Upward Agreement and Syntactic Counterfeeding in Lubukusu

Kenneth Hanson*

Abstract. What are the constraints on agreement, and how does it interact with other syntactic operations? Recent studies suggest that syntactic dependencies are TIER-BASED STRICTLY LOCAL (TSL) over Minimalist Grammar derivation trees, a computationally restrictive model which closely fits many aspects of the formal typology (Graf 2022, Hanson 2025). Using this framework, I provide an analysis of upward complementizer agreement in Lubukusu (Diercks 2013), and show that it correctly predicts lack of agreement with hyperraised subjects without any additional assumptions. I argue further that the phenomenon should be understood as a kind of SYNTACTIC COUNTERFEEDING, which together with the existence of feeding, bleeding, and counterbleeding, suggests the need for a syntactic framework which can handle variable operation ordering.

Keywords. complementizer agreement; directionality; syntactic counterfeeding; tier-based strictly local languages, Minimalist Grammars

1. Introduction. The issue of directionality in agreement has received much attention in recent syntactic theory (Pesetsky & Torrego 2007, Baker 2008, Zeijlstra 2012, a.o.). From a computational perspective which models linguistic dependencies as TIER-BASED STRICTLY LOCAL (TSL) computations over strings and trees (Graf 2022), the direction of syntactic operations is predicted to vary parametrically, just as long-distance phonological harmony may be progressive or regressive. This appears to be bourne out in the realm of complementizer agreement, which descriptively may target the embedded subject or the subject of the containing clause depending on the language. This is illustrated by the contrast between the West Flemish (Germanic) and Lubukusu (Bantu) examples below. Note: C1/C2 indicate noun class.

(1) Downward complementizer agreement in West Flemish
Kpeinzen da-j [$_{CP}$ (gie) morgen goat].
I.think that-**you** **(you)** tomorrow go
'I think that you'll go tomorrow.' (Diercks 2013, ex. 3)

(2) Upward complementizer agreement in Lubukusu
Ba-ba-ndu ba-bolela Alfredi [$_{CP}$ ba-li *pro* a-kha-khile].
C2-**C2**-people C2-said C1.Alfred **C2**-that *pro* C1-FUT-conquer
'The people told Alfred that he will win.' (Diercks 2013, ex. 1a)

In this paper, I analyze the Lubukusu agreement pattern as described by Diercks (2013) using the framework in Hanson (2025), which applies the TSL perspective to syntactic agreement. Specifically, I provide an analysis which takes upward agreement at face value, in contrast to Diercks' analysis of "indirect agreement" via an empty operator at the embedded clause edge. Of particular interest is the lack of complementizer agreement with hyperraised subjects; only a non-agreeing complementizer is allowed, as shown in (3).

* This work was partly supported by the National Science Foundation under Grant No. BCS-1845344, and by the Institute for Advanced Computational Science at Stony Brook University. The analysis of upward agreement originated in Hanson (2025), but the discussion of syntactic counterfeeding is new. Authors: Kenneth Hanson, Stony Brook University (mail@kennethhanson.net).

(3) Agreeing complementizer incompatible with hyperraising
Sammy a-lolekhana mbo (*a-li) a-likho a-lwala.
c1.Sammy c1-appears that (***c1**-that) c1.PROG c1-be.sick
'Sammy appears to be sick.' (lit. 'Sammy seems that is sick.') (Diercks 2013, ex. 75)

This follows immediately from the present analysis, which uses Minimalist Grammar (MG) DERIVATION TREES for the syntactic representation: the complementizer looks upward for agreement but in the derivation tree the hyperraised subject can be seen only in its base position in the embedded clause. In comparison, Diercks' analysis operates within mainstream Minimalism and requires further assumptions, as the hyperraised subject occupies the usual subject position and should therefore be able to bind the operator.

Lack of agreement in derived positions can be seen as a kind of COUNTERFEEDING IN THE SYNTACTIC DERIVATION, representing yet another formal parallel with phonology, where such phenomena are well known. However, unlike variable directionality, it does not derive solely from the TSL nature of syntactic computations, but emerges only in combination with other assumptions, in this case the derivation tree representation. Such counterfeeding is a basic prediction of the present framework, but to my knowledge this is the first time that a concrete example has been identified as such. Looking more broadly, we find that the other classical rule ordering effects—feeding, bleeding, and counterbleeding—also have direct analogs in syntax. For example, differential argument marking can be fed by movement, and case marking can bleed agreement, while LF reconstruction can be seen as counterbleeding of binding by movement. However, contemporary generative frameworks are not usually designed with parametrization of operation order in mind, but instead prefer certain orders over others. In mainstream Minimalism, for instance, movement is expected to feed subsequent syntactic operations by default, whereas the present model predicts counterfeeding. This study therefore suggests the need for a more flexible system, perhaps closer in some ways to classic Transformational Grammar.

The remainder of this paper proceeds as follows. First, I introduce the idea of a TSL computation and how it can be applied to syntactic agreement, adopting the model from Hanson (2025). Next, I show how the model straightforwardly accounts for the Lubukusu complementizer agreement data, inclusive of the hyperraising construction. Finally, we consider some examples of other types of interactions between syntactic operations, and how they might be accommodated in a modern derivational theory of syntax.

2. TSL syntax. In this section, I briefly introduce TSL string constraints, the syntactic framework of MG derivation trees, and a TSL model of movement and agreement which uses path-based constraints.

2.1. TSL COMPUTATIONS. TSL computations encapsulate a notion of relativized locality which meshes well with many linguistic phenomena, including surface phonotactics (Heinz et al. 2011, McMullin 2016), the UR to SR map (Burness et al. 2021), syntactic derivations (Graf 2022, Hanson 2025), and more. In general, we delimit a subset of SALIENT ELEMENTS (segments, syntactic heads, etc.) which are relevant to a given dependency, determined only by their label. The non-salient elements are ignored completely and the remainder treated as adjacent, forming a TIER PROJECTION. All constraints must be stated within a FINITE WINDOW on the tier.

A classic example is sibilant harmony with coronal blocking (based on Slovenian), as described by McMullin & Hansson (2016). In such a language, sibilants such as [s] and [ʃ] must harmonize on the surface except when a coronal obstruent such as [t] intervenes. So, a word like 'sakaʃa' would be ruled out while 'sataʃa' is permitted. The set of salient elements in this example is {s, ʃ, t}, and the local constraints on the tier ban mismatched sibilants: {*sʃ, *ʃs}. Together, these two components constitute a TSL grammar. This is summarized in Figure 1.

All elements: {s, ʃ, t, k, a}
Salient elements: {s, ʃ, t}
Tier constraints: {*sʃ, *ʃs}

	Word	Tier Proj.
✓	sakasa	ss
✓	ʃakaʃa	ʃʃ
✗	sakaʃa	sʃ
✓	sataʃa	stʃ

Figure 1. TSL grammar (left) with example words and tier projections (right) for sibilant harmony with blocking (McMullin & Hansson 2016)

Notice that when disharmony is allowed, the sibilants are not actually adjacent on the tier, and so no constraints are violated. Also, it is worth emphasizing that the tier projection itself is just a convenient visualization, and need not be literally constructed. The key insight is that we can treat long-distance dependencies in language as if they are local, modulo the irrelevant items, in a way that is mathematically precise and restricted in expressivity. Interestingly, when we do this, we find that the constraint window we need contains no more than two elements. Such computations, and the class of the patterns they produce, are called TIER-BASED STRICTLY 2-LOCAL, or TSL-2. This seems to be true not only in phonology (McMullin 2016), but also syntax (Graf 2022, Hanson 2025), giving us a unified model with a natural set of parameters: what goes on the tier for each dependency, and what pairs of adjacent elements are allowed?

2.2. MG DERIVATION TREES. A derivation tree is a static representation of the process by which the phrase structure tree is built, in this case according to the rules of a Minimalist Grammar, or MG (Stabler 1997, 2011). MGs, in turn, are a particular formalization of ideas from the Minimalist Program (Chomsky 1995) whose mathematical properties are well-understood. Both the MG formalism and the derivation tree format have several variations in the literature; the version used here is based on Graf & Kostyszyn (2021).

Let us start with a concrete example. The MG derivation tree for the sentence *There seem to be some problems* is shown in Figure 2 along with the corresponding phrase structure tree. Expletive *there* is assumed to originate in the *v*P and raise to Spec-TP (the "low origin" analysis) while verbal agreement targets the logical subject instead; while not crucial to the approach, this analysis allows an easy illustration of the parametrization of visibility.

I wish to highlight several key aspects of the representation:

1. Every node is a LEXICAL ITEM. There are no phrasal nodes, though each subtree represents a phrase as usual, making these trees a kind of dependency tree.
2. The daughters of a node are its ARGUMENTS arranged in reverse order of selection. The order among sisters corresponds to asymmetric c-command in the phrase structure tree.
3. Every node appears only in its BASE POSITION, that is, the position of external Merge. Intermediate and final positions are not represented.
4. Long-distance dependencies such as movement and agreement are indicated with FEATURES (or diacritics), as explained below. I have added dashed lines connecting dependent items as a visual aid, but they are not part of the representation.

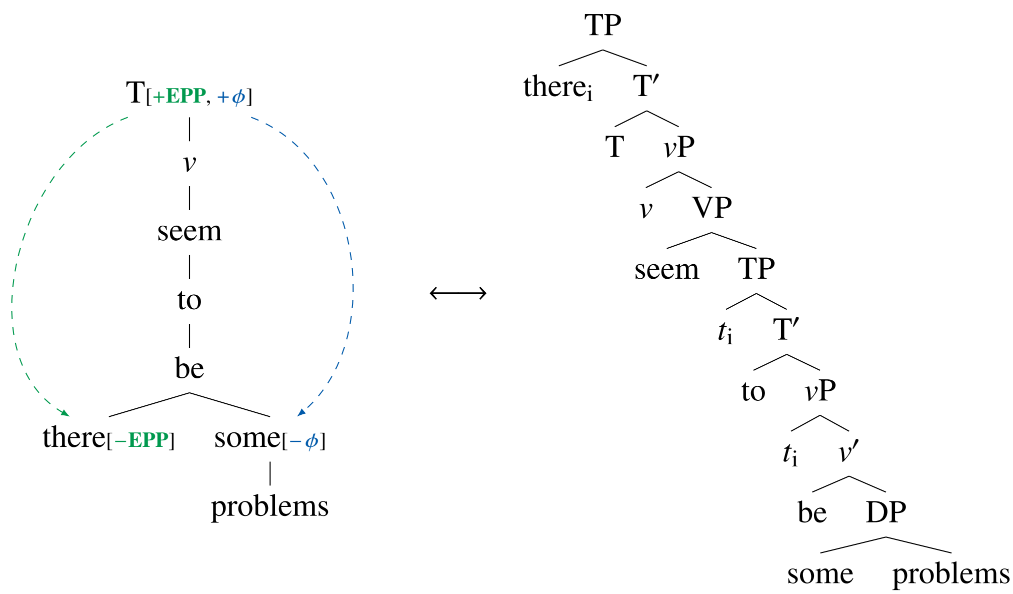


Figure 2. Derivation tree and phrase structure tree

Properties #1 and #2 produce an unusually concise representation which is convenient for analysis. While there are differences between this and other types of derivation trees, they are not relevant to the issues in this paper. #3 is important as this is the source of counterfeeding in the present analysis, as described in Section 3.1. #4 is also crucial, since the features are what allow us to describe movement/agreement configurations in formal terms, so we focus our attention on them now.

MG features (or diacritics) are declarations of what happens *in the present derivation*. In standard MGs, they come in two varieties: Merge features, which control selection, and Move features, which control movement. Since we are not interested in the details of selection, I omit all Merge features. For Move, positive features mark landing sites, and negative features mark movers. In the current example, the fact that *there* bears −EPP and T bears +EPP means that *there* will move to Spec-TP in the finite clause. (Successive cyclic movement is not conventionally represented in MG, which equates to the view that it is not feature-driven.) Here, I have intentionally chosen the rather meaningless label 'EPP' for this type of movement since for our purposes it does not matter what actually drives movement of the subject.

To handle agreement, we extend the above notation as in Hanson (2025). Adopting the terminology of Chomsky (2000), positive features indicate PROBES, which copy a feature from some other node, and negative features represent GOALS, which provide a value. Here, I have chosen ϕ to indicate verbal agreement for person/number/gender features, so T bears $+\phi$ and the D head it agrees with bears $-\phi$. To be clear, this does not mean that other D heads do not have ϕ-features in the usual sense, only that they have not transmitted their features in an Agree operation.

To briefly summarize, we have a concise representation which tells us everything we need to know about which elements of a syntactic structure are related via long-distance dependencies. Next, we aim to specify just the set of licit movement/agreement configurations in a given language using a TSL grammar.

2.3. TSL AGREEMENT. Following Hanson (2025), we assume that long-distance dependencies follow a path through the derivation tree corresponding to the COMPLEMENT SPINE, that is, they

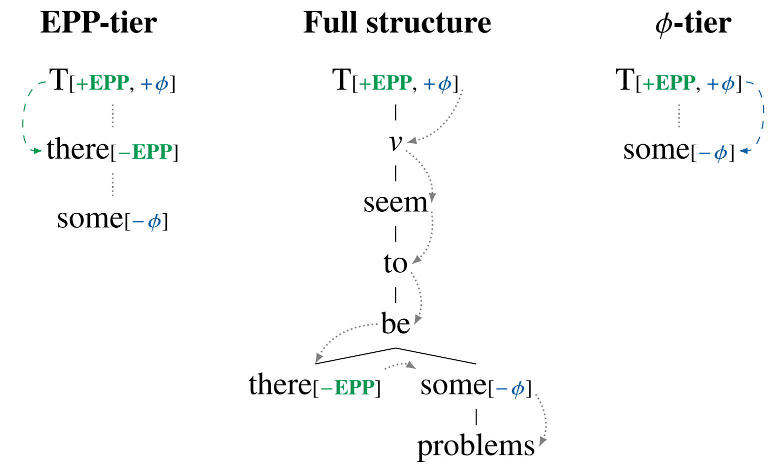


Figure 3. MG dependency tree for *There seem to be some problems*, with path tiers for EPP-movement and ϕ-ageement

can see the root of each specifier or adjunct, but not inside.[1] (In addition to the main complement spine, each complex specifier and adjunct has its own spine, which we ignore for simplicity.) Along this path, parents are ordered before children, and left sisters before right sisters, such that nodes which were first merged later in the derivation are ordered earlier. Graf & Shafiei (2019) call this ordering relation DERIVATIONAL COMMAND, or d-command, since it directly encodes derivational prominence as defined by the order of external Merge. This is depicted in the center of Figure 3 with the complement spine indicated using dotted arrows.

A path, of course, is just a string, which means that we can formulate a TSL grammar exactly as we did previously. In general, we need one grammar per dependency, as it is well established that visibility conditions vary (Rizzi 1990, Bobaljik 2008, Keine 2019, a.o.). Although the approach was developed specifically to model agreement, it works just as well for the aspects of movement which are relevant here. For English, the visible elements for verbal agreement include finite T and all lexical D heads, while the tier for EPP-movement additionally includes existential *there*, which can move even though it is assumed not to agree.[2] These tier projections for the current example are shown on the left and right sides of Figure 3, and are oriented vertically to emphasize that they correspond to derivational prominence. In each case, the dependent items are adjacent and in the correct order.

For both of these dependencies, the constraints are identical: the landing site for EPP movement must appear immediately before (=above) the mover on the tier; likewise, the probe for agreement must appear immediately before its goal. But we could just as easily imagine a grammar in which the constraints are slightly different. In particular, the required order of the dependent elements could be reversed. Though theoretically controversial, this is completely unproblematic from the present perspective, and for agreement at least, appears to be attested. This brings us to the complementizer agreement data.

3. Lubukusu complementizer agreement. Now, we apply the model developed in the previous section to the Lubukusu complementizer agreement data. I defer to Diercks (2013) regarding the

[1] This is probably too strong a constraint as stated, but can plausibly be weakened slightly without allowing rampant violations of the left branch and adjunct island conditions. See Graf & De Santo (2019) for details.
[2] The system can also handle optional default singular agreement. See Hanson (2025).

empirical generalizations, diverging only on the interpretation of the agreement dependency. For brevity, only the key illustrative examples from the introduction are treated.

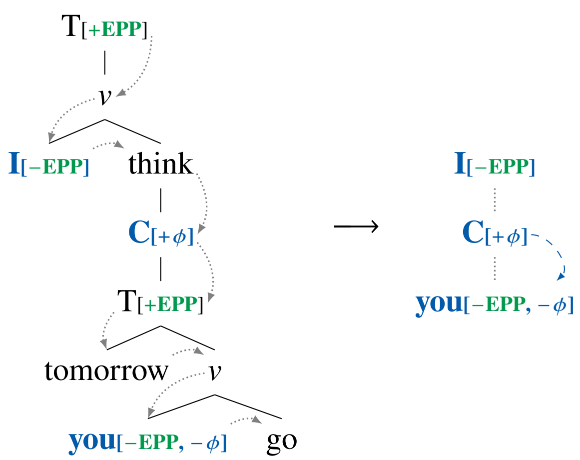
3.1. The direction of complementizer agreement. Recall that in West Flemish the complementizer agrees with the embedded subject, while in Lubukusu it agrees with the subject of the containing clause. The data showing this contrast is repeated below.

(4) Kpeinzen da-j [$_{CP}$ (gie) morgen goat].
I.think that-**you** **(you)** tomorrow go
'I think that you'll go tomorrow.' (West Flemish)

(5) Ba-ba-ndu ba-bolela Alfredi [$_{CP}$ ba-li *pro* a-kha-khile].
c2-c2-people c2-said c1.Alfred **c2**-that *pro* c1-FUT-conquer
'The people told Alfred that he will win.' (Lubukusu)

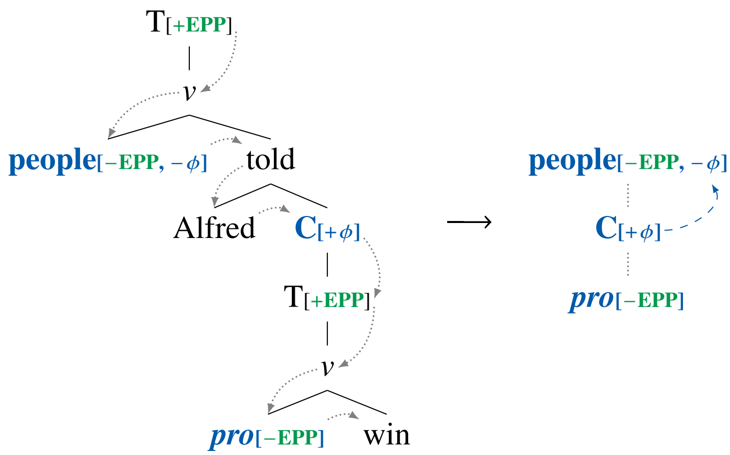
For brevity, we abstract away from verbal agreement, focusing just on complementizer agreement. For West Flemish, we have a tier containing all C and D heads, and C agrees downward. This is essentially the same pattern as verbal agreement in English. Assuming a fairly standard structural analysis, the derivation tree and the complementizer agreement tier for (4) are shown below. Note that I assume that *tomorrow* is an adverb/PP, and hence invisible.

(6) West Flemish: C agrees downward



For Lubukusu, there are two key differences. First, agreement is upward, which means that C agrees with the item which *precedes* it on the tier. Second, as shown by Diercks (2013), agreement is strictly subject oriented, skipping over the the intervening object. While this appears at first glance to be a locality violation, there is an extremely straightforward solution: not all D heads appear on the tier, only those that bear the −EPP diacritic. This is illustrated in (7) below. For simplicity, I treat *people* and *Alfred* as simplex D heads.

(7) Lubukusu: only the moved subject is visible, C agrees upward



This analysis works because I follow Diercks in assuming that the Lubukusu subject does in fact move to Spec-TP; the difference is that while for Diercks it is the position itself that is crucial, in the derivation tree it is the associated movement feature that we have direct access to. This is also how case-sensitive agreement is handled in a TSL analysis: we do not include every D head on the tier, only those that bear an appropriate case. In this way, we can handle relativized minimality in its classic form as well as more fine-grained visibility conditions which have been the focus of much recent research.

3.2. No upward agreement with hyperraised subjects. Now, let us tackle the hyperraising data. Diercks shows that the promoted subject of a passive construction can control complementizer agreement just like a canonical subject. However, a subject which raised to Spec-TP from within the embedded CP cannot. Since such clauses are finite, I refer to this simply as the hyperraising construction.[3] The key illustrative data point from the introduction is repeated below.

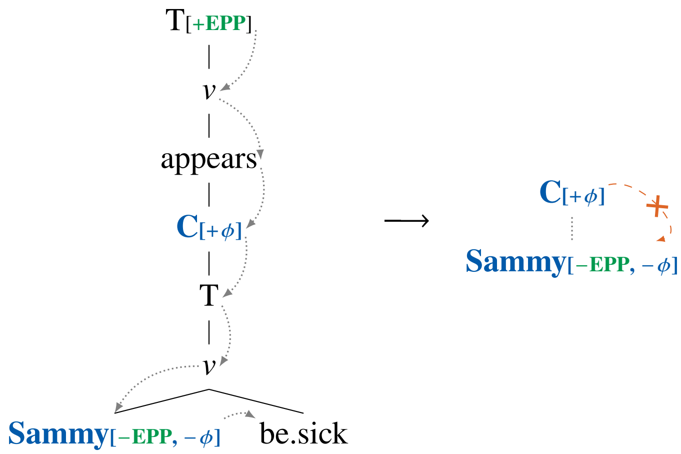
(8) Sammy a-lolekhana mbo (*a-li) a-likho a-lwala.
c1.Sammy c1-appears that (***c1**-that) c1.prog c1-be.sick
'Sammy appears to be sick.' (lit. 'Sammy seems that is sick.')

The present analysis provides a straightforward explanation for this fact: C looks upward for its goal in the derivation tree, but the hyperraised subject is found only in its base position, below C. In formal terms, the goal must precede the probe on the tier, but in such a structure the goal instead follows the probe, which is not allowed. The derivation tree and complementizer agreement tier for the preceding example are shown in (9). In contrast, finite T does agree downward, and as such it is no coincidence that it agrees with the hyperraised subject as usual (not depicted).

It is important to understand that this state of affairs derives specifically from the combination of directionality (as encoded by the TSL grammar) and representation (the derivation tree). If the same TSL grammar was applied to a surface-oriented representation, we would incorrectly predict that complementizer agreement should be possible, as the subject would appear in an agreeing position. This is an excellent example of how we must consider the properties of both the individual components of the formal system as well as their interactions.

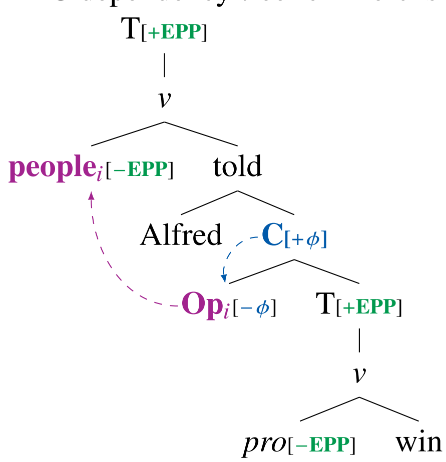
[3] It is also possible for a DP to undergo hyperraising to object position, then to move to matrix subject position via passivization. In this case, complementizer agreement is also impossible, just as in the simple hyperraising case.

(9) Hyperraised subject is below C and cannot agree



Conversely, if we were to adapt Diercks' analysis into the present framework, the extra assumptions he requires to handle the hyperraising data could be eliminated. Let us briefly see why this is. In Diercks' "indirect agreement" analysis, C does not agree with the matrix subject, but with a covert operator in Spec-CP, which is bound by the subject. Thus, agreement is downward in present terms, since the specifier is a child of the C head, but binding is upward in that the features of the operator (or at least its index) are determined from above. This is schematized below, with indices and an arrow added to highlight the binding dependency.

(10) MG dependency tree for Diercks' analysis of (5)



Complementizer agreement in the hyperraising construction is therefore ruled out for the same reason as before, since the hyperraised subject is not in an appropriate position to bind the operator. The only difference is that the constraints on binding are violated rather than those on agreement. Although I do not find Diercks' arguments for the indirect analysis convincing (see Hanson 2025 for explanation), if one were to adopt this variant of his analysis then the formal points I have made regarding agreement are simply transferred to the domain of binding.[4]

Regardless of the specific analysis, this particular example is insightful because it cannot easily be interpreted in terms of locality effects. For example, if agreement on matrix T failed

[4] It goes without saying that this discussion has no bearing on other potential cases of upward agreement. For example, if case is assigned first to the head of DP, as is commonly assumed, then the head noun and other concordial elements agree upward after case is assigned—another paradox of timing.

in cases of hyperraising (contrary to fact), we might posit that agreement and movement apply downward simultaneously, and that agreement is subject to stricter locality constraints than movement; in this case, no reference to pre/post-movement position is required. But here distance is not the issue, rather, there is a contradiction in that C receives its features from above while the hyperraised DP originates below. The system of TSL grammars over MG derivation trees predicts that agreement should respect the pre-movement configuration, at least sometimes, and this appears to be bourne out.

4. Operation ordering in syntax. To briefly review what we have just seen, complementizer agreement in Lubukusu can be understood as involving counterfeeding of agreement by hyperraising: the hyperraised subject occupies the usual subject position, yet fails to agree. In addition to fulfilling a basic prediction of the present framework, the broader significance of this data is that it completes the typology of classic rule ordering effects, of which the other three types are much better known within syntax.

Let us briefly go through a few examples. First, we have FEEDING, in which the application of one operation creates the context for another to apply. A classic example is differential object marking, which in many (but not all) instances can be understood as a direct effect of the object moving into a position which is sufficiently close to the subject, as in the analysis of accusative case in Sakha by Baker & Vinokurova (2010). Thus, movement can feed case assignment.

Next is BLEEDING, in which the application of one operation *removes* the context for another to apply. This can be found in languages with case-sensitive agreement, such as Hindi (Mahajan 1990, Bhatt 2005). In this language, only nominative (=unmarked) DPs can agree. Normally, the subject takes precedence for verbal agreement, but if it bears some other case, such as ergative or dative, then a more distant DP, such as a matrix or embedded object, can become eligible. If those other DPs are also case-marked, then agreement may fail entirely. Here, case assignment bleeds agreement.

Then, we have the opposite state of affairs, in which the application of one operation *fails* to affect the application of another, which is known as COUNTERFEEDING. This is represented by Lubukusu, in which hyperraising fails to feed complementizer agreement. In a rule ordering system, this is derived by having agreement apply before movement; in the present system, it derives from the fact that agreement follows the configuration in the derivation tree rather than the phrase structure tree.

Finally, COUNTERBLEEDING is the situation in which one operation removes the context for another, yet the latter operation applies anyway. This can be found in the realm of binding as a type of reconstruction effect. For example, *wh*-movement of a pronoun does not ameliorate a Principle B violation in the pre-movement position. In GB parlance, we would say that the pronoun is reconstructed in its base position at LF; under a Minimalist view, we might instead say that it never vacated that position in the first place.

The typology represented by the preceding examples is summarized in Table 1. While I could have included additional examples—movement of a dative DP bleeds visibility in Icelandic (feeding agreement with a more distant nominative DP), scrambling preserves case as determined by base argument position in Japanese, etc.—these suffice for illustration. All of the classical ordering effects appear to be attested, even if not every ordering of every pair of operations can be found. Furthermore, it is not obvious that any of these effects can be seen as more basic, and the others as exceptional.

Operations	Ordering phenomenon	Ex. Language
Mvmt. + Case	Object-shift **feeds** accusative marking	Sakha
Case + Agreement	Case marking **bleeds** V-agreement	Hindi
Mvmt. + Agreement	Hyperraising **counterfeeds** C-agreement	Lubukusu
Mvmt. + Binding	*wh*-Movement **counterbleeds** Principle B	English

Table 1. Examples of operation ordering in syntax

Yet, as I mentioned in the introduction, contemporary generative frameworks are not set up in such a way that all such interactions are equally easy to derive. In mainstream Minimalism, Movement feeds subsequent operations targeting the post-movement position by default. The base position is subject to counterbleeding under the copy theory of movement, while other interactions may potentially be derived under specific assumptions about the timing of copy deletion. Additionally, it is reasonable to suppose that when a single functional head is the trigger for several operations, their order might be parameterized. However, it is tricky to alter the timing of operations which are triggered by different heads.

For other frameworks, the state of affairs can be completely different. As mentioned above, interactions with movement in the present system involve counterfeeding and counterbleeding by default, as only the base position is directly visible. Once again, it is often possible to derive other operation orders with extra work. For example, differential object marking can be handled by paying attention to whether the object DP bears the relevant movement feature. But if we really need to know exactly which positions a phrase moves to, something more is needed.

If the above characterization of the empirical landscape is on the right track, then what we really want is a system which can control the application of syntactic operations in an extremely general manner, much like classic Transformational Grammar but in the context of a continuous derivation. At the same time, we do not want to lose the restrictiveness of the TSL model. This might be achieved in the form of a TSL tree-to-tree function which maps the derivation tree to the corresponding phrase structure tree. Graf (2023) shows how both copies and traces can be handled in such a system, but much additional work is needed to determine whether the approach can be extended to case, agreement, and their interactions with movement. At present, there is a tension between the need for operation ordering, strict cyclicity, and computational restrictiveness which has yet to be resolved.

5. Conclusion. The TSL model predicts strong formal parallels between phonology, morphology, and syntax, as well as similar types of variation. This study finds evidence for variation in directionality in agreement, as well as evidence of counterfeeding in the syntactic derivation. It also exposes a key difference between the perspective of mainstream Minimalism and the derivation tree-oriented perspective which is familiar to practitioners of MG. While the derivation tree perspective provides an elegant explanation of the Lubukusu data (and perhaps other counterfeeding and counterbleeding interactions), it may struggle with interactions which are more surface-oriented, which are also extremely common. This is roughly the opposite of the situation in mainstream Minimalism. To me, this suggests that neither approach is quite right, but an appropriately flexible fusion might result in a more adequate model. This study thus highlights the need for a more systematic empirical and formal study of operation ordering effects in syntax.

References

Baker, Mark C. 2008. *The syntax of agreement and concord* (Cambridge Studies in Linguistics 115). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511619830.

Baker, Mark C. & Nadya Vinokurova. 2010. Two modalities of case assignment: case in Sakha. *Natural Language & Linguistic Theory* 28(3). 593–642. https://doi.org/10.1007/s11049-010-9105-1.

Bhatt, Rajesh. 2005. Long distance agreement in Hindi-Urdu. *Natural Language & Linguistic Theory* 23. 757–807. https://doi.org/10.1007/s11049-004-4136-0.

Bobaljik, Jonathan. 2008. Where's phi? Agreement as a postsyntactic operation. In Daniel Harbour, David Adger & Susana Béjar (eds.), *Phi theory*, 295–328. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780199213764.003.0010.

Burness, Phillip Alexander et al. 2021. Long-distance phonological processes as tier-based strictly local functions. *Glossa: a journal of general linguistics* 6(1). https://doi.org/10.16995/glossa.5780.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/9780262527347.001.0001.

Chomsky, Noam. 2000. Minimalist inquiries: the framework. In R. Martin, D. Michaels & J. Uriagereka (eds.), *Step by step: Essays on Minimalist Syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press. https://mitpress.mit.edu/9780262516839/step-by-step/.

Diercks, Michael. 2013. Indirect agree in Lubukusu complementizer agreement. *Natural Language & Linguistic Theory* 31(2). 357–407. https://doi.org/10.1007/s11049-013-9187-7.

Graf, Thomas. 2022. Subregular linguistics: bridging theoretical linguistics and formal grammar. *Theoretical Linguistics* 48(3–4). 145–184. https://doi.org/10.1515/tl-2022-2037.

Graf, Thomas. 2023. Subregular tree transductions, movement, copies, traces, and the ban on improper movement. In Tim Hunter & Brandon Prickett (eds.), *Proceedings of the Society for Computation in Linguistics (SCiL) 2023*, 289–299. Amherst, MA: University of Massachusetts Amherst. https://doi.org/10.7275/TK1N-Q855.

Graf, Thomas & Aniello De Santo. 2019. Sensing tree automata as a model of syntactic dependencies. In Philippe de Groote & Frank Drewes an Gerald Penn (eds.), *Proceedings of the 16th Meeting on the Mathematics of Language (MoL 16)*, 12–26. Association for Computational Linguistics. https://doi.org/10.18653/v1/W19-5702.

Graf, Thomas & Kalina Kostyszyn. 2021. Multiple wh-movement is not special: the subregular complexity of persistent features in Minimalist Grammars. In Allyson Ettinger, Ellie Pavlick & Brandon Prickett (eds.), *Proceedings of the Society for Computation in Linguistics (SCiL) 2021*, 275–285. Amherst, MA: University of Massachusetts Amherst. https://doi.org/10.7275/2dfv-v219.

Graf, Thomas & Nazila Shafiei. 2019. C-command dependencies as TSL string constraints. In Gaja Jarosz et al. (eds.), *Proceedings of the Society for Computation in Linguistics (SCiL) 2019*, 205–215. Amherst, MA: University of Massachusetts Amherst. https://doi.org/10.7275/4rrx-x488.

Hanson, Kenneth. 2025. Tier-based strict locality and the typology of agreement. *Journal of Language Modeling*. To appear.
Heinz, Jeffrey, Chetan Rawal & Herbert G. Tanner. 2011. Tier-based strictly local constraints for phonology. In Dekang Lin, Yuji Matsumoto & Rada Mihalcea (eds.), *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, 58–64. https://aclanthology.org/P11-2011/.
Keine, Stefan. 2019. Selective opacity. *Linguistic Inquiry* 50(1). 13–62. https://doi.org/10.1162/ling_a_00299.
Mahajan, Anoop Kumar. 1990. *The A/A-bar distinction and movement theory*. Cambridge, MA: Massachusetts Institute of Technology dissertation.
McMullin, Kevin. 2016. *Tier-based locality in long-distance phonotactics: learnability and typology*. Vancouver, B.C.: University of British Columbia dissertation.
McMullin, Kevin & Gunnar Ólafur Hansson. 2016. Long-distance phonotactics as tier-based strictly 2-local languages. In Adam Albright & Michelle A. Fullwood (eds.), *Proceedings of the 2014 Annual Meeting on Phonology*. Washington, DC: Linguistic Society of America. https://doi.org/10.3765/amp.v2i0.3750.
Pesetsky, David & Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In Simin Karimi, Vida Samiian & Wendy K. Wilkins (eds.), *Phrasal and clausal architecture: Syntactic derivation and interpretation*. Amsterdam: Benjamins. https://doi.org/10.1075/la.101.14pes.
Rizzi, Luigi. 1990. *Relativized minimality* (Linguistic Inquiry Monographs 16). Cambridge, MA: MIT Press.
Stabler, Edward P. 1997. Derivational minimalism. In Christian Retoré (ed.), *Logical Aspects of Computational Linguistics: First International Conference, LACL '96, Nancy, France, September 23-25, 1996. Selected Papers* (Lecture Notes in Computer Science 1328). Berlin, Heidelberg: Springer. https://doi.org/10.1007/BFb0052152.
Stabler, Edward P. 2011. Computational perspectives on Minimalism. In Cedric Boeckx (ed.), *Oxford handbook of linguistic Minimalism*, 617–643. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199549368.013.0027.
Zeijlstra, Hedde. 2012. There is only one way to agree. *The Linguistic Review* 29(3). 491–539. https://doi.org/10.1515/tlr-2012-0017.