To be clear, the relation is transitive, so the verb also d-commands every node contained by its arguments.

Alternatively, we can define d-command in purely geometric terms. There are two primitive relations in the dependency tree which represent structural prominence: dominance and left-sister. The problem we need to address is the fact that the specifier of a node can enter into a long-distance dependency with a node embedded in the complement. This cannot be expressed using either relation alone, so we combine them into a new relation, defined in (18). This is visualized on the right side of Figure 7. Notice that unlike c-command, d-command is inherently asymmetric.

This brings us to the TSL-2 grammar. In present terms, only the subject can bear the wh^- feature in this structure since it is the closest possible mover. This means that the tier for wh -movement contains not only those items which actually bear the wh^- feature, but also those items which *could have* borne it, in this case other wh -heads. As shown in [Hanson \(2025c\)](#), this assumption combined with the fact that we are now using c-strings is what creates the effect of relativized minimality. For concreteness, the c-strings from the current example and their wh -tiers are presented below.

(20) C-strings for wh -movement and minimality violation

- a. Who did you say she saw?

C-string: $who_{[wh^-]} \cdot T \cdot C \cdot say \cdot you \cdot T \cdot did_{[wh^+]}$

wh -tier: $who_{[wh^-]} \cdot did_{[wh^+]}$

- b. * Who did you say saw what?

C-string: $what_{[wh^-]} \cdot saw \cdot who \cdot T \cdot C \cdot say \cdot you \cdot T \cdot did_{[wh^+]}$

wh -tier: $what_{[wh^-]} \cdot who \cdot did_{[wh^+]}$

I have taken some liberties in presentation, since [Hanson \(2025c\)](#) does not use plain c-strings; the version from that work will be explained shortly. For now, I draw attention to the fact that the form of the constraint is once again identical: the mover must not be followed by anything other than a landing site. This is shown in the grammar below.

(21) C-string grammar for wh -movement

Tier: all wh -heads with $[wh^-]$, all wh -heads w/o $[wh^-]$, all $C_{[wh^+]}$

Constraints: $\left\{ \begin{array}{l} *D_{[wh^-]} \cdot who \\ *D_{[wh^-]} \cdot what \\ *D_{[wh^-]} \cdot where \\ \dots \end{array} \right\}$

One might reasonably ask whether islands (and horizons) can be integrated into the c-string grammar so that we can dispense with a-strings altogether. This is possible to an extent, but there are subtleties involved. These are discussed in the next section.

At this point, it should be fairly clear how the analysis generalizes to the other examples in Section 2. For example, assuming that the class of potential EPP-movers includes all DPs, the minimality constraint for EPP-movement bans a mover followed by another D head on its c-string tier. However, I have not yet provided a full derivation for agreement. In anticipation of the discussion to follow, I do so now.

So far we have assumed standard MGs, which model only selection and movement. (We touched on adjunction but did not go into detail.) In order to describe the constraints on agreement, we need to add a new operation. [Hanson \(2025c\)](#) extends the notation for movement to agreement: f^+ marks the ‘probe’ (e.g., finite T), which receives the morphological features of another node, and f^- marks the ‘goal’ (usually some D head), which transmits its features (cf. [Chomsky, 2000](#)). In the most common examples of verbal agreement, some combination of ϕ -features (i.e., person/number/gender features) are transmitted together, so we lump these together with diacritics ϕ^+ / ϕ^- for simplicity.

And so, we revisit the example of minimality in verbal agreement from (9), repeated below with the relevant c-strings and their agreement tiers. As usual, the dependency trees are provided in Figure 9. (We assume that all ϕ -features are collected on the head of the DP, as is standard.) Object agreement is ruled out because the subject D head intervenes, violating the constraint $D[\phi^-] \cdot D$.

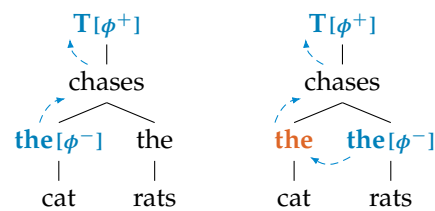


Figure 9. C-strings for licit agreement (left) and relativized minimality violation (right).

(22) C-strings for verbal agreement and minimality violation

- a. The cat **chases** the rats.

C-string: the $[\phi^-]$ · chases · T $[\phi^+]$

ϕ -tier: the $[\phi^-]$ · T $[\phi^+]$

- b. * The cat **chase** the rats.

C-string: the $[\phi^-]$ · the · chase · T $[\phi^+]$

ϕ -tier: the $[\phi^-]$ · the · T $[\phi^+]$

Once again, I must emphasize that not every D head is equally visible for agreement. In addition to existential *there* mentioned earlier, [Hanson \(2025c\)](#) also highlights case-sensitivity ([Bobaljik, 2008](#); [Preminger, 2014](#)) as a common cross-linguistic variable. Similar to the diversity of island effects, the TSL-2 model can be said to predict such variation. I cannot go into detail here regarding agreement, so I refer the reader to the works cited. Instead, I will survey some important points of variation in island effects (beyond the mere existence of multiple island types) in Section 6, which have not yet been given their due in the subregular syntax literature.

To summarize, a representative selection of syntactic blocking effects is collected in Table 2. These include all of those given a complete derivation in Sections 3.5, 4.1 and 4.2, plus a few which were omitted for brevity.

Table 2. Summary of syntactic locality constraints treated in this section.

Dependency	Constraint	Relation	Model	Tier Blocker	ex. Banned Substring
<i>wh</i> -movement	island	dominance	a-string	whether, etc.	D[wh $^-$] · whether
<i>wh</i> -movement	minimality	d-command	c-string	all <i>wh</i> -heads	D[wh $^-$] · who
EPP-movement	horizon	dominance	a-string	all C heads	D[epp $^-$] · C
EPP-movement	minimality	d-command	c-string	all D heads	D[epp $^-$] · D
ϕ -agreement	horizon	dominance	a-string	all C heads	D[ϕ^-] · C
ϕ -agreement	minimality	d-command	c-string	all D heads	D[ϕ^-] · D

4.3. The Complement Spine

For our final path-based model, I briefly describe the variant of the c-string model used in the typological study of agreement in [Hanson \(2025c\)](#), termed SPINAL C-STRINGS since they trace the main complement spine of dependency tree or of some complex left branch or adjunct therein. For illustration, a structure with three non-trivial complement spines is shown in Figure 10. Compared to ordinary c-strings, there are several small but important differences. First, they point downwards rather than upwards. Second, the number of strings we must manage is much fewer than the number of nodes, with minimal overlap between them: only the head of a complex left branch (or adjunct) is ever duplicated across spines. Third, except for the main spine, the others do not extend to the root of the tree. Essentially, what we have done is to successively chop off the branches of the tree until we are left with a set of strings.

As in the original c-string model, it is necessary for every spinal c-string associated with a given dependency tree to be well formed. One advantage of this setup is that it

how the constraint is formulated, so we focused on ancestor and command strings due to their conceptual simplicity. Thus, the same TSL-2 analysis which encompasses so many of the island constraints also subsumes horizons, relativized minimality, and perhaps other constraints not yet considered.

Similar to some of the formal issues raised in Section 3.6, the fact that we are now entertaining several different types of paths, not just a-strings, but c-strings and spinal c-strings, raises important questions regarding where the line should be drawn. I do not have a definitive answer. That said, these various formal models may not be quite as different as they appear at first glance—see [Graf and Hanson \(2025\)](#) for an automata-theoretic view that unifies several of them, including daughter strings, a-strings, and c-strings, into a coherent framework.

This concludes our survey of past work on islands and related locality conditions as TSL-2 constraints. The remainder of this article builds on this foundation, extending the analysis of islands and sketching some additional ideas for further research.

5. Some Islands Are Different

Some of the other classic island constraints, specifically the left branch condition and the coordinate structure constraint, appear not to admit a TSL-2 analysis, at least not with tree tiers. This is acknowledged by [Graf \(2022b\)](#), who suggests that they may admit some sort of path-based analysis, or that they derive from some other computational factor. Given how new this line of research is, it is difficult to draw firm conclusions at the moment. However, there are reasons to believe that the misbehaving islands should *not* be analyzed as TSL-2 locality constraints, or indeed, as locality constraints of any sort. This division also correlates with other well-established empirical findings, as the TSL-2 islands are known to be subject to extensive gradience and variation, much like other blocking effects, whereas the non-TSL-2 islands are not.

Consider the left branch condition, which says that the leftmost constituent of a larger nominal cannot be moved. This covers pairs like the following in English:

- (24) Left Branch Condition
- a. Whose book did you read ___?
 - b. * Whose did you read [___ book]?

Leaving aside the fact that such constructions are fine in other languages, this is very different from the other island constraints: descriptively, it is a ban on the extraction of a specifier (or a head) of a certain category. But tree tiers do not tell us whether a given node represents specifier or not, nor do any of the path-based models we have considered. This is not a problem with our hypothesis about locality, but with the way in which movement interacts with the tree geometry.

The issue with the coordinate structure constraint is similar, though it takes a bit of explanation to see why. It should actually be broken into two constraints: a ban on movement *of* conjuncts (the Conjunct Constraint of [Grosu 1973](#)), and a ban on movement *from* conjuncts (the Element Constraint).

- (25) Conjunct Constraint
- a. They ordered coffee and apple pie.
 - b. * What drink did they order ___ and apple pie?
 - c. * What dessert did they order coffee and ___?
- (26) Element Constraint
- a. Alan loves baseball and Betty hates golf.
 - b. * What sport does Alan love ___ and Betty hate(s) golf.

- c. * What sport does Alan love baseball and Betty hate(s) ____.

Upon initial inspection, these facts have a straightforward TSL-2 analysis: if there is a Conj(unction) element whose arguments are the individual conjuncts, then we simply project the Conj head onto the tier, just like any other island node. The problem is that the constraint is suspended in cases of across-the-board (ATB) movement, which only applies to movement from within a conjunct.

(27) Across-the-board movement obviates the Element Constraint

- a. What sport does Alan love ____ and Betty hate ____?
- b. It is hockey that Alan loves ____ and Betty hates ____.
- c. The sport that Alan loves ____ and Betty hates ____ is hockey.

The full pattern, taking into account across-the-board movement, has no possible TSL-2 analysis. If the Conj head is visible, then the landing site has no access to its movers; if it is invisible, then there is no way to restrict ATB movement to just this context. Furthermore, unlike other island constraints, there is no variation among different A' dependencies, nor any noticeable gradience—coordinate structure constraint violations are always very strong, and across-the-board movement is completely fine. It seems more likely that conjunction structures are not blockers for movement in the sense we have developed, rather, both parts of the coordinate structure constraint have some other origin.

First, the immobility of conjuncts is analogous to the left branch condition in that it requires reference to some grammatical property which is not modeled in the present system, namely that of being a conjunct. I do not espouse any particular explanation for why this should be the case. I note only that if we were to add this information to our grammar, then the constraint might be TSL-2, but it would be a restriction on the kinds of elements which can bear particular licensee features, not a locality effect in any sense.

As for the fact that movement is possible out of all conjuncts, but not just one, I am partial to the explanation given by [Graf and De Santo \(2019\)](#), who propose that the MG dependency tree must be recognizable by a SENSING TREE AUTOMATON. This basic idea is this. The landing site and the mover are linked via a deterministic top-down procedure with finite look-ahead, which means that it must be known in advance which path leads to the mover. Normally, the mover is not deeply embedded in some specifier or adjunct; only paths which recursively traverse the complement are possible, so there is no problem. However, if coordinate structures are inherently symmetric then there is no principled way to choose among the possible paths. Therefore, the only choice is to send the signal for the unfilled dependency down *all* conjuncts, such that it must be satisfied in each independently. This implies that **What sport does Alan love ____ and Betty hates golf* is ungrammatical for the same reason that **What sport does Betty hate golf* is ungrammatical: there is no gap to be filled, or in MG terms, there is no mover to be paired with the landing site.

We still need an explanation for why across-the-board movement is possible at all, since this contradicts the usual movement rule as encoded in the tree tier analysis. While details have yet to be worked out, my interpretation is that the TSL-2 constraints should interact with the restrictions of the sensing tree automaton such that the daughter string language for the landing site ends up being applied to the contents of each conjunct separately.¹⁹

Moving along, the same explanation also potentially covers the subject island, or more correctly, the SPECIFIER ISLAND, as it would ban extraction from all specifiers, not just subjects. The same goes for the adjunct island as well. (To be clear, we have not explained why the complement is the privileged path, but given that it is, the rest follows.) Recall that these restrictions are very strong for agreement, which motivated the spinal c-string model. However, when it comes to movement, the subject and adjunct islands are actually more

porous than than the coordinate structure constraint (see Section 6). Since they otherwise seem amenable to a TSL-2 explanation, I leave the categorization of these islands open.

Finally, the conjunct constraint and left branch condition could be implemented in this system since the sensing tree automaton can identify the relevant nodes, though this view provides little insight into why they should exist. This is not much of an improvement over simply adding a diacritic to the dependency tree.

Thus, from a computational perspective it appears that the constraints traditionally grouped together as islands are likely not a uniform class. Many of them are TSL-2 blocking effects, formally identical to other well-attested locality restrictions. Others are not locality restrictions at all and have some other origin, which may nonetheless be computational in nature. Our exploration of variation in the following section will further support this conclusion. The relative roles of the TSL-2 model of locality and the sensing tree automaton model in explaining each of the islands discussed here are summarized in Table 3.

Table 3. Summary of island types and their potential explanations as TSL-2 blockers or as consequences of the sensing tree automaton model.

Island Type	TSL-2 Blocking Effect?	Sensing Tree Automaton Explanation?
<i>wh</i> -island	yes	no
Relative clause island		
Frozen structure constraint		
Complex NP island		
Adjunct island	maybe	maybe
Subject island		
Left branch condition	no	no
Conjunct constraint		
Element constraint	no	yes

6. Variation in Island Effects

A variety of complications to the classic island constraints have been unearthed over the years. Building on the foundation set by Graf (2022b), I expand the analysis to show that it accounts for several dimensions of variation. As with our earlier discussion of gradience, this is important for several reasons. First, the existence of such variation can be said to be predicted by the model, since there is no reason why the constraints may not differ slightly across A' dependencies, different languages, or indeed, individual speakers. Second, it is especially interesting that the subset of the classic islands which are TSL-2 are precisely those which are subject to extensive variation, whereas those we distinguished on computational grounds in the previous section are not. This suggests that the TSL-2 generalization identifies a meaningful natural class of phenomena defined by their computational characteristics.

The following survey is far from exhaustive, but should be sufficient to demonstrate how grammatical variation works in the TSL-2 model. Specifically, I discuss how lexical and cross-linguistic variation can be handled by adjusting the tier contents (Section 6.1), and how asymmetries across movement types fit into the system (Section 6.2). In then briefly compare variation in the TSL-2 islands with the variation which has been described for other syntactic locality constraints (Section 6.3), and with the non-TSL-2 islands identified in the previous section (Section 6.4).

Along the way, I note some instances of gradient acceptability, which correlate with the existence of variation in general. For the most part, these can be relegated to extra-grammatical factors, and so for ease of exposition I will treat each contrast as if categorical. If,

however, one adopts the stance that gradient grammaticality is required, then the weighted generalization of a TSL-2 tree grammar described in Section 3.4 provides a solution that otherwise preserves the insights of the present model.

6.1. Lexical Variation

Above, we assumed that it was some property of the clause that defines most islands. In particular, declarative complement clauses are assumed to allow extraction while interrogatives forbid it; this is the *wh*-island. In fact, this is possible only for the complements of certain verbs such as *say/think/believe*, which are called BRIDGE VERBS. In contrast, complements of factive verbs such as *know* tend to resist extraction; these are sometimes called FACTIVE ISLANDS. In addition, other miscellaneous verbs that take a clausal complement are considerably less tolerant of extraction, including *whisper*, *mutter*, *complain*, etc.

(28) Bridge and non-bridge verbs (Chaves, 2022)

- a. What did John say that Mary bought ____? (bridge verb)
- b. ?? What did John know that Mary bought ____? (factive verb)
- c. ?? What did John whisper that Mary bought ____? (manner-of-speaking verb)

Chaves (2022) notes that the effect seems to be at least partly pragmatic in nature. (Another example of this type is discussed below.) On the other hand, precisely which verbs robustly serve as bridge verbs is subject to cross-linguistic and individual variation, which motivates the acquisition-based account in Heuser et al. (2024). In the latter case, it is simple to state the relevant distinction in the TSL-2 grammar: bridge verbs are absent on the tier, but other verbs are present. Given that we already assume some nouns to be visible, this is hardly a major leap.

If some but not all members of a major category may be islands, then certain other asymmetries can also be accounted for. For example, some PPs are islands, but not others (cf. Chomsky, 1973):

(29) Lexical variation in *wh*-movement from PPs

- a. Who did you see a picture of ____?
- b. What did you receive a request for ____?
- c. * Who did you admire a book by ____?
- d. * What color do you buy the car in ____?

To the extent that the specific blockers for each movement type must be learned, this appears to be technically feasible. Heuser et al. (2024) describe an algorithm which learns syntactic blockers from positive data by tracking the bigrams that occur along movement paths in a phrase-structure tree. When applied to English data, it learns that *wh*-movement of a verbal subject/object is possible, movement out of classic islands is impossible, and movement out of miscellaneous structures such as the VP-CP and NP-PP structures just described is available for only certain lexical combinations. Additionally, as noted in Hanson (2024), the information tracked by their algorithm is extremely similar to that needed by a formal TSL-2 learner, and so it can be adapted to output a TSL-2 grammar while retaining its core characteristics.²⁰

Let us consider one last example of this type. It has been observed that extraction from infinitival island structures is relatively acceptable. Between the pairs of examples below, extraction from the infinitive is considerably better than extraction from the fully tensed clause. Thus, we might posit that the infinitive C head (or T head, depending on the analysis) is missing from the tier despite being interrogative.

- (30) Infinitival and tensed island structures (Richards, 2014)
- a. ? What are you wondering [whether to buy ___]?
 - b. * What are you wondering [whether you should buy ___]?

As for the fact that extraction from the infinitive is judged less than perfect, this could be due to processing difficulties. This explanation is compelling since many instances of gradient acceptability correlate with processing difficulty; see, e.g., Hofmeister and Sag (2010), De Santo (2020), and many of the works cited by Chaves (2022). At the same time, it also raises important questions regarding the distribution of work among these various competing explanations. It may well be the case that some islands (particularly the coordinate structure constraint) are universal while others are learned, and that the latter group includes both the lexical-distributional type as well as processing effects which have been grammaticalized via some other mechanism, with pure processing effects on top of this. Attempting to disentangle these factors completely is far beyond the scope of the present article. What is important is that to the extent that islands must be encoded in the grammar, the TSL-2 model is up to the task.

6.2. Variation by Movement Type

While to a first approximation islands are said to apply to all A' dependencies, it turns out that some island constraints are weaker for certain movement types. For example, Chaves (2022) observes that several of the classic islands are ameliorated for relative clause formation, henceforth RC-movement. Some of the examples he cites include:

- (31) Acceptable examples of RC-movement out of an island
- a. ... Nemaha, a town_i that [to describe ____i as tiny] would be to overstate its size. (Subject island)
 - b. This is the watch_i that I got upset [when I lost ____i] (Adjunct island)
 - c. I met the person who_i Robin knew [that Alex helped ____i]. (Factive island)

Similarly, topicalization and cleft formation may show severely weakened island effects in some instances. These facts are not difficult to understand in the TSL-2 model. Although these other processes are related to *wh*-movement, they are not literally the same type of movement. They are therefore regulated by different tiers, and may have slightly different sets of visible items. In the case of the present example, RC-movement appears to be subject to fewer restrictions than *wh*-movement, and so there are fewer items which appear on the tier. Alternatively, the same items are present but with different projection probabilities, as discussed in Section 3.4.

Although not emphasized in early discussions of islands, it has always been assumed in the subregular literature that each long-distance syntactic dependency has its own tier with its own visibility conditions and constraints, and that these may vary across languages and speakers. For Graf (2022b), this is the reason why seemingly surprising phenomena such as multiple *wh*-movement, the *that*-trace effect, and extraction morphology can be found; they are not difficult to encode in a TSL-2 grammar, so there is no particular reason why they should not exist. Following the same logic, it would be more surprising if each one of these movement tiers was identical in structure given there is no *a priori* reason that this must be the case, much like how it would be surprising if there were no islands at all.

We might consider the effect of D[iscourse]-linking (Cinque, 1990; Pesetsky, 1987, a.o.) in similar terms. In brief, D-linked *wh*-phrases are those whose range of possible answers are pre-established in the discourse. D-linking can subvert the usual locality restrictions, including both superiority and islands. The effect is lexically conditioned, such that *which*-phrases are more likely to be interpreted as D-linked out of context, leading to improved acceptability.

- (32) D-linking and superiority (Pesetsky, 1987)
- Who did you persuade ___ to read what?
 - * What did you persuade who to read ___?
 - Which book did you persuade which man to read ___?
- (33) D-linking and *wh*-islands (Richards, 2014)
- * What do you wonder how to solve ___?
 - ? Which problem do you wonder how to solve ___?

The exact nature of the effect is debated. Szabolcsi and Zwarts (1992) argue for a semantic account; in this case it is a curious coincidence that the semantics should break down in precisely the manner predicted by the TSL-2 model. If the effect is instead pragmatic in nature, then following the intuitions of Rizzi (2001) and Starke (2001), we might posit that movement of D-linked *wh*-phrases is actually a different subtype of A' movement from non-D-linked movement, such that the former may skip over items related to the latter, including not just movers (subverting superiority/minimality) but also *wh*-islands. Without getting into the reason why the asymmetry should be this way, we can encode the difference in visibility conditions in our grammars just as before.²¹ If this approach is on the right track, then any gradience in acceptability can be explained in terms of the relative difficulty of accessing the D-linked interpretation.

Finally, I would like to address the often-cited claim that arguments are somewhat easier to extract from islands than adjuncts.

- (34) Adjunct/argument asymmetry (Richards, 2014)
- ?? What are you wondering [whether you should buy ___]?
 - * Why are you wondering [whether you should buy it ___]?

The full data is extremely complex, and various accounts have been provided, some syntactic (Huang, 1982; Rizzi, 1990, a.o), some not (Szabolcsi & Zwarts, 1992). Taken at face value, the effect as traditionally described does not seem very amenable to the present model, including via probabilistic tier projection. One possible but unsatisfying solution is to make the daughter string language a WEIGHTED LANGUAGE and penalize moving adjunct daughters of an island node more severely than arguments. Alternatively, it is conceivable that the adjunct/argument distinction is not actually the right cut, given that some cases of adjunct extraction are okay given an appropriate supporting context:

- (35) Licit cases of adjunct extraction across a *wh*-island (Szabolcsi, 2006)
- How are you wondering [whether to solve the problem ___]?
(OK when choosing from salient checklist)
 - How did [no student solve the problem ___]?
(OK when choosing from salient checklist)

In this case, we might consider unifying the phenomenon with D-linking. As suggested by Szabolcsi (2006), the reason that why/how questions are more susceptible to island violations is because of the difficulty in accessing a D-linked interpretation, not in the fact that they are adjuncts. This seems more desirable than altering the daughter string languages, since the latter are only tangentially involved in the derivation of locality effects.

6.3. Other Blockers Are Also Variable

As mentioned previously, other types of syntactic blockers are subject to similar variation. By definition, horizons are relativized to individual syntactic dependencies; in addition, while Keine (2019, 2020) attempts to pare down the space of possible horizons

such that they can be predicted by the position of the landing site (or agreeing head), there remains room for parametric variation. For example, TP is clearly invisible to verbal agreement dependencies in English (cf. *There seem* [_{TP} *to be some problems with this theory*]), but it is opaque in Hindi. Similarly, some exceptions to relativized minimality are best captured by assuming that visibility of potential targets for agreement is parameterized by other features such as case (Bobaljik, 2008; Preminger, 2014), and that these once again may vary in a semi-arbitrary manner across languages.

Such variation is to be expected under the TSL-2 hypothesis. Indeed, it is absolutely pervasive. Hanson (2025c) explores the formal variation in visibility and other aspects of agreement, including those just mentioned, and Hanson (2025b) does the same for case. This in turn directly mirrors what we find in phonology (McMullin & Hansson, 2016) and other aspects of movement (Graf, 2022b). Thus, the closer we look, the more islands appear just like other blocking elements. We revisit this parallel in Section 7, where we will examine cases of optionality which occur in the presence of certain blocking elements, reminiscent of the kinds of gradience observed with island structures.

6.4. The Non-TSL Islands Are Not So Variable

The island types identified in Section 5 as not being reducible to TSL-2 locality conditions—the left branch condition and both parts of the coordinate structure constraint—are also much less susceptible to variation or gradient acceptability than the others (Chaves, 2022). While some languages do not enforce the left branch condition, it is exceptionless in those that do, and it cannot be ameliorated by pragmatic factors, repeated exposure, or anything of the sort. The same seems to be true of the Conjunct Constraint. Finally, while the Element Constraint is suspended in cases of across-the-board movement, this does not seem to depend on the movement type, and in any case has an independent explanation based on the symmetry of coordination structures.²² Overall, the facts involving variation reinforce the conclusion of the original computational analysis: islands are not a unified class. Most islands are TSL-2; these behave as expected of any linguistic blocking category, being finely parameterized and subject to extensive variation. The remainder do not seem like locality effects at all, and do not exhibit the same types of variability as the first group.

7. Extending the Model

Having explored the TSL-2 model of islands in considerable detail, let us step back and consider the wider linguistic landscape once again. By taking seriously the predictions of the model and the parallels with other long-distance dependencies, we were able to understand some additional complexities of the islands data beyond the original investigation by Graf (2022b). We might also ask how what we have learned could bring insight into the relation between movement and other syntactic locality phenomena. Building on the above discussion of movement along with existing TSL-2 analyses of agreement (Hanson, 2025c) and case (Hanson, 2023, 2025b), I ask whether long-distance agreement is subject to the subject and adjunct islands (Section 7.1) and how optionality in cross-clausal case/agreement configurations should be handled (Section 7.2). Each of these reveals a small but possibly important asymmetry between movement on the one hand, and case and agreement, on the other. Finally, Section 7.3 addresses the possibility of optional movement.

7.1. Islands for Agreement?

I mentioned in Section 2 that agreement is subject to similar locality restrictions as movement. In particular, both Rizzi's (1990) relativized minimality and Keine's (2019) probe horizons are thought to apply to each. Minimality is why (all else being equal) the structurally superior subject usually agrees with the finite verb, and finite C is usually

a horizon for such agreement. But what about long-distance (cross-clausal) agreement? By now, it is clear that such agreement exists. Some important work that established this beyond a reasonable doubt include the analysis of Tsez by Polinsky and Potsdam (2001) and that of Hindi-Urdu by Bhatt (2005).²³ Representative examples are provided below.

(36) Examples of long-distance agreement

a. Tsez (Polinsky & Potsdam, 2001)

Eni-r [CP uŽ-ā **magalu** b-āc'-ru-li] **b-iy-xo**
 mother-DAT boy-ERG bread.**III**.ABS III-eat-PSTPTCP-NMLZ **III**-know-PRS

'The mother knows the boy ate the bread.'

b. Hindi-Urdu (Bhatt, 2005)

Ram-ne [VP **roṭii** khaa-nii] chaah-**ii**
 Ram-ERG bread.**F** eat-INF.F want-PFV.**FSG**

'Ram wanted to eat bread.'

In all cases I am aware of, such agreement crosses a complement clause, finite (and nominalized) in the case of Tsez, and non-finite in the case of Hindi. Polinsky and Potsdam (2001) explicitly note the impossibility of agreement into an adjunct clause in Tsez, and the same seems to be true of Hindi-Urdu (Rajesh Bhatt, p.c.). As discussed earlier, the TSL-2 analysis of agreement in Hanson (2025c) also assumes that agreement across the boundary of a specifier or adjunct is completely impossible, which is what enabled the analysis in terms of spinal c-strings.

This state of affairs is reminiscent of the subject and adjunct islands, and suggests yet another direct parallel for relativized blocking across dependencies, in addition to horizons and relativized minimality. However, given that there are various known exceptions to the subject and adjunct islands for movement, it seems puzzling that there should be none for agreement. Furthermore, while we will see an example of case assignment into a complement clause momentarily, I am unaware of analogous instances of case assignment into a subject or adjunct clause. Thus, there may be a small formal distinction between movement and the other long-distance syntactic dependencies.

Additionally, it is possible for other structural information can cross a specifier or adjunct boundary. In particular, licensing operations such as binding or obligatory control can do so without issue. In contrast, when some element outside has needs which can only be met by something inside—as is the case for most movement and agreement—then dependency formation may be blocked. This is explained by the Graf and De Santo (2019) model as a licensor can send information needed by the licensee down every path, similar to what happens with across-the-board movement.²⁴

As it stands, our uncertainty regarding status of the subject and adjunct islands has increased. Two reasonable explanations for the effects are available: they are relativized blockers, as analyzed by Graf (2022b), or they are effects of the asymmetry of tree-based dependencies, as in Graf and De Santo (2019). The preponderance of the evidence suggests a two-pronged explanation for movement: amelioration of the subject and adjunct islands parallels other types of variation in the tree tier, while across-the-board movement derives from the top-down path restriction. Agreement, on the other hand, seems amenable to a purely path-based model, though we might then expect to find some analog of across-the-board movement. Given the otherwise strong symmetry between movement and agreement, this issue deserves a deeper investigation.

7.2. Optional Case and Agreement

Aside from movement and agreement, case has received considerable attention from a TSL-2 perspective. All existing TSL-2 analyses of case assignment utilize tree tiers. Vu et al.

(2019) is the first such analysis; that in [Hanson \(2023, 2025b\)](#) uses the same dependency structures we have been assuming, so I follow the latter presentation. Similar to the analysis of movement, we have one or more case tiers, each including the relevant case assigners ($\approx$ licensors) and assignees ($\approx$ licensees), plus any known blockers. Case assigners then determine the cases of all tier daughters of the relevant category, usually nominals.

[Hanson \(2023\)](#) uses this system to analyze a range of case patterns in Japanese. Our attention will be focused on optionality in cross-clausal case assignment as found in the finite ECM construction, in which the embedded subject received accusative case as if it was an object; nominative case is also available for the embedded subject.

(37) Finite ECM ([Kishimoto, 2018](#))

Ken ga [Eri {ga/o} kawaii to] omotteiru.

Ken **NOM** Eri {**NOM/ACC**} be.cute C think

‘Ken thinks that Eri is cute.’

While a syntactician’s first instinct is to look for a major structural difference—e.g., the ECM subject really is a matrix object—this does not seem to be tenable here (cf. [Kishimoto, 2018](#)). However, once we admit the possibility of long-distance case assignment, then such constructions can be treated in terms of selective opacity, much like we did for movement and agreement. This is exactly what was proposed in [Hanson \(2023\)](#): the embedding predicate optionally selects an special complementizer which is transparent for case assignment.²⁵ In effect, the variation that we see on the surface is encoded in the syntactic lexicon, the same as a verb that takes an optional object, or how a *wh*-element optionally bears the *wh* feature.

Without denying that this is reasonable solution in general, from the present perspective it seems suspicious that we find optionality in precisely these contexts, where a blocking element comes into play. Therefore, I would like to sketch an alternative based on the idea of probabilistic tier projection introduced earlier. Recall that while [Torres et al. \(2023\)](#) interpreted the score assigned to each structure as representing gradient wellformedness, this need not be the case. For example, we could collapse all non-zero scores to a value of 1, resulting in a categorical tree language which allows both structures. Given that both structures are fully grammatical, this seems appropriate. Since the exact details do not matter, I adopt this version for demonstration.

The proposed analysis is extremely similar to that in [Hanson \(2023\)](#). Simplifying to remove some irrelevant details, the relevant nodes are the verb (which is a case assigner), the nominals (the case assignees), and any complementizers (which are blockers); all others are invisible, including the predicate adjective. Now, in each of the relevant structures, the embedding predicate always selects the special “ECM” complementizer. (This is where we diverge from the original analysis, in which the ECM predicate can also select the normal embedded complementizer.) This complementizer is semi-permeable to case assignment, unlike the normal embedded complementizer, which is fully opaque. Since the exact projection probability does not matter for us, I arbitrarily set it to 0.5. (We could even say that projection is discrete with three values: no, maybe, and yes.) We always project other complementizers, the verb, and the nominal heads, but nothing else.

Next, the accusative case rule says that when there are two nominals in the same verbal case domain, the first is marked nominative and the second is accusative. Also, nominals with no case assigner at all are nominative by default. (This is a dependent case analysis, similar to that in [Baker \(2015\)](#); traditional case theory would also give the same result here, but the TSL-2 analysis is more convoluted.) So, the daughter string language of the verb allows an initial nominative, possibly followed by an accusative, and complementizers allow only nominatives among their daughters. In essence, the tier sister relation is being used as a proxy for c-command. This is not foolproof, but it will suffice for the data in question.²⁶

Now, we have two derivations whose dependency trees differ only in the case of the embedded subject, as illustrated in Figure 11. For each, there are two case-tier projections, of which only one is licit. For the structure in which the embedded subject is assigned nominative case, only the tier in which the complementizer appears is licit; for the accusative subject, the opposite is true. Since one licit case-tier is sufficient, both derivations are licit (modulo any other constraints), and we have derived the effect of optionality. One can also imagine how the projection probability could be altered to model relative production frequencies, provided that final scores are normalized in some way so as to generate a stochastic tree language.

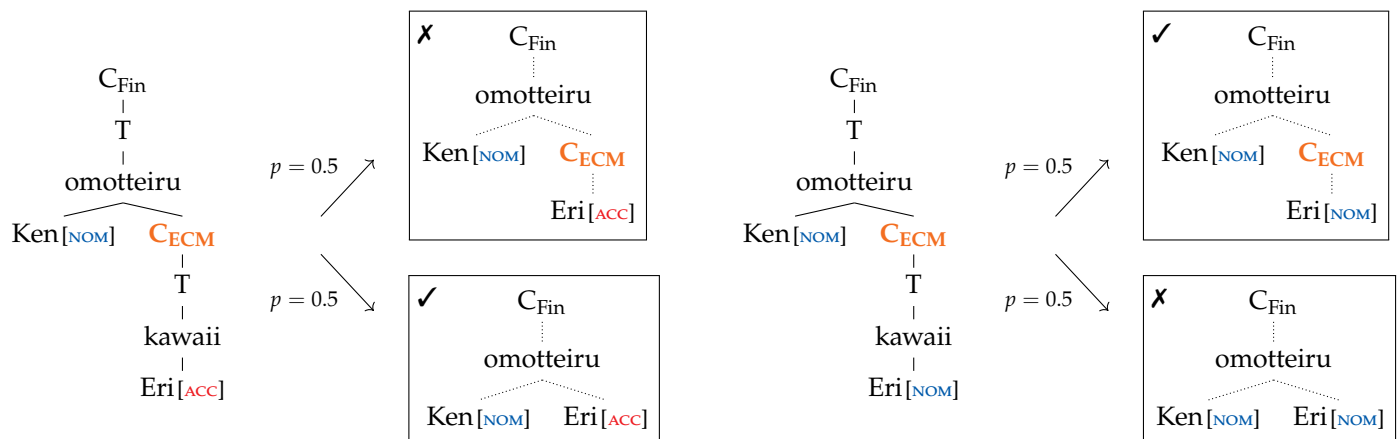


Figure 11. Optional ECM via probabilistic tier projection. (Left) the embedded subject is accusative when the complementizer is absent. (Right) the embedded subject is nominative when the complementizer is present.

This may not turn out to be the best way of handling such phenomena, but it is an interesting one, being a natural extension of the existing formal system. Furthermore, it seems to me that it could also be used to handle optional agreement such as that found with existential clauses in English, in which many speakers accept (38a) as well as (38b).

- (38) a. There **seem** to be some problems with this theory.
 b. % There **seems** to be some problems with this theory.

To emphasize that the approach is not specific to tree tiers, let us use the spinal c-string model of agreement as in Section 4.3. For brevity, I simply outline the analysis, which is diagrammed in Figure 12. First, either infinitival *to* or existential *there* would be semi-opaque; let us assume it is the latter. Additionally, we assume that singular morphology in (38b) results from lack of agreement with anything. As before, we mark the agreeing T (the probe) with ϕ^+ and D (the goal) with ϕ^- , and the TSL-2 constraints require that an element with ϕ^+ must be immediately followed by ϕ^- and vice versa. Additionally, finite T must agree with the nearest target, if possible, otherwise it does not agree with anything (cf. Preminger, 2014); this amounts to banning non-agreeing T followed by a non-agreeing D head. The net result is the same as before: exactly one of two agreement tiers is licit for each derivation, so optionality results.

Such phenomena seem less amenable to a lexical solution compared to the blocking of case by the complementizer—do we really want to say that there are two *there*'s whose only difference is their visibility? Hanson (2025c) presents a third alternative, which is that the finite verb agrees with both *there* and the logical subject, with optional realization relegated to the morphological component, but then this does not extend as well to the earlier example with case. Probabilistic tier projection is therefore appealing as a possible

means of unifying the optionality which is often found in long-distance dependencies.²⁷ I leave a more thorough investigation of this idea to future research.

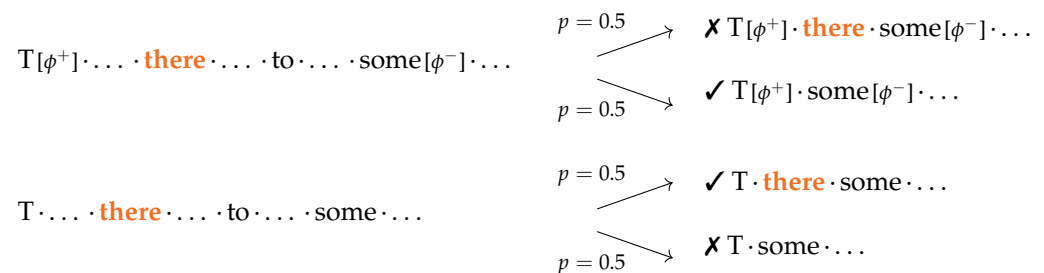


Figure 12. Optional agreement with probabilistic tier projection. (**Top**) agreement is allowed when *there* is absent. (**Bottom**) non-agreement is allowed when *there* is present.

7.3. What About Optional Movement?

Earlier, I noted that agreement and case seem to be more strictly confined to the complement spine, while movement admits more exceptions. The immediately preceding investigation of optionality suggests a second disconnect, which is that failure to form a grammatical movement dependency tends to result in ungrammaticality (or reduced grammaticality), while failure of agreement or case assignment results in a default form (cf. Preminger, 2014).

Recall the *wh*-movement example from Section 3.4: **Who did you wonder whether Mary saw?* This has just one derivation which can be constructed from our MG lexicon, which includes a *wh*-movement dependency between matrix CP and the embedded object. If *whether* is a blocker, then the sentence is ungrammatical, or at least degraded. But there is no alternative structure in which movement simply doesn't happen. That is to say, English speakers do not revert to *wh*-in-situ as a fallback. In MG terms, the relevant complementizer always bears the feature wh^+ , so movement is mandatory.

Compare this to what happened in the case and agreement examples above. Assignment of accusative case in Japanese is obligatory in ordinary transitive clauses, and becomes optional in the finite ECM construction; subject-verb agreement in English is obligatory in ordinary clauses, and only becomes optional in the non-canonical case where existential *there* is the subject. In contrast, it is difficult to find good examples of movement which is obligatory in simple structures and completely optional in those with a blocker.

Overall, the locality profile of these different syntactic operations is extremely similar, suggesting a common computational basis, but they are not the same in all details. There seems to be a principled difference between movement and other long-distance syntactic dependencies, and this might be related to what happens in the case of failed dependency formation. Further investigation of the behavior of syntactic islands in this manner, comparing them directly with other locality phenomena which are computationally similar, might help to pin down exactly where the parallel breaks down and why.

8. Conclusions

In this article, I have examined the case for treating islands as part of a larger family of relativized blockers, which we now understand to be elements which interfere with the requirements of the TSL-2 grammars regulating long-distance dependencies. No special mechanisms are needed to encode most island constraints, as they have the same form as other syntactic locality constraints affecting not only movement, but also agreement and case, as well as locality constraints in phonology. In this sense, islands are not unique, though they remain interesting due to their interactions with other aspects of language and in the details of how they do and do not exactly mirror other long-distance linguistic

dependencies. We also saw that certain islands, specifically the left branch condition and coordinate structure constraint, are quite different from the others from a computational perspective, and that this divide coincides with previously established findings related to variation and gradience—the TSL-2 islands appear much like other well-known locality constraints, while the non-TSL-2 islands seem to be more rigid and intertwined with other aspects of syntactic structure. Several questions remain, such as the precise relation between this computational factor and the others which are implicated in the facts surrounding islands, and whether there is a principled difference between islands and other syntactic blocking effects. Further probing of optionality and gradience from the perspective of the TSL-2 model may yield further insights.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: My thanks to three anonymous reviewers whose insightful comments greatly improved the presentation of this article. In particular, the comments by Reviewer 1 led to a much more detailed coverage of path-based subregular models and a better framing of several conceptual issues. In addition, I thank Bob Frank for a very helpful discussion of the implementation of syntactic optionality presented here.

Conflicts of Interest: The author declares no conflicts of interest.

Notes

- ¹ Such dependencies have also been studied from the perspective of the mapping from underlying to surface forms. See [Burness et al. \(2021\)](#) for a study of this mapping in terms of tier-based strictly local functions.
- ² In most examples, visible elements are bolded and blocking elements are additionally colored orange. Other visible elements may also be colored for visual emphasis. Blanks indicate gaps formed by movement, and agreeing noun phrases are underlined.
- ³ Commonly cited effects of successive cyclic movement can be recast as constraints on movement paths. See, e.g., [Bouma et al. \(2001\)](#) for an HPSG implementation and [Graf \(2022b\)](#) for a version with TSL MGs.
- ⁴ Originally, EPP refers to the Extended Projection Principle, which states that the specifier of TP must be filled. Here, I assume no particular theory for what motivates this movement.
- ⁵ The Ban on Improper Movement serves this role in earlier transformational theories. The Williams Cycle ([Williams, 2002](#)) also falls in this category.
- ⁶ For present purposes, the Minimal Link Condition ([Chomsky, 1995](#)) and Minimal Search ([Chomsky, 2000](#)) are equivalent.
- ⁷ The MG lexicon is not a theory of the human lexicon. It can accommodate any theory which provides a finite list of atoms.
- ⁸ I use the same notation for MG features as [Graf \(2022b\)](#). This is the same as that used in many works in subregular syntax, but differs slightly from [Stabler \(1997\)](#) and other sources.
- ⁹ The dependency tree omits the Merge and Move nodes found in a conventional MG derivation tree. This is similar to the system in [Kobele \(2002\)](#) and particularly [Kobele \(2012\)](#). Also see [Brody \(2000\)](#), on which [Kobele \(2002\)](#) is based, and [Adger \(2025\)](#) for a related idea.
- ¹⁰ We distinguish three groups of nodes: landing sites for the relevant movement type, movers of the same type, and other visible nodes such as islands. A landing site must have one mover in its daughter string; other nodes must *not* have any movers among their daughters. Formally, let l be a landing site for feature F , let m be a mover, and let o be any other node (i.e., an island). Also let $\bowtie/\bowtie$ be the left/right edge markers. Defined as a regular expression, the daughter string language $ds(x)$ for each element type x is:

$$ds(l) := \overline{m}^* m \overline{m}^* \quad ds(m) := \overline{m}^* \quad ds(o) := \overline{m}^*$$

The language $ds(l)$ is TSL-2. Its tier contains all elements m and the banned substrings are $\{\bowtie\bowtie, mm\}$. The other languages are SL-1, banning all unigrams m .

- ¹¹ See [Graf \(2018\)](#) and [Hanson \(2025a\)](#) regarding the complexity of adjunction in the context of classical MG derivation trees and dependency trees, respectively.

- Freezing effects should be further distinguished from so-called “criterial freezing” (Rizzi, 2006, 2010), which prohibits subsequent movement after a phrase reaches in a certain position in the functional spine.
- Another problem is that since even sentences without island violations are not judged to be perfect, the model fitting algorithm has no choice but to assign a fairly high projection probability to declarative C. For the proposed study, it would be important to factor out at least some of the irrelevant factors depressing acceptability to get a more accurate baseline.
- A reviewer asks whether we could model gradient islands solely using weighted daughter string languages, with no change to the tier projection function. This would not work because an island node need not occur along a movement path. The weight of a mover under the island node should be non-zero only when the island node is a daughter of the landing site, and zero otherwise, but this is impossible with TSL-2.
- The extra power of IBPS allow some additional constraints on movement to be expressed, notably freezing effects, but conversely it makes much weaker the typological predictions than the TSL-2 model. At the time Shafiei and Graf (2020) were writing, it was not yet apparent how pervasive TSL-2 constraints are in syntax.
- My thanks to a reviewer for reminding me of this.
- I could have instead chosen the a-string starting from the noun rather than the determiner. This makes no difference. It is only when we examine an a-string which does not contain the movement path that the problem discussed in Section 3.5 arises.
- I have omitted the EPP-movement features to reduce visual noise. Also, a null complementizer is used in both structures for consistency and because doing otherwise would create a *that*-trace violation.
- A reviewer asked about the precise relation between the TSL-2 grammar and the sensing tree automaton. The regular tree languages are the same class which is recognized by non-deterministic top-down finite tree automata. TSL-2 (either with tree tiers or paths) and sensing tree automata each restrict this class in a different way (i.e., they are incomparable).
- Regarding the VP-CP case, recall that we can distinguish verbs that take a CP complement from those that take a nominal complement by their MG features.
- For Starke (2001), it is because it is a more specific type of movement. For Rizzi (2001), it is because D-linked movement involves “nominal” features, while general *wh*-movement does not.
- The Element Constraint also seems to be suspended in examples like the following, borrowed from Chaves (2022):
- (i) a. Who did Sam pick up the phone and call ____ ?
 - b. How much can you drink ____ and still stay sober?
- Here, there is reason to think that such examples break symmetry of coordination, eliminating the source of the constraint. This seems rather like the explanation for the coordinate structure constraint given by Graf and De Santo (2019), and rather unlike from the kinds of variation discussed in this section.
- Icelandic also has a kind of long-distance agreement (Andrews, 1982), though the data is rather complicated; see Kučerová (2016) and references therein.
- A similar idea can be found in McFadden et al. (2019).
- It is common to attribute such locality phenomena to properties of the complementizer (cf. Lohninger et al., 2022). Another common solution is to posit that the transparent clause is “smaller”, lacking the highest function projection of an exploded CP. In absence of further considerations, it is difficult to distinguish these two styles of analysis.
- As discussed in Hanson (2025b), what we really want is a tree tier which combines directly relativized dominance with relativized d-command. A complete definition of such a system has not yet been formulated.
- As remarked by a reviewer, it is specifically the non-determinism introduced by probabilistic tier projection which is doing most of the work here. So, might imagine a spectrum of increasing generalization, from deterministic to non-deterministic grammars, and then to weighted grammars.

References

- Adger, D. (2025). *Mereological syntax: Phrase structure, cyclicity, and islands*. MIT Press.
- Andrews, A. D. (1982). Long Distance Agreement in Modern Icelandic. In P. Jacobson, & G. K. Pullum (Eds.), *The nature of syntactic representation* (pp. 1–33). Springer. [CrossRef]
- Baker, M. C. (2015). *Case* (No. 146). Cambridge University Press. [CrossRef]
- Barker, C., & Pullum, G. K. (1990). A theory of command relations. *Linguistics and Philosophy*, 13, 1–34. [CrossRef]
- Bhatt, R. (2005). Long Distance Agreement in Hindi-Urdu. *Natural Language & Linguistic Theory*, 23, 757–807. [CrossRef]
- Bobaljik, J. (2008). Where’s Phi? Agreement as a Postsyntactic Operation. In D. Harbour, D. Adger, & S. Béjar (Eds.), *Phi theory* (pp. 295–328). Oxford University Press. [CrossRef]
- Bouma, G., Malouf, R., & Sag, I. A. (2001). Satisfying constraints on extraction and adjunction. *Natural Language & Linguistic Theory*, 19(1), 1–65. [CrossRef]
- Brody, M. (2000). Mirror theory: Syntactic representation in perfect syntax. *Linguistic Inquiry*, 31(1), 29–56. [CrossRef]

- Burness, P. A., McMullin, K. J., & Chandlee, J. (2021). Long-distance phonological processes as tier-based strictly local functions. *Glossa: A Journal of General Linguistics*, 6(1), 99. [CrossRef]
- Chaves, R. P. (2022). Sources of discreteness and gradience in island effects. *Languages*, 7(4), 245. [CrossRef]
- Chaves, R. P., & Putnam, M. T. (2022). Islands, expressiveness, and the theory/formalism confusion. *Theoretical Linguistics*, 48(3–4), 219–231. [CrossRef]
- Chomsky, N. (1973). Conditions on Transformations. In S. R. Anderson, & P. Kiparsky (Eds.), *Festschrift for Morris Halle* (pp. 232–286). Holt, Rinehart and Winston.
- Chomsky, N. (1977). On wh-movement. In W. Thomas, P. W. Culicover, & A. Akmajian (Eds.), *Formal syntax*. Academic Press.
- Chomsky, N. (1995). *The Minimalist Program*. MIT Press. [CrossRef]
- 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* (pp. 89–155). MIT Press. Available online: <https://mitpress.mit.edu/9780262516839/step-by-step/> (accessed on 6 August 2026).
- Cinque, G. (1990). *Types of A'-dependencies*. MIT Press.
- De Santo, A. (2020). MG parsing as a model of gradient acceptability in syntactic islands. In *Proceedings of the Society for Computation in Linguistics (SCiL) 2020* (pp. 53–63). University of Massachusetts Amherst. [CrossRef]
- De Santo, A., & Graf, T. (2019). Structure sensitive tier projection: Applications and formal properties. In R. Bernardi, G. Kobele, & S. Pogodalla (Eds.), *Formal Grammar: 24th international conference* (pp. 35–50). Springer. [CrossRef]
- Frank, R. (2002). *Phrase structure composition and syntactic dependencies*. The MIT Press. [CrossRef]
- Frank, R. (2004). Restricting grammatical complexity. *Cognitive Science*, 28(5), 669–697. [CrossRef]
- Frank, R. (2006). Phase theory and tree adjoining grammar. *Lingua*, 116(2), 145–202. [CrossRef]
- Frey, W., & Gärtner, H.-M. (2002). On the treatment of scrambling and adjunction in Minimalist Grammars. In G. Jäger, P. Monachesi, G. Penn, & S. Wintner (Eds.), *Proceedings of Formal Grammar 2002* (pp. 41–52). Werner Frey & Hans-Martin Gärtner.
- Graf, T. (2017). The power of locality domains in phonology. *Phonology*, 34, 385–405. [CrossRef]
- Graf, T. (2018). Why movement comes for free once you have adjunction. In *Proceedings of CLS 53* (pp. 117–136). The Chicago Linguistic Society. Available online: <https://thomasgraf.net/output/graf18cls.html> (accessed on 6 August 2026).
- Graf, T. (2022a). Subregular linguistics: Bridging theoretical linguistics and formal grammar. *Theoretical Linguistics*, 48(3–4), 145–184. [CrossRef]
- Graf, T. (2022b). Typological implications of tier-based strictly local movement. In *Proceedings of the Society for Computation in Linguistics (SCiL) 2022* (pp. 184–193). University of Massachusetts Amherst. [CrossRef]
- Graf, T. (2023). Subregular tree transductions, movement, copies, traces, and the ban on improper movement. In T. Hunter, & B. Prickett (Eds.), *Proceedings of the Society for Computation in Linguistics (SCiL) 2023* (pp. 289–299). University of Massachusetts Amherst. [CrossRef]
- Graf, T. (2026). Minimalism and computational linguistics. In E. Leivada, & K. K. Grohmann (Eds.), *The Cambridge handbook of Minimalism and its applications*. Cambridge University Press.
- Graf, T., & De Santo, A. (2019). Sensing tree automata as a model of syntactic dependencies. In P. de Groote, F. Drewes, & G. Penn (Eds.), *Proceedings of the 16th Meeting on the Mathematics of Language (MoL 16)* (pp. 12–26). Association for Computational Linguistics. [CrossRef]
- Graf, T., & Hanson, K. (2025). Automata for subregular syntax: Syntax with strings attached. In *Proceedings of the Society for Computation in Linguistics (SCiL) 2025*. University of Massachusetts Amherst Libraries. [CrossRef]
- Graf, T., & Kostyszyn, K. (2021). Multiple Wh-Movement is not Special: The Subregular Complexity of Persistent Features in Minimalist Grammars. In A. Ettinger, E. Pavlick, & B. Prickett (Eds.), *Proceedings of the Society for Computation in Linguistics (SCiL) 2021* (pp. 275–285). University of Massachusetts Amherst. [CrossRef]
- Graf, T., & Shafiei, N. (2019). C-command dependencies as TSL string constraints. In G. Jarosz, M. Nelson, B. O'Connor, & J. Pater (Eds.), *Proceedings of the Society for Computation in Linguistics (SCiL) 2019* (pp. 205–215). University of Massachusetts Amherst. [CrossRef]
- Grosu, A. (1973). On the nonunitary nature of the coordinate structure constraint. *Linguistic Inquiry*, 4(1), 88–92. Available online: <http://www.jstor.org/stable/4177753> (accessed on 6 August 2026).
- Hanson, K. (2023). A TSL analysis of Japanese case. In T. Hunter, & B. Prickett (Eds.), *Proceedings of the Society for Computation in Linguistics (SCiL) 2023* (pp. 15–24). University of Massachusetts Amherst Libraries. [CrossRef]
- Hanson, K. (2024). Tiers, paths, and syntactic locality: The view from learning. In R. Futrell, C. Mayer, & N. Zaslavsky (Eds.), *Proceedings of the Society for Computation in Linguistics (SCiL) 2024* (pp. 107–116). University of Massachusetts Amherst Libraries. [CrossRef]
- Hanson, K. (2025a). Adjunction in (T)SL syntax. In *Proceedings of the Society for Computation in Linguistics (SCiL) 2025*. University of Massachusetts Amherst Libraries. [CrossRef]

- Hanson, K. (2025b). The typology of TSL case assignment. In *Proceedings of the 18th Meeting on the Mathematics of Language*. Association for Computational Linguistics.
- Hanson, K. (2025c). Tier-based strict locality and the typology of agreement. *Journal of Language Modelling*, 13(1), 43–97. [CrossRef]
- Heinz, J. (2018). The computational nature of phonological generalizations. In L. M. Hyman, & F. Plank (Eds.), *Phonological typology* (pp. 126–195). De Gruyter Mouton. [CrossRef]
- Heinz, J., Rawal, C., & Tanner, H. G. (2011). Tier-based strictly local constraints for phonology. In D. Lin, Y. Matsumoto, & R. Mihalcea (Eds.), *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human language technologies* (pp. 58–64). Association for Computational Linguistics. Available online: <https://aclanthology.org/P11-2011/> (accessed on 6 August 2026).
- Heuser, A., Martinez, H. V., & Yang, C. (2024). *The learnability of syntactic islands* (NELS 54). MIT.
- Hofmeister, P., & Sag, I. A. (2010). Cognitive constraints and island effects. *Language*, 86(2), 366–415. [CrossRef]
- Huang, J. C.-T. (1982). *Logical relations in Chinese and the theory of grammar* [Ph.D. thesis, Massachusetts Institute of Technology].
- Kaplan, R. M., & Zaenen, A. (1989). Dependencies, constituent structure, and functional uncertainty. In M. Baltin, & A. Kroch (Eds.), *Alternative conceptions of phrase structure*. The University of Chicago Press.
- Kayne, R. S. (1984). *Connectedness and binary branching*. De Gruyter. [CrossRef]
- Keine, S. (2019). Selective opacity. *Linguistic Inquiry*, 50(1), 13–62. [CrossRef]
- Keine, S. (2020). *Probes and their horizons*. MIT Press.
- Kishimoto, H. (2018). On exceptional case marking phenomena in Japanese. *Kobe Papers in Linguistics*, 11, 31–49.
- Kobebe, G. M. (2002). Formalizing mirror theory. *Grammars*, 5(3), 177–221. [CrossRef]
- Kobebe, G. M. (2012). Eliding the derivation: A Minimalist formalization of ellipsis. In *Proceedings of the International Conference on Head-Driven Phrase Structure Grammar*. CSLI Publications. [CrossRef]
- Kučerová, I. (2016). Long-distance agreement in Icelandic: Locality restored. *The Journal of Comparative Germanic Linguistics*, 19(1), 49–74. [CrossRef]
- Lambert, D. (2023). Relativized adjacency. *Journal of Logic, Language and Information*, 32(4), 707–731. [CrossRef]
- Lambert, D., & Rogers, J. (2020). Tier-based strictly local stringsets: Perspectives from model and automata theory. In A. Ettinger, G. Jarosz, & J. Pater (Eds.), *Proceedings of the Society for Computation in Linguistics (SCiL) 2020* (pp. 159–166). University of Massachusetts Amherst. [CrossRef]
- Lohninger, M., Kovač, I., & Wurmbrand, S. (2022). From prolepsis to hyperraising. *Philosophies*, 7(2), 32. [CrossRef]
- Mayer, C. (2021). Capturing gradience in long-distance phonology using probabilistic tier-based strictly local grammars. In A. Ettinger, E. Pavlick, & B. Prickett (Eds.), *Proceedings of the Society for Computation in Linguistics (SCiL) 2021* (pp. 39–50). University of Massachusetts Amherst. [CrossRef]
- Mayer, C. (2025). Reconciling categorical and gradient models of phonotactics. In *Proceedings of the Society for Computation in Linguistics (SCiL) 2025* (Vol. 8). University of Massachusetts Amherst Libraries. [CrossRef]
- McFadden, T., Sundaresan, S., & Zeijlstra, H. (2019). Deriving selective opacity in adjuncts. In *GLOW 42, May 8–10*. University of Oslo.
- McMullin, K., & Hansson, G. Ó. (2016). Long-distance phonotactics as tier-based strictly 2-local languages. In A. Albright, & M. A. Fullwood (Eds.), *Proceedings of the Annual Meetings on Phonology* (Vol. 2). Linguistic Society of America. [CrossRef]
- Pesetsky, D. (1987). Wh-in-situ: Movement and unselective binding. In *The representation of (in)definiteness*. MIT Press.
- Polinsky, M., & Potsdam, E. (2001). Long-distance agreement and topic in Tsez. *Natural Language & Linguistic Theory*, 19(3), 583–646. [CrossRef]
- Preminger, O. (2014). *Agreement and its failures*. MIT Press.
- Reinhart, T. (1981). Definite NP anaphora and c-command domains. *Linguistic Inquiry*, 12(4), 605–635. Available online: <http://www.jstor.org/stable/4178245> (accessed on 6 August 2026).
- Richards, N. (2014). A-bar Movement. In A. Carnie, D. Siddiqi, & Y. Sato (Eds.), *The Routledge handbook of syntax*. Routledge.
- Rizzi, L. (1990). *Relativized minimality* (No. 16). MIT Press.
- Rizzi, L. (2001). Relativized minimality effects. In *The handbook of contemporary syntactic theory* (pp. 89–110). John Wiley & Sons, Ltd. [CrossRef]
- Rizzi, L. (2006). On the form of chains: Criterial positions and ECP effects. In *Wh-movement: Moving on*. The MIT Press. [CrossRef]
- Rizzi, L. (2010). On some properties of criterial freezing. In *The complementizer phase* (pp. 17–32). Oxford University Press. [CrossRef]
- Rizzi, L. (2013). Locality. *Lingua*, 130, 169–186. [CrossRef]
- Rogers, J., & Pullum, G. K. (2011). Aural pattern recognition experiments and the subregular hierarchy. *Journal of Logic, Language and Information*, 20(3), 329–342. [CrossRef]
- Ross, J. R. (1967). *Constraints on variables in syntax* [Ph.D. thesis, Massachusetts Institute of Technology].
- Shafiei, N., & Graf, T. (2020). The subregular complexity of syntactic islands. In A. Ettinger, G. Jarosz, & J. Pater (Eds.), *Proceedings of the Society for Computation in Linguistics (SCiL) 2020* (pp. 421–430). University of Massachusetts Amherst. [CrossRef]

- Sprouse, J., Caponigro, I., Greco, C., & Cecchetto, C. (2016). Experimental syntax and the variation of island effects in English and Italian. *Natural Language & Linguistic Theory*, 34(1), 307–344. (Erratum in 2016, *Natural Language & Linguistic Theory*, 34, 345–349). [CrossRef]
- Stabler, E. P. (1997). Derivational Minimalism. In C. Retoré (Ed.), *Logical aspects of computational linguistics*. Springer. [CrossRef]
- Stabler, E. P. (2011). Computational perspectives on Minimalism. In C. Boeckx (Ed.), *The Oxford handbook of linguistic Minimalism* (pp. 617–643). Oxford University Press. [CrossRef]
- Starke, M. (2001). *Move dissolves into Merge: A theory of locality* [Ph.D. thesis, University of Geneva].
- Szabolcsi, A. (2006). Strong vs. weak islands. In *The Blackwell companion to syntax* (pp. 479–531). John Wiley & Sons, Ltd. [CrossRef]
- Szabolcsi, A., & Zwarts, F. (1992). Weak islands and an algebraic semantics for scope taking. *Natural Language Semantics*, 1(3), 235–284. Available online: <http://www.jstor.org/stable/23748421> (accessed on 6 August 2026). [CrossRef]
- Torres, C., Hanson, K., Graf, T., & Mayer, C. (2023). Modeling island effects with probabilistic tier-based strictly local grammars over trees. In T. Hunter, & B. Prickett (Eds.), *Proceedings of the Society for Computation in Linguistics (SCiL) 2023* (pp. 155–164). University of Massachusetts Amherst Libraries. [CrossRef]
- Vu, M. H., Shafiei, N., & Graf, T. (2019). Case assignment in TSL syntax: A case study. In G. Jarosz, M. Nelson, B. O'Connor, & J. Pater (Eds.), *Proceedings of the Society for Computation in Linguistics (SCiL) 2019* (pp. 267–276). University of Massachusetts Amherst. [CrossRef]
- Wexler, K., & Culicover, P. (1980). *Formal principles of language acquisition*. MIT Press.
- Williams, E. (2002). *Representation theory*. MIT Press.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.