A Tier-Based Model of Syntactic Agreement

Kenneth Hanson
Stony Brook University

Abstract

Agreement configurations across languages show extensive variation, including (1) which elements agree, (2) which elements may intervene, and (3) whether elements obtain their value from above or below. Why should such variation exist, and what are its limits? A wide range of linguistic dependencies have been shown to belong to the class of *tier-based strictly 2-local* (TSL-2) (Graf 2022a). I argue that agreement is yet another TSL-2 phenomenon, providing an explanation for several general properties and parameters of variation which is based on considerations of computational efficiency (Chomsky 1995; Lambert *et al.* 2021).

1 Introduction

Syntactic agreement displays a number of paradoxical properties. It applies over a distance, but is predictably blocked by certain elements, such as the finite clause boundary, and the exact set of blockers varies across individual dependencies and languages (Keine 2019). Furthermore, agreement usually targets the closest visible element bearing the relevant features (Rizzi 1990; Rizzi 2001), yet here too, the conditions for visibility vary (Preminger 2014). In addition to these complications for locality, agreement can proceed upward or downward, at least descriptively, and can also iterate, resulting in feature spreading. Putting all of this together, it is clear that agreement is subject to a number of fine-grained parameters.

But where do the parameters come from? An intriguing answer can be found if we think about the type of computation which is required to pair up agreeing elements. In this paper, I show a range of well-known agreement patterns belong to the class of TIER-BASED STRICTLY 2-LOCAL (TSL-2) formal languages, mirroring previous results on long-distance phonotactics (McMullin 2016), movement (Graf 2022b), and case assignment (Vu *et al.* 2019; Hanson 2023). I also emphasize how the natural parameters of this formal model align closely with what we observe.

Assuming that the language faculty includes the ability to compute TSL-2 patterns, we therefore have an explanation for the existence of similar long-distance phenomena across linguistic domains (Graf 2022a). We can even partially explain why it should be TSL-2 rather than some other class of formal patterns that is relevant to natural language by appealing to considerations of computational efficiency, as proposed by Chomsky (1995). This is because TSL-2 falls among the simplest formal classes which can handle long-distance dependencies at all, and furthermore belongs to a select subset which can be learned with realistic constraints on input and memory (Lambert *et al.* 2021).

The remainder of this paper proceeds as follows. In Section 2, I introduce the TSL-2 formal languages with a simple example from phonology, and discuss how they encapsulate a notion of relativized locality which closely matches linguists' intuitions. Next, I show how syntactic dependencies can be analyzed as TSL-2 patterns in a Minimalist framework (Section 3), how the analysis derives several core properties of syntactic locality (Section 4), and how it predicts the types of formal variation mentioned above (Section 5). From there, we consider the close parallel between syntactic agreement and long-distance harmony (Section 6), closing with a discussion of the strengths and limitations of the model (Section 7).[1]

2 TSL patterns

In what follows, we model words as strings of symbols, and the constraints on those strings as FORMAL LANGUAGES, which are sets of strings, possibly infinite. A formal language is TSL if it can be described in terms of forbidden substrings over a TIER PROJECTION containing just the salient elements, treating the rest as invisible; the set of projecting elements and the set of constraints together constitute a TSL grammar. A TSL grammar/language is TSL-2 if all constraints reference only two elements.

For example, suppose we have a language with (symmetric) sibilant harmony, such that words like 'sasakasa' and 'ʃaʃakaʃa' are licit, but 'sasakaʃa' and 'saʃakasa' are not. Furthermore, suppose that harmony is blocked by [t], similar to Slovenian, so that words like 'sasataʃa' are licit even though they violate strict harmony.

The tier for this pattern consists of segments {s, ʃ, t}, which are the sibilants and the blocker. All other segments are invisible. The constraints on the tier are {*sʃ, *ʃs}. As a result, harmony is enforced except when [t] intervenes. Since the substrings we need to consider have a maximum length of two, this pattern is TSL-2. Several licit and illicit words and their tiers are presented in (1), with constraint violations highlighted.

(1) Example words and tiers for sibilant harmony with blocking

	Word	Tier Projection	Tier Substrings
✓	**s**a**s**aka**s**a	sss	{ss}
✓	**ʃ**a**ʃ**aka**ʃ**a	ʃʃʃ	{ʃʃ}
✗	**s**a**s**aka**ʃ**a	ssʃ	{ss, **sʃ**}
✗	**s**a**ʃ**aka**s**a	sʃs	{**sʃ**, **ʃs**}
✓	**s**a**s**a**t**a**ʃ**a	sstʃ	{ss, st, tʃ}

The TSL languages were originally defined by Heinz *et al.* (2011) to model long-distance phonotactics, as we have done here. Although inspired by autosegmental phonology (Goldsmith 1976), the tier of a TSL language is conceptually distinct, and now has several equivalent mathematical characterizations (Lambert & Rogers 2020). At its heart, this notion of a tier encapsulates a mathematical concept that has been termed RELATIVIZED ADJACENCY by Lambert (2023). This can be visualized in multiple ways, including a moving window of attention which skips invisible

[1]This is an informal presentation of ideas which are worked out more completely in a formal system based on MG derivation trees (Hanson 2024a); I refer readers interested in the technical details to this and other cited works.

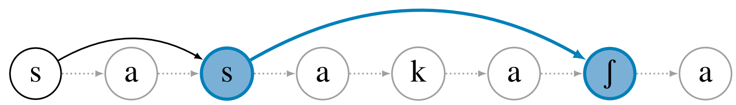


Figure 1: Tier-based string model for the word 'sasakaʃa', which violates the constraint *sʃ and is therefore illicit according to our example grammar. Although the offending elements are not adjacent in the base word, they are adjacent on the tier.

elements, or as a relational structure which consists of arcs which do the same. The latter visualization is shown in Figure 1.

TSL languages have since been generalized to functions to model phonological maps (Burness *et al.* 2021) and to trees to model syntactic constraints (Graf 2022a). In each case, they have been found to provide a tight upper bound for most long-distance linguistic dependencies, and the TSL-2 languages provide a particularly good fit for the observed typology (McMullin 2016; Graf 2022b). This is because they derive a characteristic property of linguistic locality, which is that a single intervener of the right type is enough to break a long-distance dependency. In TSL-2 terms, the presence of such an intervener means that dependent elements can never appear in the same window, as with the above example of blocking by [t]. In the next section, we will see how this idea can be applied to syntactic agreement, and in the following sections, we explore some consequences of the model.

3 A TSL model of agreement

We begin with a few syntactic assumptions. I adopt a version of bare phrase structure, in which all projections of a head are in some sense "the same element". I also assume feature driven selection, movement, and agreement, and some mechanism for case assignment/licensing. Note that I use the term *agreement* rather than *Agree* to emphasize that we are primarily interested in feature matching/copying and not other phenomena that are sometimes subsumed under Agree.

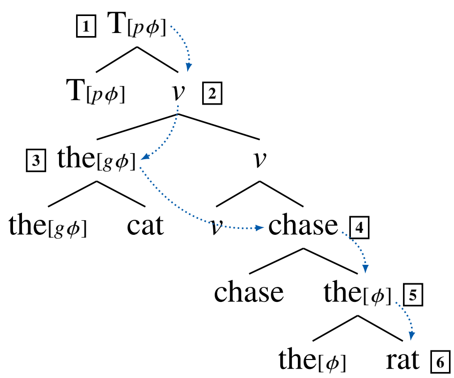
Next, in order describe the class of licit agreement configurations in a given language as a TSL-2 formal language, we use the following diacritics to indicate (potentially) agreeing elements:

- *p*F indicates a PROBE, bearing initially unvalued features of type F
- *g*F indicates a GOAL, which values some probe *in the current derivation*
- F marks other elements valued for F, but which ultimately do not agree

Finally, we need to specify the set of licit agreement configurations. Adapting the approach of Graf & Shafiei (2019), I propose that the SEARCH PATH of an agreement probe, annotated as just described, must obey a TSL-2 grammar. I assume that the search path itself follows the COMPLEMENT SPINE of the tree, based on ideas from Graf & De Santo (2019). As an example, consider the point at which finite T is merged in a simple transitive sentence such as *The cat chases the rats*, shown in (2) below. We start at the root and follow the complement spine, adding each head in the position first encountered, yielding the string $\mathrm{T}_{[p\phi]} \cdot v \cdot \mathrm{D}_{[g\phi]} \cdot \mathrm{V} \cdot \mathrm{D}_{[\phi]} \cdot \mathrm{N}$.

Graf & Shafiei (2019) call such paths command strings or c-strings. The ordering relation relation they encapsulate is derivational command, or d-command, which represents the relative height of each maximal projection, such that a head precedes its specifier and complement, and the specifier also precedes the complement. For our purposes, the main advantage of adopting d-command over the more conventional c-command is that it allows us to model domain-based locality (such as probe horizons) and intervention-based locality (such as relativized minimality) together in a uniform manner.

(2) The cat **chases** the rats.



Path:	$\mathbf{T}_{[p\phi]} \cdot v \cdot \mathbf{D}_{[g\phi]} \cdot \mathrm{V} \cdot \mathrm{D}_{[\phi]} \cdot \mathrm{N}$
Tier:	$\mathbf{T}_{[p\phi]} \cdot \mathbf{D}_{[g\phi]} \cdot \mathrm{D}_{[\phi]}$
Violations:	n/a

Technically, every node has its own c-string, but I assume that the only c-strings which are relevant for agreement are those that trace the main spine, or the spine of a complex specifier or adjunct; see Hanson (2024a) for details. Also, note that not all syntactic constraints can be handled by c-strings; for some, a tree-based analog of TSL, as in Graf (2022b), is needed.

We can now state a general principle for what constitutes a well-formed configuration for subject-verb agreement in a language like English:

(3) A probe must be immediately followed by its goal on a tier projected from its command string, and a goal immediately preceded by its probe.

This will be modified somewhat for certain agreement patterns in other languages. Additionally, we have the following substantive principle:

(4) The tier regulating agreement for feature F contains all elements which potentially agree in feature F, and all blockers for agreement in feature F.

In the case of English, the tier for ϕ-agreement therefore contains finite T and all DPs (the agreeing elements) as well as finite C (a blocker), as will be discussed in detail below. The constraints in the TSL-2 grammar therefore include those in (5).

(5) Constraints for English:

$*\mathrm{T}_{[p\phi]} \cdot \mathrm{D}_{[\phi]}$	–	closest D fails to agree
$*\mathrm{T}_{[p\phi]} \cdot \mathrm{C}$	–	clause boundary intervenes
$*\mathrm{D}_{[\phi]} \cdot \mathrm{D}_{[g\phi]}$	–	non-closest D agrees
$*\mathrm{D}_{[g\phi]} \cdot \mathrm{D}_{[g\phi]}$	–	multiple goals share probe

If we project the tier for the current example, we find that the probe immediately precedes the goal, so no constraints are violated. This is summarized on the right side of (2). Notice that even if there were many other irrelevant elements in the structure, such as verbal auxiliaries and adjuncts, all of these would be invisible on the tier, so finite T and the closest DP remain adjacent in the appropriate sense.

Note that, in general, each type of probe has has its own tier and constraints; *wh*-movement, for example, is not sensitive to non-*wh* DPs, nor is it hindered by declarative C. This is in accord with recent theories of Agree in which locality restrictions are relativized to individual probes (Deal 2015; Keine 2019, a.o.).[2] Thus, it is most accurate to say that the full system, including all movement, case, and agreement dependencies, is multi-TSL, or MTSL (De Santo & Graf 2019).

4 Consequences for locality

Let us now see how this analysis derives several basic properties of syntactic locality:

- **Minimality:** a potential goal intervenes on the tier

$$\mathrm{T}_{[p\phi]} \ldots \mathbf{D}_{[\phi]} \ldots \mathrm{D}_{[g\phi]}$$

- **Invisibility:** a hypothetical goal does not appear on tier

$$\mathrm{T}_{[p\phi]} \ldots \text{there} \ldots \mathrm{D}_{[g\phi]}$$

- **Blocking:** a non-agreeing element intervenes on the tier

$$\mathrm{T}_{[p\phi]} \ldots \mathbf{C} \ldots \mathrm{D}_{[g\phi]}$$

English is perhaps not the best test case due to the limited scope of agreement in this language, but it will suffice for illustration. In the next section, we will see more interesting examples from other languages.

4.1 Minimality

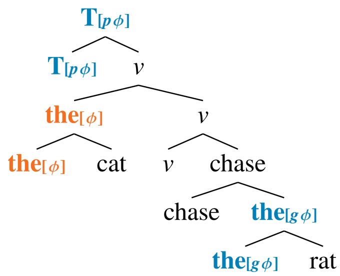
The principle of relativized minimality (Rizzi 1990; Rizzi 2001), in its modern form, says that two elements X and Y cannot enter a dependency involving feature F if there is a (visible) intervening element Z which also bears F. Put another way, a probe must take the closest potential goal, with exceptions as discussed below.

In the previous example, we saw how the TSL-2 constraints on agreement allow finite T to agree with the closest visible DP. Now, we confirm that violations of relativized minimality are ruled out by constructing a derivation in which finite T agrees with the object instead of the subject, shown in (6).

---

[2] Phases and similar notions of absolute locality are of little interest here. The phase impenetrability condition (Chomsky 2001) may force movement to the phase edge, but we must still state which categories are transparent to which dependencies. See Keine (2019) for elaboration on this point.

(6) * The cat **chase** the rats.



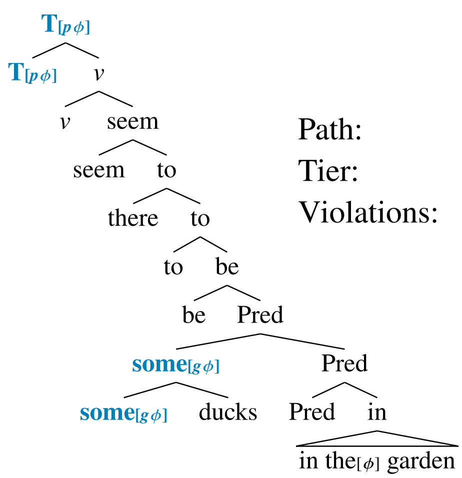
Path:	$\mathbf{T}_{[p\phi]} \cdot v \cdot \mathbf{D}_{[\phi]} \cdot \mathrm{V} \cdot \mathbf{D}_{[g\phi]} \cdot \mathrm{N}$
Tier:	$\mathbf{T}_{[p\phi]} \cdot \mathbf{D}_{[\phi]} \cdot \mathbf{D}_{[g\phi]}$
Violations:	$*\mathrm{T}_{[p\phi]} \cdot \mathrm{D}_{[\phi]}$, $*\mathrm{D}_{[\phi]} \cdot \mathrm{D}_{[g\phi]}$

In this case, the non-agreeing subject D head intervenes on the tier, breaking adjacency between the probe and intended goal. As a result, agreement is impossible.

4.2 Invisibility

The TSL model leaves open the possibility that certain DPs may fail to appear on the tier, creating apparent exceptions to relativized minimality. The canonical example is case-sensitive agreement, discussed in the next section, but we can demonstrate something similar in English with existential *there*, as shown in (7).

(7) There **seem** to be some ducks in the garden.



Path:	$\mathbf{T}_{[p\phi]} \cdot v \cdot \mathrm{V} \cdot \mathrm{T} \cdot \mathrm{there} \cdot v \cdot \mathbf{D}_{[g\phi]} \cdot \mathrm{P} \cdot \mathrm{D}_{[\phi]} \cdot \mathrm{N}$
Tier:	$\mathbf{T}_{[p\phi]} \cdot \mathbf{D}_{[g\phi]} \cdot \mathrm{D}_{[\phi]}$
Violations:	n/a

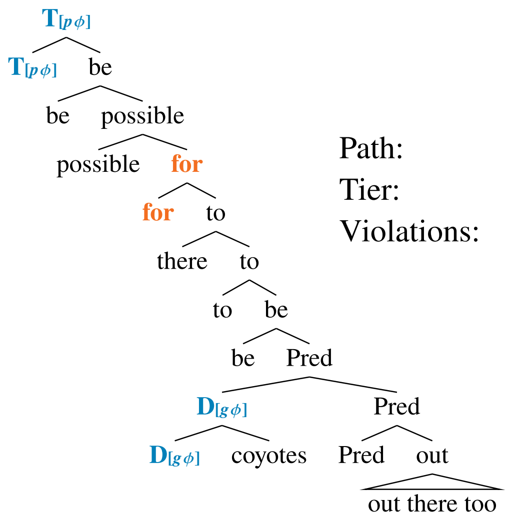
In this case, there is a plausible substantive explanation for the invisibility of *there*: that it is deficient in ϕ-features. Whatever the reason, *there* is omitted from the tier, allowing long-distance agreement when it would otherwise be impossible.[3]

4.3 Blocking

We can also imagine a situation in which a non-agreeing element nonetheless projects on the tier, blocking agreement where it intervenes. For example, C appears to block long-distance agreement in English, as in (8). Assume for the sake of demonstration that expletive "it" is inserted late and does not agree.

[3] Long-distance agreement is optional for many English speakers. See Hanson (2024a) for discussion and two possible TSL-2 analyses.

(8) * It **are** possible for there to be <u>coyotes</u> out there too.
cf. It **is** possible...



Path: $\mathbf{T}_{[p\phi]} \cdot v \cdot \mathrm{V} \cdot \mathbf{C} \cdot \mathrm{T} \cdot \text{there} \cdot v \cdot \text{Pred} \cdot \mathbf{D}_{[g\phi]} \cdot \ldots$

Tier: $\mathbf{T}_{[p\phi]} \cdot \mathbf{C} \cdot \mathbf{D}_{[g\phi]}$

Violations: $*\mathrm{T}_{[p\phi]} \cdot \mathrm{C}$, $*\mathrm{C} \cdot \mathrm{D}_{[g\phi]}$

This effectively implements Keine's (2019) notion of a HORIZON, which is a blocking category relativized to the individual probe. For example, C is a horizon for EPP-movement in English, but not in a language which allows hyperraising. It is also presumably a horizon for ϕ-agreement in English, though not for *wh*-movement or relative clause formation. A major advantage of this approach to blocking is that it is not overly reliant on other analytical assumptions, and therefore less susceptible to falling apart if those assumptions change.

4.4 Summary

We have seen how a variety of syntactic locality phenomena derive from the assumption that agreement is TSL-2 over c-strings. It is important to note that neither the tier projection nor the finite window alone is adequate. Tiers allow long-distance dependencies to be treated as if local (relativized adjacency). The finite constraint window limits the power of the system by eliminating arbitrary global constraints; see Section 7 for further discussion. Together, they provide a simple system which is just powerful enough to handle both invisibility and blocking, the hallmarks of a long-distance linguistic dependency.

5 Parameters of variation

The parameters for a TSL-2 language (the tier elements and constraints) correspond neatly to variation in long-distance dependencies, particularly:

- Visibility – which elements are relevant and which are ignored?
- Iteration – if you allow AB and BB, then you get ABB, ABBB, etc.
- Directionality – do we ban AB or BA?

We have already seen several examples of variation in visibility (existential *there* and probe horizons). In this section, we will add case-sensitive agreement. Next, we will treat concord in the DP as an instance of iteration/chaining. Finally, we will discuss how mirroring the constraints flips the direction of agreement.

5.1 Case-sensitive agreement

In many Indo-European languages, there is a strong correlation between agreement and case, since both verbal agreement and nominative case target the highest DP in the finite clause. This correlation comes apart in other languages. In Hindi, the verb agrees with the highest nominative argument, which may not be the subject. This can be seen by comparing an imperfective sentence, in which the subject is nominative, with the corresponding perfective, in which it is ergative.

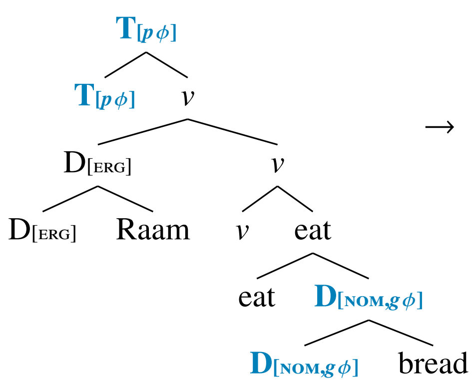
(9) Hindi verbal agreement ignores ergatives (Mahajan 1990)

a. Raam roṭii khaataa thaa.
Raam.M.NOM bread.F.NOM eat.IPFV.M be.PST.M
'Raam ate bread (habitually).'

b. Raam-ne roṭii khaayii.
Raam.M-ERG bread.F.NOM eat.PFV.F
'Raam ate bread.'

The TSL analysis is simple: we project D only if nominative; otherwise, both the tier and its constraints are unchanged compared to English. The derivation for (9b), which features object agreement, is shown below. Note: we ignore agreement on the non-finite verb for simplicity.[4]

(10) *Raam ate bread.* (Ergative subject, object agrees)



$\rightarrow$ $\mathbf{T}_{[p\phi]} \cdot v \cdot \mathrm{D}_{[\textsc{erg}]} \cdot \text{eat} \cdot \mathbf{D}_{[\textsc{nom},g\phi]} \cdot \text{bread}$

$\downarrow$

$\mathbf{T}_{[p\phi]} \cdot \mathbf{D}_{[\textsc{nom},g\phi]}$

Compared to the earlier example of existential *there*, it is not plausible to claim that ergative DPs lack ϕ-features, so any attempt to maintain relativized minimality in its strictest form must add auxiliary assumptions to ensure to ensure that the

[4]There is much more to be said about Hindi verbal agreement, in particular the fact that infinitival agreement is parasitic on successful agreement by the finite verb (Bhatt 2005). While space prevents me from providing a full analysis here, the basic idea is that agreement proceeds in two steps: first, T agrees with the closest visible DP, then all verbs along the agreement path are valued if applicable. The tiers for these two steps are different, as the verbs must be ignored when setting up the initial relation between T and its goal, as we have done here.

ϕ-features of a DP are made accessible to the outside only when they should be. In contrast, in the TSL analysis, no further complications are required: we simply state the visibility conditions directly as we did in our treatment of horizons.

Related to this point, it is also important to note that there is nothing inherent about ergative DPs which renders them invisible for agreement. In fact, they not *not* invisible in the related language Nepali, as shown in (11).

(11) Agreement with ergative in Nepali (Coon & Parker 2019)

a. Maile yas pasal-mā patrikaā kin-ē.
1SG.ERG DEM store-LOC newspaper.ABS buy-**1SG**
'I bought the newspaper in this store.'

b. Ma thag-ī-ē.
1SG.ABS cheat-PASS-**1SG**
'I was cheated.'

Both Hindi and Nepali also have dative-marked DPs, which are invisible in both languages. Neither of these facts is a problem for the TSL analysis. For Hindi, we projected only $D_{[NOM]}$. For Nepali, we project $D_{[NOM]}$ and $D_{[ERG]}$ but not $D_{[DAT]}$. Generally speaking, there is an implicational hierarchy of visibility for agreement, such that if a given case is visible, then so are the cases above it in the hierarchy. This is shown in (12).

(12) **Case visibility hierarchy** (Bobaljik 2008)
Unmarked (NOM, ABS) > Marked (ACC, ERG) > Oblique (DAT, GEN)

We can encode the constraints corresponding to each position in the hierarchy, though the contents of the hierarchy itself require a separate explanation, such as the substance of the features involved. This is not specific to TSL-2 languages: there can be no formal reason why nominatives, ergatives, and datives occupy the specific positions in the case visibility hierarchy that they do. For example, in an analysis which treats invisible cases as involving a PP shell, we must still stipulate which cases correspond to which structures, and in which languages. Even so, a formal explanation can be given for the general prevalence of implicational hierarchies throughout language via the principle of MONOTONICITY; see Graf (2019) for discussion.

5.2 Concord in the DP

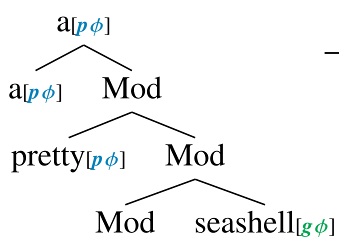
In previous examples, we considered only agreement between a single pair of elements. But it is also common for a chain of elements to share some feature or features. From the perspective of the TSL-2 grammar, this is quite unexceptional: we simply permit multiple probes in sequence when appropriate. Consider gender concord in the DP, as shown in the German example in (13).

(13) Ich habe [eine hübsche Muschel] gefunden.
I have [a.**F** pretty.**F** seashell.**F**] found
'I found a pretty seashell.'

Let us assume that gender is inherent on the noun (or a suitably low functional projection), and spreads to other heads and adjuncts along the functional spine. In the TSL-2 grammar, we simply relax the constraints in order to allow multiple probes in sequence. Specifically, we must allow permit $\mathrm{D}_{[p\phi]} \cdot \mathrm{A}_{[p\phi]}$ and $\mathrm{A}_{[p\phi]} \cdot \mathrm{A}_{[p\phi]}$. Naturally, we also allow $\mathrm{D}_{[p\phi]} \cdot \mathrm{N}_{[g\phi]}$ and $\mathrm{A}_{[p\phi]} \cdot \mathrm{N}_{[g\phi]}$.

A sample derivation for just the DP *a pretty seashell* is provided in (14) below. For simplicity, I use a functional head Mod to introduce adjectival modifiers, which are invisible on the tier. Other analyses, such as direct adjunction or an elaborated functional hierarchy, would work just as well.

(14) Gender concord in German



$\rightarrow \quad \mathrm{D}_{[p\phi]} \cdot \mathrm{Mod} \cdot \mathrm{A}_{[p\phi]} \cdot \mathrm{N}_{[g\phi]}$

$\downarrow$

$\mathrm{D}_{[p\phi]} \cdot \mathrm{A}_{[p\phi]} \cdot \mathrm{N}_{[g\phi]}$

Aside from this type of concord, part of the full analysis for Hindi verbal agreement also involves sequences of probes on verbal heads. Yet another example would be iterated complementizer agreement, as in Dinka (Van Urk 2015); see Hanson (2024a) for an analysis of the latter. Also note that the morphological realization of determiners and adjectives in German depends in part on case and definiteness. Assuming these to originate on the D head, we need concord to be upward looking for these features. We turn to this issue next.

5.3 Upward agreement

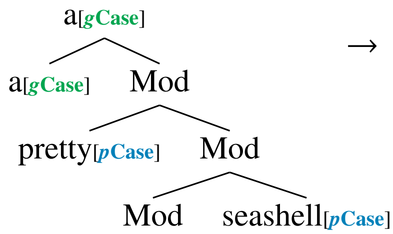
Directionality of agreement is a famously thorny issue in Minimalism. For our purposes, the critical fact is that the element which is intuitively the origin of some feature may appear above or below others which copy that feature. Here, I maintain the assumption that the initially unvalued element is the probe. In this case, all we need to do from the perspective of the TSL-2 grammar is to mirror the constraints.

Consider again the example of DP-internal concord in German, repeated in (15), this time with annotation for case.

(15) Ich habe [eine hübsche Muschel] gefunden.
I have [a.ACC pretty.ACC seashell.ACC] found

Since case, by hypothesis, case is inherent on D (or is assigned to D before spreading to the rest of the DP), categories like A and N need to probe upward to find a value. In the TSL-2 grammar, this means that we allow $\mathrm{D}_{[g\mathrm{Case}]} \cdot \mathrm{N}_{[p\mathrm{Case}]}$ instead of $\mathrm{D}_{[p\mathrm{Case}]} \cdot \mathrm{N}_{[g\mathrm{Case}]}$, and $\mathrm{D}_{[g\mathrm{Case}]} \cdot \mathrm{A}_{[p\mathrm{Case}]}$ instead of $\mathrm{D}_{[p\mathrm{Case}]} \cdot \mathrm{A}_{[g\mathrm{Case}]}$. As before, we also allow successive probes. How this plays out in our running example is shown in (16). We can handle definiteness agreement on the adjective in the same way, letting definiteness spread onto the adjectives from above. In this way, the present example also demonstrates how multiple tiers, each regulating a different feature, ensure that each dependency obeys its language-specific constraints.

(16) Case concord in German



$\rightarrow$ $\mathrm{D}_{[g\mathrm{Case}]} \cdot \mathrm{Mod} \cdot \mathrm{A}_{[p\mathrm{Case}]} \cdot \mathrm{N}_{[p\mathrm{Case}]}$

$\downarrow$

$\mathrm{D}_{[g\mathrm{Case}]} \cdot \mathrm{A}_{[p\mathrm{Case}]} \cdot \mathrm{N}_{[p\mathrm{Case}]}$

So, we have seen that variation in directionality is yet another slight variant on the basic TSL-2 template for agreement, just like variation in visibility and presence or absence of chaining. However, there is something odd about upward agreement in the present system, since the relevant c-string does not even exist until the point in the derivation when the goal is first merged, whereas we assumed earlier that it was the point when the probe is merged that matters. In contrast, in the formal system used by Graf & Shafiei (2019) and related work, the entire derivation is available in a static representation known as a DERIVATION TREE, so the issue simply does not arise.

There are, of course, various solutions to the problem. We could modify the lexicon so that probing always proceeds downward (Pesetsky & Torrego 2007) or allow valued features to 'probe' downward and value a lower goal (Adger 2003). The TSL-2 analysis is compatible with these alternatives, but the existence of upward agreement would become accidental rather than a prediction of the formalism. Alternatively, we could asserting that every step of the derivation provides new chance for agreement, reminiscent of the theory of cyclic Agree (Béjar & Rezac 2009). I wish to suggest that the best route forward may be to adopt tier-based FUNCTIONS, similar to those used by Burness *et al.* (2021) to model phonological maps. In such a system, feature copying would take place each time an appropriate configuration arises. I leave exploration of this possibility to future work.

5.4 Summary

In this section, we have seen how various small modifications of the simple TSL-2 grammar, initially formulated to model subject-verb agreement, each correspond to common instances of cross-linguistic variation. This is summarized in Table 1.

Example Phenomenon	**Tier Elements**	**Tier Constraints**
Subject-verb agreement (canonical)	All T/D/C	Strict pairing of $p\phi$ and $g\phi$
Case-sensitive agreement	All T/C, D only if right case	(as above)
Concord within DP	All D/Adj/N	Allow sequential $p\phi$
Upward agreement	(as above)	Swap order of $p\phi$/$g\phi$

Table 1: Variation in agreement configurations and the corresponding parameters of the TSL-2 grammar.

There are several other possible points of variation which we have not explored. For example, we have not considered cases where a single probe shares multiple goals, or where no goal is found, resulting in default agreement. I treat these and several others in Hanson (2024a).

6 Parallels across domains

A key strength of this style of computational analysis is its ability to abstract away from differences in substance and representation, revealing deep parallels between linguistic phenomena across domains (Graf 2022a). Syntactic agreement bears several striking similarities to long-distance harmony in phonology, and this is readily apparent in the TSL analysis, where in each case all participants and blockers appear on the tier, and other elements are omitted. This derives the locality profile we observe, whereby a single blocker of the relevant type is always enough to disrupt agreement/harmony (McMullin & Hansson 2016). Upon initial inspection, the parameters of variation are also extremely similar, as phonological harmony also shows semi-arbitrary variation in visibility, directionality, and presence or absences of chaining. This is summarized in Table 2.

There are also important differences between these phenomena. For example, opaque vowels do not merely block harmony, but restart it by transmitting their own features, which contrasts with both coronal blocking of sibilant harmony and horizons for agreement. Furthermore, neither the perspective of surface phonotactics taken in Section 2 nor the perspective of a phonological map corresponds perfectly with the perspective of derivational Minimalism. Even so, the similarities are undeniable, the present analysis provides a compelling reason for why this should be so: both phenomena involve TSL-2 computations implementing feature matching/copying in their respective domains, and are therefore predicted to show similar formal variation.

7 Strengths and limitations of the model

The preceding discussion has highlighted several strengths of the current model: it is grounded in simple, well-understood mathematics; it is applicable to a wide range of empirical phenomena; it is predictive in that the natural parameters of the TSL-2 grammar for agreement correspond to actual variation in agreement configurations; finally, it facilitates comparisons between agreement and other

Parameter	**ϕ-agreement**	**Vowel harmony**
Participants	Probe and most DPs	Most vowels
Invisible	Non-DPs, some DPs	Consonants, some vowels
Blockers	Finite C, some DPs	Some vowels
Directionality	Downward/upward	Progressive/regressive
Chaining	Concord/no concord	Spreading/“icy targets”

Table 2: Some parallels between ϕ-agreement and vowel harmony.

long-distance linguistic dependencies.

What I have not emphasized so far is that the system is also highly restrictive. In particular, it cannot implement arbitrary global constraints involving boolean logic, nor can it count constraint violations. Graf (2022b) provides these examples of linguistically preposterous island constraints which are ruled out, among others:

(17) Non-TSL-2 and unattested island types:
  a. Gang-up islands: A mover can escape n islands, but not $n + 1$.
  b. Configurational islands: XP is an island iff it is inside an embedded clause.
  c. Cowardly islands: XP is an island iff there are at least n XPs in the same clause.
  d. Discerning islands: XP is an island only for movers that contain a PP.

We can construct similarly ludicrous agreement patterns which are not TSL-2:

(18) Non-TSL-2 and unattested agreement restrictions:
  a. Gang-up horizons: A probe can see past n horizons, but not $n + 1$.
  b. Configurational direction: T probes down if it is in a matrix clause, and up if it in an embedded clause.
  c. Cowardly goals: A DP is visible iff there are at least n DPs in the same clause.
  d. Discerning goals: A DP can transmit its features iff it contains a PP.

So while TSL-2 seems to be capable of quite a lot, there is far more that it cannot do, and broadly speaking, what it can do aligns with what we in fact observe. At the same time, there are a number of outstanding puzzles, particularly for the path-based approach described here. For example, we know that certain restrictions on movement, such as the requirement that every landing site have a mover, cannot be fully implemented with c-strings, as discussed by Graf & Shafiei (2019). In addition, it is unclear how best to handle apparent violations of c-command (i.e. sub-command) and exceptions to the complement spine generalization (e.g. left branch extraction). Some syntactic constraints are straightforwardly handled with a tree-based analog of TSL (Graf 2022b), but that system cannot enforce command-based constraints. It would appear that some combination of the two systems is necessary.

Additionally, the present model on its own cannot answer certain questions involving the substance of linguistic elements, which are also of considerable interest to theoreticians. We discussed the case-sensitivity hierarchy earlier, but a more basic question is why case matters for ϕ-agreement at all. Similarly, why do probes seem to look downward more often than upward? But perhaps most pressing is the acquisition question: how do children identify the visible elements and constraints for each dependency? Certain formal learnability results such as those in Lambert *et al.* (2021) may help inform the construction of realistic acquisition models, but we will need to make of other aspects of the acquisition theory in order to answer these questions definitively; see (Hanson 2024b) for discussion.

8 Conclusion

We have seen that many well-known agreement patterns are TSL-2 when analyzed in terms of command strings. If we vary the tier projection and constraints slightly, we can account for a wide range of variation across languages, and the exact range of variation is highly similar to other long-distance linguistic phenomena. The principal parameters of the model, namely visibility, directionality, and chaining, are all realized within the domain of ϕ-agreement; we need not even look to other types of agreement in order to find them.

As previous computational analyses of movement and case, we have only just scratched the surface. At this point, we might ask whether there exist any agreement dependencies which are not TSL-2 under any reasonable analysis. It is also not yet known just how far we can take the parallels with other long-distance linguistic dependencies. Recent research in Minimalist theory given much attention to the relation between agreement and case assignment; perhaps the TSL perspective can shed light on this issue.

In summary, the TSL-2 analysis of agreement provides independent support to conclusions reached in other ways, such as the fact that visibility is parameterized, while revealing further insights that might otherwise not be obvious, such as the close parallel between agreement and phonological harmony. The results described here therefore lend further support to the claim that language achieves the dazzling variety seen on the surface using just a few very simple types of computations.

Acknowledgments

This work was partly supported by NSF Grant BCS-1845344 and by the Institute for Advanced Computational Science at Stony Brook University. Thanks also to Thomas Graf, Sandhya Sundaresan, Tom McFadden, John Bailyn, and audiences at the NYU Linguistics Brown Bag, CLS 60, and the Workshop on Myopia in Grammar at Leipzig University for their comments and feedback.

References

Adger, D. 2003. *Core syntax: A Minimalist approach*. Oxford: Oxford University Press.

Béjar, S., & M. Rezac. 2009. Cyclic agree. *Linguistic Inquiry* 40. 35–73.

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

Bobaljik, J. 2008. Where's phi? Agreement as a postsyntactic operation. In *Phi Theory*, ed. by D. Harbour, D. Adger, & S. Béjar, 295–328. Oxford: Oxford University Press.

Burness, P. A., K. J. McMullin, J. Chandlee, P. Burness, & K. McMullin. 2021. Long-distance phonological processes as tier-based strictly local functions. *Glossa: a journal of general linguistics* 6.

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

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

Coon, J., & C. Parker. 2019. Case interactions in syntax. In *Oxford Research Encyclopedia of Linguistics*. Oxford University Press.

De Santo, A., & T. Graf. 2019. Structure sensitive tier projection: Applications and formal properties. In *Formal Grammar*, ed. by R. Bernardi, G. Kobele, & S. Pogodalla, 35–50, Berlin, Heidelberg. Springer.

Deal, A. R. 2015. Interaction and satisfaction in ϕ-agreement. In *Proceedings of NELS 45*, 179–192, Amherst, MA. GLSA Publications.

Goldsmith, J. 1976. *Autosegmental phonology*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Graf, T. 2019. Monotonicity as an effective theory of morphosyntactic variation. *Journal of Language Modelling* 7. 3–47.

Graf, T. 2022a. Subregular linguistics: bridging theoretical linguistics and formal grammar. *Theoretical Linguistics* 48. 145–184.

Graf, T. 2022b. Typological implications of tier-based strictly local movement. In *Proceedings of the Society for Computation in Linguistics 2022*, 184–193, Amherst, MA. University of Massachusetts Amherst.

Graf, T., & A. De Santo. 2019. Sensing tree automata as a model of syntactic dependencies. In *Proceedings of the 16th Meeting on the Mathematics of Language*, 12–26, Toronto, Canada. Association for Computational Linguistics.

Graf, T., & N. Shafiei. 2019. C-command dependencies as TSL string constraints. In *Proceedings of the Society for Computation in Linguistics 2019*, 205–215. University of Massachusetts Amherst Libraries.

Hanson, K. 2023. A TSL analysis of Japanese case. In *Proceedings of the Society for Computation in Linguistics 2023*, 15–24, Amherst, MA. University of Massachusetts Amherst Libraries.

Hanson, K., 2024a. Tier-based strict locality and the typology of agreement. To appear in *Journal of Language Modeling*.

Hanson, K. 2024b. Tiers, paths, and syntactic locality: The view from learning. In *Proceedings of the Society for Computation in Linguistics 2024*, 107–116, Amherst, MA. University of Massachusetts Amherst Libraries.

Heinz, J., C. Rawal, & H. G. Tanner. 2011. Tier-based strictly local constraints for phonology. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human language technologies*, 58–64. Association for Computational Linguistics.

Keine, S. 2019. Selective opacity. *Linguistic Inquiry* 50. 13–62.

Lambert, D. 2023. Relativized adjacency. *Journal of Logic, Language and Information* 32. 707–731.

Lambert, D., J. Rawski, & J. Heinz. 2021. Typology emerges from simplicity in representations and learning. *Journal of Language Modelling* 9. 151–194.

Lambert, D., & J. Rogers. 2020. Tier-based strictly local stringsets: Perspectives from model and automata theory. In *Proceedings of the Society for Computation in Linguistics 2020*, 159–166, Amherst, MA. University of Massachusetts Amherst.

Mahajan, A. K. 1990. *The A/A-bar distinction and movement theory*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

McMullin, K. 2016. *Tier-based locality in long-distance phonotactics: learnability and typology*. Vancouver, BC: University of British Columbia dissertation.

McMullin, K., & G. Ó. Hansson. 2016. Long-distance phonotactics as tier-based strictly 2-local languages. In *Proceedings of the Annual Meetings on Phonology*, volume 2, Washington, DC. Linguistic Society of America.

Pesetsky, D., & E. Torrego. 2007. The syntax of valuation and the interpretability of features. In *Phrasal and Clausal Architecture: Syntactic Derivation and Interpretation*, ed. by S. Karimi, V. Samiian, & W. K. Wilkins, 262–294. Amsterdam: Benjamins.

Preminger, O. 2014. *Agreement and its failures*. Cambridge, MA: MIT Press.

Rizzi, L. 1990. *Relativized minimality*. Cambridge, MA: MIT Press.

Rizzi, L. 2001. Relativized minimality effects. In *The Handbook of Contemporary Syntactic Theory*, chapter 4, 89–110. John Wiley & Sons, Ltd.

Van Urk, C. 2015. *A Uniform Syntax for Phrasal Movement: A Dinka Bor Case Study*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Vu, M. H., N. Shafiei, & T. Graf. 2019. Case assignment in TSL syntax: A case study. In *Proceedings of the Society for Computation in Linguistics 2019*, 267–276, Amherst, MA. University of Massachusetts Amherst Libraries.