Two arguments for a positive vowel harmony imperative

Samuel R. Bowman

April 1, 2013

Abstract: In this paper I provide arguments from two case studies in support of Kimper's (2011) new Trigger Competition approach to vowel harmony, which is based on autosegmental linking and the positively formulated SPREAD constraint. The first argument is based on Hungarian vacillation, in which sequences of vowels that individually do not participate in backness harmony can conspire to create free variation or to participate fully. Trigger Competition's violable preference for local harmony over long-distance harmony, as well as its formalization of differing degrees of vowel transparency, make this simple to account for. The second is based on an account of Seto backness harmony, a system which has not yet been addressed in the constraint-based phonology literature, and which I claim presents a distinctly difficult case for conventional theories due to its paired neutral vowels.

1 An introduction to Trigger Competition

Trigger Competition is a novel framework that accounts for vowel harmony by proposing that any segment in a word can spread its feature values onto any other segment, but that locality, vowel quality, and other factors contribute to deciding which segment, if any, can trigger harmony on each potential target. This paper presents two pieces of evidence for a version of this system with two implementational changes.

The competition referred to by the name Trigger Competition is implemented using the positive SPREAD constraint, which assigns rewards (rather than violations) to candidates in which segments share their feature values. Factors which impact how desirable a particular instance of spreading is, like the proximity between the segments involved, are encoded as multipliers which affect the reward score assigned by the constraint: rather than stipulating additional constraints and mechanisms on top of harmony, Trigger Competition enriches the harmony imperative directly.

Kimper proposes a multiplier which scales the reward according to the quality of the vowel that triggers spreading. This multiplier is meant to capture how well cued the harmonic feature is for a given segment, with the understanding that the grammar uses harmony to suppress some of the ambiguity associated with these weakly cued segments by allowing them to jointly cue the same feature value. This innovation allows the grammar a much richer vocabulary of possible behaviors when dealing with the notorious problem of neutral vowels—those which are not systematically subject to harmonic alternation. In particular, this makes it possible for the grammar to treat vowel transparency as a scalar phenomenon rather than a binary one, allowing it to account for cases where generally transparent vowels become opaque and propagate harmony in certain environments.

The constraint is implemented within Serial Harmonic Grammar (SHG, Pater et al., 2008, Pater, 2010, Mullin, 2010), a variant of the weighted constraint system Harmonic Grammar which adds stepwise evaluation from Harmonic Serialism. A grammar in SHG makes incremental changes to a linguistic form until no more productive changes can be made. At each iteration, the grammar generates every possible single change to the input. The grammar then chooses to make the change which has the highest harmony score: the weighted sum of the change's violations of the grammar's various constraints. This repeats until there are no changes which improve upon the harmony score of the existing form. At this point, that form is chosen as the output, and optimization ends.

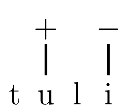
What constitutes a single change for the purposes of a derivation step is an open question, and an essentially empirical one. I follow Kimper's approach, never fully enumerated, which seems to reflect a tentative consensus that changing contrastive feature values possibly by way of linking is a valid single step, as is the epenthesis of unmarked and minimally specified segments, and that the epenthesis of a fully specified, autosegmentally linked segment is not a valid step.

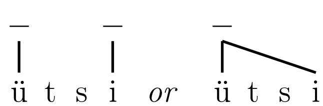
1.1 The constraint and representation

Serial Harmonic Grammar, under Kimper's formulation, inherits Harmonic Serialism's tolerance of positively formulated constraints, which reward candidates for demonstrating certain structures, rather than penalizing them. The constraint enforcing harmony is one of these:

(1) SPREAD(±F): For a feature F, assign +1 for each segment linked to F as a dependent. (Kimper, 2011)

In the representation used here, one of the segments linked to a feature value node F is the head, and additional segments linked to it are dependents. A set of segments is considered to be linked if they all share the same specification for F on the tier for feature F, and thereby all display the same value of F. A candidate can establish a link between two segments with different values of F, but this forces one of them to change in its value for F (represented as in 2), and introduces a faithfulness violation. The grammar does not require that segments that have identical values for F be linked to one another. It is possible for each to have its own specification, or for the two to share one, as in (3):

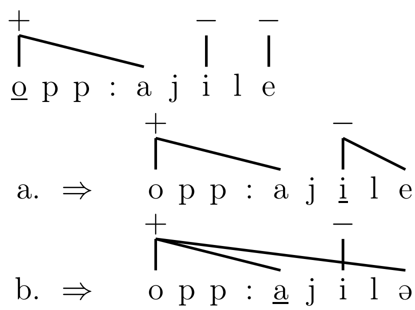
(2) 

(3) 

The underlying forms that I use here have one feature value specification on the back tier for each segment, leaving each segment as the head of its own feature value domain, though

I allow that inputs may contain already-linked segments[1].

Breaking from more mainstream uses of autosegmental representations (Goldsmith, 1976), Kimper's system allows for crossing lines: two segments can be linked to the same feature node even if there are additional segments between them that are linked to a different feature node, as in (4b). This is key to the framework's mechanism for transparent vowels. I demonstrate this here with two possible ways of realizing harmonic spreading with the input /opp:ajile/ ('teacher-PL-ALL'), one in which [i] is opaque (4a) and one in which it is transprent (4b):

(4)


This representational scheme shares its explicit tolerance of non-local spreading with another scheme that has been proposed for harmony, called Agreement by Correspondence (ABC, Walker, 2001). Kimper provides some arguments against the specific cluster of harmony constraints used for ABC, but it is may be possible to adapt an analysis based on Trigger Competition and the SPREAD constraint to use ABC-style representations.

1.2 Triggers

In Trigger Competition, *triggers* are the segments through which an autosegmental linking operation is initiated: spreading proceeds from a trigger, spreads the feature value of the trigger, and assigns rewards conditioned on the identity of that trigger. Whether or not a segment is a trigger at a given point in the derivation is necessarily determined in conjunction with the rewards assigned by the harmony constraint, SHARE, to a given candidate. Any effects involving the identity of the target are determined by other constraints: SHARE does not discriminate among vowel types for targets.

A segment is only treated as a trigger in the specific tableau in which it spreads, and does not retain this property, even if harmonic spreading continues in the same word. Because of this, it does not correspond well to the conventional notion of a domain head. In the representation that I use here, triggers are marked with underlines in the candidates in which they appear. The precise visual placement of the feature node is of no theoretical significance: in order to maximize continuity between steps, I leave the node for a feature value directly above the first segment to spread that value.

Kimper allows for transparent vowels by allowing association links to connect two segments to a feature value, even if there are segments in between them and even if those intervening segments are linked to some other feature value (as in 4b). Simply allowing

[1]A study of loanwords and other potential domains for disharmonic stems would be valuable here: It may be possible to use Kimper's link faithfulness constraints build a grammar that preserves harmony patterns that appear in the input which would not otherwise be generated.

for arbitrarily distant links, though, misses the basically local tendency of vowel harmony and predicts transparency effects to be extremely widespread. To remedy this, Kimper introduces a constant scaling factor k which diminishes the reward assigned by the harmony imperative SPREAD for each unit of distance between the two segments being linked:

(5) **Scaling factor: non-locality**
For a trigger α and a target β, multiply the reward earned for the dependent segment β by a constant k (such that $1 > k > 0$) for each unit of distance d intervening between α and β. (Kimper, 2011)

Kimper proposes that the mora should be the unit of distance, but does not provide any evidence to motivate this proposal, nor does the proposal seem crucial to any part of his analysis. In order to simplify presentation and to better address variation in Hungarian, I adopt the vowel segment as the unit for this paper. Spreading past long and short vowels incurs the distance multiplier once, and spreading past consonants of any kind or number incurs no penalty. The only crucial function of the spreading factor is that for any two potential vowel triggers for a given target, it must always be preferable to link with the nearer one.

Kimper introduces a second, similar, scaling factor to account for the theoretically central notion of trigger strength:

(6) **Scaling factor: trigger strength**
For a trigger α, a target β, and a feature F, multiply the reward earned for the dependent segment β by a constant x (such that $x > 1$) for each degree i to which α is perceptually impoverished with respect to $\pm$F. (Kimper, 2011)

This captures the following intuition:

> Segments which are better triggers (by virtue of being perceptually impoverished, and therefore in most need of the advantages conferred by harmony) are more likely to be opaque, and segments which are poor triggers are more likely to be treated as transparent. (Kimper, 2011)

In Hungarian backness harmony, /a/ is a stronger trigger than /e/, and /e/ is in turn a stronger trigger than /i/. The prediction made by this claim is that in any environment in which /i/ will trigger harmony, /e/ will, and in any environment in which /e/ will trigger harmony, /a/ will as well. In particular, the scaling factor captures this in a way that is directly tied to a perceptual property of those three vowels: a typical canonical [a] manifests backness only rather poorly, [e] manifests frontness somewhat more clearly, and [i] more clearly still.

There is currently no mechanistic means by which one can choose i values for vowel types. Without doing so, it is further impossible to set the multiplier x. For the time being, I avoid this problem by simply setting individual *trigger strength values* for each segment—hypothesized to be the product of the unknown i values for each segment and the unknown x value—though I come back to this matter in section 3.9.

1.3 Motivating a positive non-local harmony constraint

Positive constraints are fairly unusual in constraint-based grammar, and deserve some skepticism. However, neither of the case studies discussed here can be implemented using any negatively formulated constraint that I am aware of in the literature, and it is possible that no negative constraint can achieve these effects: Spread's sensitivity to both distance and trigger strength rely on its status as a positive constraint.

There have also been extensive discussions in the literature about whether it is possible or desirable for vowel harmony to be strictly local, and I follow Kimper in taking a stance against strict locality in this paper, and allow the grammar to generate interactions between non-adjacent vowels. Kimper provides two broad arguments against locality, one from grammatical pathologies like sour grapes that emerge from strictly local constraints and another from the experimental observation that the perceptual accuracy advantage provided by harmony systems does not appear to be diminished by non-local effects like transparent vowels.

It is not possible, when looking at harmony constraints, to fully disentangle the effects of positive or negative polarity from the effects of other choices about constraint design like locality: negative constraints cannot generally be made into usable positive constraints by simply rewarding the opposite of the penalized structure. For example, one might try to propose a negative counterpart to Spread that penalized unlinked vowels, rather than rewarding linked ones. However, Kimper's treatments of distance and trigger strength would not be available, since the novel constraint would only refer to the segments that ceased to be isolated, and not to the segments that initiated the linking. It may be possible to incorporate factors like distance into such a constraint, but this would require a wholly novel approach, which could introduce new and different problems.

Similarly, a local constraint like Agree cannot be straightforwardly made positive for comparison with Spread. An attempt to do so might yield a constraint which rewards neighboring vowels which agree, but this would predict many behaviors not seen in the typology of Agree: for example, adding a suffix to a word would not change the number of violations assigned by Agree is that suffix shared a feature value with the last vowel in the stem, but the new positive constraint would have to assign an additional reward, since there would be an instance of agreement between the last two vowels which was not previously present.

In the absence of any viable negative constraint to compare with Spread, it is necessary to compare the positive non-local constraint directly with the common but substantially different negative local constraints like Agree and Align, as I do here.

Kimper points out problematic predictions made by both of these, and disputes their empirical adequacy. While I leave open the question of whether either constraint makes appropriate predictions about any kind of harmony system, I claim that neither can capture the phenomena discussed in the following sections: negative local constraints cannot be made to model variable degrees of non-locality, as is required by Hungarian, nor can they be made to generate transparent vowels that are harmonically paired in the inventory, as is required by Seto. The positive non-local constraint Spread is the only constraint that I am aware of capable of capturing both these facts.

1.4 Spread in action

I demonstrate the Spread constraint here with the trivial example of a monosyllabic stem spreading harmony onto a monosyllabic affix:

(7)

/a+ä/ (+ linked to a, − linked to ä)	StemID[Bk] −1	Spread[±Bk] +1	$\mathcal{H}$
a. a+ä (+ linked to a, − linked to ä)	0	0	0
☞ *b.* a̲+a (+ linked to both vowels)	0	1	1

The reward of 1 associated with each instance of spreading is the product of the trigger strength value for each trigger (here the trigger /a/—marked with an underline—has a trigger strength of 1) and the distance factor k raised to a power representing the number of non-harmonizing vowel segments that need to be crossed to spread. Since harmony is perfectly local here, this power is 0, and the product is 1. Multiplying each violation count by the weight of each constraint and summing the results yields the total score for that candidate, denoted $\mathcal{H}$, by which the winner is selected.

There are only two candidates here: failing to harmonize incurs no constraint violation, but rightward harmony earns a reward and wins. I provisionally assume that there are no viable candidates involving leftward spreading in the languages discussed here, and address that assumption in section 3.7 below.

In the next step, the grammar has the opportunity to make another change. There are no more front vowels in the input, and so there is no motivation for any change to the vowel qualities. It should be noted that since the input to this step corresponds to the output of the previous step, there is no faithfulness violation associated with preserving the harmony established in the previous step. At this point, the derivation converges, and [a+a] is the output. A sample operation, delinking,[2] is provided only to show the impossibility of any further improvement.

(8)

/a+a/ (+ linked to both vowels)	StemID[Bk] −1	Spread[±Bk] +1	$\mathcal{H}$
a. a+a (+ linked to a, + linked to a)	0	0	0
☞ *b.* a+a (+ linked to both vowels)	0	1	1

[2]Delinking is never optimal in the grammar presented here. No constraint penalized multiply-linked features, and delinking does not immediately change feature values, so it cannot benefit in the tableau by repairing faithfulness violations.

This constraint, combined with standard markedness and faithfulness constraints, generates a tidy typology of phenomena incorporating vowel harmony, neutral (non-participating) vowels, and neutralization, wherein some segment is banned from the inventory entirely and merges to some other segment. The following factorial typology presentation is adapted from Kimper's (45), and examples and citations for the languages named can be found there.

(9) **Typology: Harmony and Neutralization**

a. Spread(atr) ≫ *(+hi,–atr) ≫ Ident(atr)
Harmony with full participation; neutralization (e.g Khalkha)

b. Spread(atr) ≫ Ident(atr) ≫ *(+hi,–atr)
Harmony with full participation; no neutralization (e.g. Kasem)

c. *(+hi,–atr) ≫ Ident(atr) ≫ Spread(atr)
No harmony; neutralization (e.g. Hebrew)

d. *(+hi,–atr) ≫ Spread(atr)≫ Ident(atr)
Harmony with non-participants; neutralization (e.g. Yoruba)

e. Ident(atr) ≫ Spread(atr) ≫ *(+hi,–atr)
Ident(atr) ≫ *(+hi,–atr) ≫ Spread(atr)
No harmony, no neutralization (e.g. English)

The broader range of possible combinations of scaling factor values and constraint weights makes it difficult to develop a more complete typology by hand, and no software toolkit for constraint-based grammar currently available is capable of generating typologies for Serial Harmonic Grammar with scaling factors. However, the typology that emerges looking only at the novel scaling factors is reasonable:

(10) **Typology: Distance and Trigger Strength**

If the constraints named in the previous typology are weighted so as to allow for harmony, then the following typological generalizations can be made about the relationship between the trigger strength and distance multipliers. I focus on the grammar's effects on the strongest trigger in the inventory, which is the least prone to be transparent and which I refer to as v_s, and the weakest (the most prone), which I refer to as v_w.

a. $k * x[v_s] < x[v_w]$
Strictly local harmony (Ọyọ Yoruba, Kimper, 2011): Local harmony from weak trigger v_w will always receive a higher reward than long-distance harmony from strong trigger v_s. Any neutral vowels will be opaque.

b. $k * x[v_s] \geq x[v_w] \geq k^n * x[v_s]$
Vowel harmony with distance-sensitive transparency (Hungarian): Long-distance harmony from strong trigger v_s will receive a higher reward than local harmony from weak trigger v_w, as long as the distance involved is less than n. Tokens of v_s cannot spread harmony past any more than n segments, because beyond this point, the distance penalty will diminish the reward for spreading so greatly that it is better to spread from a nearby token of v_w. If v_w is a neutral vowel, then individual tokens of v_w will be transparent, but sequences of n tokens of v_w will be opaque.

c. $k^n * x[v_s] \gg x[v_w]$
Vowel harmony with categorical transparency (Seto): If this inequality holds for large values of n (i.e., if k is 1 or near 1), then distance effects cease to appear. Long-distance harmony from strong trigger v_s will receive a higher reward than local harmony from weak trigger v_w in all cases. If v_w is a neutral vowel, it will be transparent.

To summarize, if the distance factor is zero or near zero, then the language will not display transparency. If it is one or near one, then the language will allow harmony to spread past an arbitrarily large number of transparent vowels. The higher the trigger strength factor for some vowel is relative to the faithfulness and markedness constraints penalizing harmonic alternations, the more different vowel types and contexts will show alternations in the presence of that triggering vowel.

It should be noted that the introduction of Spread does have side effects outside of the purview of harmony. Since deleting a vowel mora eliminates any reward that vowel might have been assigned by Spread, this approach predicts the existence of languages where vowels will reduce to schwa in some context *only if they are not eligible to participate in harmony.* While this may sound like a pathology, Kimper shows that precisely this phenomenon appears in Kera.

1.5 Motivating serial evaluation

With the positive Spread selected as the harmony constraint, the choice of a serial grammar is unavoidable: serial evaluation is necessary for any positively formulated constraint in order to prevent a host of pathologies surrounding the infinite goodness problem. Essentially, if a constraint rewards a structure (such as a linked vowel), and the grammar can generate that structure (such as by epenthesizing and linking a vowel), then under many constraint weightings, the language will generate outputs with infinitely many copies of the structure:

(11)

− + /ä+a/	Spread[±Back] ...	DepC ...	$\mathcal{H}$
a. − + ä+a	0	0	...
b. ☹ − ä+ä	1	0	...
c. − ä+ä ä	2	1	...
d. ☞ − ä+ä ä ä ä ä...	∞	∞	...

In a standard serial constraint-based grammar like the one presented here, however, it is not possible to both epenthesize and link in a single tableau, so in the first tableau where a candidate with epenthesis is considered, that candidate will receive no reward from SPREAD, and will be no more likely to win than it otherwise would be, preventing harmony from producing unbounded epenthesis.

More subtly, situating a positive harmony constraint in a serial grammar prevents another set of pathologies. No language has been observed to modify segments to improve their suitability as triggers *only* when they actually serve as triggers. A parallel grammar in which the reward for spreading is associated with trigger strength would predict this to be a fairly common occurrence: any time the grammar prefers harmony from a moderate trigger over faithfulness, it will also prefer harmony from a strong trigger over harmony from a moderate trigger. If the faithfulness constraints protecting the features distinguishing moderate triggers from strong ones (e.g., height or roundness) are not weighted highly enough, then these moderate triggers will be transformed into strong ones. In a serial grammar, a solution is available: this transformation would constitute its own step in the derivation that would follow harmony. Basing the reward for harmonic spreading on the strength of the trigger at the step where spreading occurs prevents the harmony constraint from inducing this transformation.

1.6 Motivating weighted constraints

Trigger Competition is implemented in a version of the weighted constraint system Harmonic Grammar, Optimality Theory's most immediate ancestor. Superficially, the choice of weighted constraints over ranked constraints seems like a substantial increase in complexity. For a given set of n constraints Optimality Theory defines only $n!$ different possible rankings, but Harmonic Grammar defines an infinite number of weightings, each theoretically (though not practically) representing a different grammar. Furthermore, Harmonic Grammar introduces a new means of constraint interaction in the gang effect[3]. Ultimately, though, this choice does not pose a considerable additional burden to the speaker. As Bane et al. (2010) show, for a given number of constraints, the VC dimensions of both types of constraint-based grammar are identical and linear in the number of constraints. This measure assures us that, even in the worst case, both can be learned from a similar, and typically manageable, number of examples.

Practically, no analysis based on a SPREAD constraint or anything similar can function in ranked-constraint Optimality Theory. One key obstacle to this is the treatment of non-local spreading: Trigger Competition distinguishes between varying degrees of locality using the distance multiplier. The reward for spreading crucially decreases with increasing distance, but it must also remain positive—since spreading at any distance is always better than doing nothing when no other constraint is violated. These two limitations on the theory require that the grammar be able to distinguish arbitrarily small real numbers, which OT, with its integer-valued violation counts, fundamentally cannot do.

Furthermore, these fine distinctions can sometimes lead to gang effects, wherein high violation scores from low-weighted constraints cause a constraint to win even though it is dispreferred by a highly weighted constraint. These are necessary with Trigger Competition

[3]This is said to occur when many violations of low-weighted constraints collectively have a greater effect on the output than does a highly weighted constraint

in order to allow harmony from strong triggers to overcome faithfulness constraints which would prevent spreading from weaker triggers. Optimality Theory uses strict domination between candidates, rather than weighting, and disallows interactions of this kind.

1.7 Motivating violation count multipliers

On the assumption that phenomena like non-local spreading and variable trigger strength need to be modeled directly, and cannot be made to emerge naturally from an analysis of other facts, the multipliers that Kimper introduces are a simple way to capture them. While they do add some complexity to the basic structure of the model in that they add a new *kind* of parameter, they serve to considerably reduce the number of parameters that are required by any particular model in practice. In particular, assuming using multipliers on the weight of a constraint requires the learner to acquire fewer parameter values than would be the case if these phenomena were represented by large hierarchies of constraints of forms like SPREAD[FROM:E,DISTANCE:1], SPREAD[FROM:E,DISTANCE:2], SPREAD[FROM:Ä,DISTANCE:2]...

1.8 Motivating operation-specific triggers

Mullin (2011) advocates for an approach to autosegmental spreading in harmony in which it is the head of a harmonic domain which initiates spreading. The notion of the durable head—the segment which initially carries the spreading feature—is familiar in the autosegmental phonology literature (i.e., Goldsmith, 1976), and it seems reasonable to posit that the identity of the head of a harmonic domain might impact the ability of the harmonic domain to spread.

However, combining this representational technique with the understanding that segments differ in the degree to which they induce harmony can lead to problems, especially when combined with scaling factors for distance and trigger strength. If it were possible to perform operations subsequent to harmony that improve the reward associated with harmony, we would expect triggers to change identity to become stronger triggers, but only if they harmonize, or even only if they harmonize with a sufficient number of targets for the change to be worth the reward. In addition, the grammar would be able to generate languages in which segments (consonants or vowels) were eligible for deletion *only* if they came between two harmonizing segments without themselves participating in harmony, since doing so would reduce the distance penalty associated with the two harmonizing segments.

To avoid this and related pathologies, Kimper's approach rests firmly on a different notion of autosegmental spreading. While he allows the grammar to refer to durable domain heads for other purposes, the SPREAD constraint assigns rewards on the basis of operation-specific triggers. The trigger is defined as follows:

(12) For a feature $\pm F_i$ and a segment S, create an autosegmental association between F_i and S. If S is associated with another instance $\pm F_j$, the link between S and $\pm F_j$ is dissolved. For each instance of an operation creating a new autosegmental link (spreading)...

  a. The target is a segment which is not associated with F in the input, but is associated with F in the output (candidate).

b. The trigger is the segment already associated with F in the input which is linearly closest to the target.

The reward associated with an instance of spreading is then determined by the target and its relationship with the trigger at the point in the derivation at which spreading occurs.

1.9 An aside on cyclicity

I assume solely for expositional clarity that phonological domains are evaluated cyclically, in the style of Kiparsky (2000). That is, each time an affix is added to a form, that stem–affix pair is evaluated, with the output of that evaluation serving as the input to future evaluations.

It is not crucial to this analysis that the affixes be added one-by-one, as long as faithfulness constraints can distinguish between roots and affixes. In a cyclic grammar, a constraint, denoted here as StemIdent[Back] preferentially enforces faithfulness to the stem at each cycle of evaluation. Kimper uses a similar constraint that makes reference to the monomorphemic root rather than the cycle-specific stem of affixation:

> I assume that root control is accomplished via positional faithfulness (Beckman, 1998) – a ranking of Ident(ATR)$_{Rt}$ $\gg$ Ident(ATR)$_{Aff}$ will ensure that harmony that targets affixes is consistently preferred over harmony that targets roots.

Were words built all at once, with all affixes present from the start of the derivation, my analysis could be readily implemented using this constraint pair in this ranking.

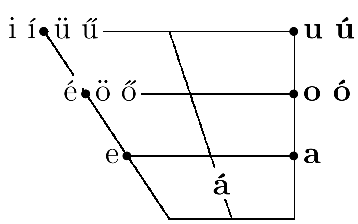
I will now demonstrate the effectiveness of this system using two case studies: Hungarian vacillation and Seto neutral vowels.

2 Hungarian vacillation and tied candidates

Hungarian backness harmony shows a pair of related phenomena involving neutral vowels which are difficult to account for in any conventional generative approach to vowel harmony. In this section I provide a somewhat simplified sketch of the harmony system that spotlights these problematic phenomena, and show how they can be accounted for in Trigger Competition.

In the below chart, I present the structure of the vowel inventory presented in Ringen and Vago (1998), depicted in Hungarian orthography. Segments in bold type count as back vowels from the prospective of harmony, and contrastively long vowels are marked with acute (´) accents:

(13)



I follow Benus et al.'s (2003) description of the behavior of Hungarian transparent vowels. They find that the language has three transparent vowels, /i/, /í/, and /é/, as well as one partially transparent vowel, /e/. These four vowels show an unusually complex pattern of behaviors: if a single transparent vowel intervenes between a back vowel and a suffix vowel, the suffix vowel will take on back harmony, as in (14a). If an /e/ intervenes, then front or back harmony is possible, as in (14b). If two transparent vowels intervene, then both are again possible, as in (14c). Finally, if a transparent vowel and an /e/ both intervene, only front harmony is possible, as in (14d).

(14)
a. *papír-ban* ∼ **papír-ben* 'paper-INESS' (transparent vowel)
b. *ágnes-ban* ∼ *ágnes-ben* 'Agnes-INESS' (/e/)
c. *oxigén-ban* ∼ *oxigén-ben* 'oxygen-INESS' (two transparent vowels)
d. **kabinet-ban* ∼ *kabinet-ben* 'administration-INESS' (T.V. plus /e/)

Kimper mentions these phenomena in a footnote, but does not attempt to address them:

> In Hungarian, harmony is optionally opaque across a series of non-participants — the present discussion abstracts away from this variation. (Kimper, 2011)

In the implementation of his system used here, though, the phenomena emerge quite readily: a reasonable attempt to capture the rest of Hungarian yields a grammar which generates the categorical effects that we observe, and generates ties for the cases of variation. We need only assume that most of the back vowels are strong triggers, that /e/ is a less strong trigger, and that the three transparent vowels (/i/, /í/, and /é/) are weaker still. If the ratios between the three different values are similar to the ratio expressed by the distance multiplier, then precisely this phenomenon emerges.

In its outline, my analysis follows the intuition behind the presentation of Hungarian vacillation using local agreement in Hayes and Londe (2006)[4]. I present a simplified sketch of the analysis below: I don't establish a set of inventory constraints, but the results are no less tenable when the full harmony system of the language is implemented.

Parameter	Value
k (distance multiplier)	0.5
x[default] (trigger strength multiplier)	4
$x[e]$ (trigger strength multiplier)	2
$x[é, í, i]$ (trigger strength multiplier)	1

Table 1: Spreading parameter values for Hungarian in Hungarian orthography

These simple constraints enforce harmony in the typical case, shown here with the example of *haz-nak* 'house-DAT':

[4] Hayes and Londe (2006) achieve a fairly close fit to observed frequencies in their model. It would be potentially informative to choose a particular stochastic grammar framework and attempt to weight the constraint set presented here to model the same frequency data presented there.

(15)

+ − /h á z + n e k/	StemID[±Bk] −5	Spread[±Bk] +1	$\mathcal{H}$
a. + − h á z + n e k	0	0	0
☞ *b.* + h á z + n a k	0	4	4

A single transparent vowel allows harmony to propagate past it:

(16)

+ − − /p a p í r + b e n/	StemID[±Bk] −5	Spread[±Bk] +1	$\mathcal{H}$
a. + − − p a p í r + b e n	0	0	0
b. + − p a̲ p Í r + b e n	1	4	−1
☞ *c.* + − p a̲ p í r + b a n	0	$4 \times .5^1 = 2$	2
d. + − p a p í̲ r + b e n	0	1	1

Even though candidate *c* incurs a distance penalty (a multiplier of .5) for skipping the middle vowel, it still gains a greater reward than the local harmony candidate by spreading from a better trigger (strength 4, rather than 1). At this point, no further improvement is possible, and the derivation converges with back harmony:

(17)

+ − /p a p í r + b a n/	StemID[±Bk] −5	Spread[±Bk] +1	$\mathcal{H}$
a. + p a̲ p í r + b a n	1	$4 \times .5^1 + 4 = 6$	1
☞ *b.* + − p a p í r + b a n	0	$4 \times .5^1 = 2$	2
c. + − p a p í̲ r + b e n	0	1	1

A single token of the translucent vowel /e/ leads to tied harmony:

(18)

+ − − /á g n e s+b e n/	StemID[±Bk] −5	Spread[±Bk] +1	$\mathcal{H}$
a. + − − á g n e s+b e n	0	0	0
b. + − á̲ g n a s+b e n	1	4	−1
☞ *c.* + − á̲ g n e s+b a n	0	$4 \times .5^1 = 2$	2
☞ *d.* + − á g n e̲ s+b e n	0	2	2

Upon selecting either output, the grammar can make no more changes that affect the vowel qualities of the output, and the derivation converges. This yields a tie, which can be reasonably interpreted as allowing some form free variation. For one possible such interpretation, if we assume that numeric computation within the grammar is imprecise—as is generally true of cognitive processes—then the slightest perturbation in the harmony score of either candidate will yield a single winner, and we can expect both front and back harmony to surface.

Multiple consecutive normal transparent vowels also yield tied harmony, though by way of a somewhat more complex derivation:

(19)

+ − − − /o x i g é n+b e n/	StemID[±Bk] −5	Spread[±Bk] +1	$\mathcal{H}$
a. + − − − o x i g é n+b e n	0	0	0
b. + − − o̲ x ɨ g é n+b e n	1	4	−1
☞ *c.* + − − o̲ x i g é n+b a n	0	$4 \times .5^2 = 1$	1
☞ *d.* + − − o x i g é̲ n+b e n	0	1	1
☞ *e.* + − − o x i̲ g é n+b e n	0	1	1

If candidate *c* is selected, then in the next step of the derivation the two word-internal front vowels vacuously harmonize with one another, without affecting back harmony in the suffix:

(20)

+ − − /o x i g é n+b a n/	StemID[±Bk] −5	Spread[±Bk] +1	$\mathcal{H}$
a. + − − o x i g é n+b a n	0	$4 \times .5^2$ $= 1$	1
☞ *b.* + − o x i̲ g é n+b a n	0	$4 \times .5^2 + 1$ $= 2$	2
c. + − o̲ x ɨ g é n+b a n	1	$4 \times .5^2 + 1$ $= 2$	−3

At this point, no further improvement is possible, and we have an output with back harmony. However, if candidate *d* is selected in the first step, then a similar stem-internal harmony takes place:

(21)

+ − − /o x i g é n+b e n/	StemID[±Bk] −5	Spread[±Bk] +1	$\mathcal{H}$
a. + − − o x i g é n+b e n	0	1	1
☞ *b.* + − o x i̲ g é n+b e n	0	$1 + 1$ $= 2$	2
c. + − o̲ x ɨ g é n+b e n	1	$1 + 4$ $= 5$	0

Again, no further improvement is possible, and front harmony results. If candidate *e* wins in the original tableau, then similar derivations make both front and back harmony possible with the same two output structures as result from choosing *c* and *d*. Ultimately a sequence of a transparent vowel and an /e/ yields free variation in this grammar, with an equal number of outputs for both back and front harmony.

Finally, this system correctly predicts that a transparent vowel followed by the translucent /e/ will categorically select for front suffixes:

(22)

/k a b i n e t + b e n/	StemID[±Bk] −5	Spread[±Bk] +1	$\mathcal{H}$
a. k a b i n e t + b e n	0	0	0
b. k a b i n e t + b e n	0	$4 \times .5^2 = 1$	1
c. ☞ k a b i n e t + b e n	0	2	2
d. k a b i n e t + b e n	0	1	1
e. k a b ɨ n e t + b e n	1	4	−1

In the next step the two front vowels in the stem are linked, and the derivation converges, with no optionality and front harmony on the affix:

(23)

/k a b i n e t + b e n/	StemID[±Bk] −5	Spread[±Bk] +1	$\mathcal{H}$
a. k a b i n e t + b e n	0	2	2
b. ☞ k a b i n e t + b e n	0	$1 + 2 = 3$	3
c. k a b ɨ n e t + b e n	1	$4 + 2 = 6$	1

Building a Trigger Competition grammar that generates these phenomena necessarily makes predictions about longer strings of transparent vowels: a token of a normal back vowel at a distance of greater than two cannot beat local spreading from even the weakest transparent vowel, so the grammar rules out sequences like [a.i.i.i+nak].

This analysis requires nothing more than the most basic components of a Trigger Competition grammar: the harmony constraint Spread, and some other constraint (IdentStem here, though constraints like *ɨ would be needed for a full account) to prevent transparent vowels from undergoing back harmony. By contrast, in any system which treats different vowel types identically for the purposes of the harmony constraint, and in any system which does not explicitly recognize the scalar effect of locality on harmony, an analysis like this would require considerable additional theoretical technology.

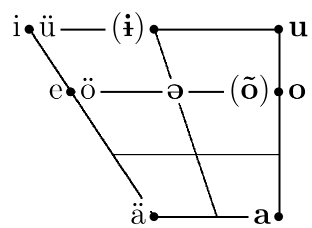
3 Seto and paired neutral vowels

Seto, a Finno-Ugric language (or dialect) spoken in southern Estonia, shows a set of behaviors which render straightforward analysis under any conventional framework impossible. In the following section, I sketch out the properties that make Seto difficult to analyze, and show how it can be comprehensively captured using a version of Trigger Competition. The analysis presented here represents the first generative account that I am aware of for Seto vowel harmony.

3.1 An introduction to Seto vowel harmony

I draw my data from Kiparsky and Pajusalu's (2001) unpublished manuscript (henceforth K&P), key pieces of which are excerpted in this section. I depict the inventory described in that paper in (24) with two notational conventions: segments in parentheses only appearing in syllables with primary stress, which always falls on the first syllable of the word, and segments in bold type count as back vowels from the prospective of harmony:

(24)



The orthography corresponds fairly closely to the IPA standard, with the exception of the orthographic *ü* (IPA /y/), *ö* (IPA /ø/), ä (IPA /æ/), and *õ* (IPA /ɤ/). All vowel symbols in this paper, except where noted otherwise, are drawn from the orthography, with one exception: I use /ə/ for non-initial lax vowels, where the orthography would use 'õ', the same symbol as is used for a word-initial back vowel. Expressing this contrast simplifies the analysis of some of the unusual properties of /e/.

These vowels are subject to the following harmonic alternations:

(25) ü–u, ö–o, ä–a, e–õ_{init}, e–$\text{ə}_{\text{non-init}}$

Harmony spreads rightwards: Initial syllables determine the harmonic feature values of subsequent syllables, and stems determine the harmonic feature values of suffixes. There are no attested prefixes. Harmonic alternations and related neutralization in Seto disyllabic words are summarized in the Table 2, and more complex phenomena emerging in longer words are described in the following sections.

Seto features two transparent vowels: /i/ is transparent in all positions, and allows harmony to progress past it. /e/ is transparent only word-initially, and allows front and back vowels to follow it. Stems containing only transparent vowels require suffixes with front vowels, suggesting that there may be something to be gained by an analysis which does not exclude them from harmony entirely.

Seto also features one opaque vowel: /o/ can occur after vowels of either feature value, but allows only back vowels and transparent vowels vowels to follow it.

V1: \ V2 input:	/i/	/e/	/ä/	/ü/	/ö/	/ɨ/	/õ/	/a/	/u/	/o/
[i]	i	e	ä	ü	+ o	+ i	∼ e	∼ ä	∼ ü	o
[e]	i	e	ä	ü	+ o	+ i	+ e	∼ ä	∼ ü	o
[ä]	i	e	ä	ü	+ o	+ i	+ e	+ ä	+ ü	o
[ü]	i	e	ä	ü	+ o	+ i	+ e	+ ä	+ ü	o
[ö]	i	e	ä	ü	+ o	+ i	+ e	+ ä	+ ü	o
[ɨ]	i	+ ə	+ a	+ u	+ o	+ i	ə	a	u	o
[õ]	i	+ ə	+ a	+ u	+ o	+ i	ə	a	u	o
[a]	i	+ ə	+ a	+ u	+ o	+ i	ə	a	u	o
[u]	i	+ ə	+ a	+ u	+ o	+ i	ə	a	u	o
[o]	i	+ ə	+ a	+ u	+ o	+ i	ə	a	u	o

Table 2: Vowel cooccurrence patterns in disyllabic words: each cell shows the effect of vowel harmony on the second vowel (indicated in the top row) when placed after the first vowel (indicated in the leftmost column). Vowels marked with a '+' show evidence of neutralization or harmonic alternation. Vowels marked with a '∼' display the depicted alternating behavior only in affixes. (K&P)

Positional neutralizations appear to interact with harmony: the phoneme represented orthographically as *õ* surfaces as /õ/ word-initially and /ə/ elsewhere, and shows some unique harmonic behaviors non-initially—harmonizing after initial /e/ and in some non-harmonic affixes—that appear to be limited to lax vowels. The presence of /ɨ/ word-initially further complicates the analysis of harmony, as I will discuss below.

While there are other reflexes of stress and footing in the language, harmony is not subject to effects of stress or linear position. The first syllable has a distinctive inventory, and no preceding syllable to compel it to harmonize, but besides that, the cooccurrence restrictions on a sequence of vowels is the same regardless of the linear position of that sequence or what secondary stress (if any) is assigned to it.

These cooccurrence patterns are demonstrated below. When other processes do not intervene, harmony is categorical within word forms. According to K&P, the following starred words below are impossible in Seto, despite being minimally different from their harmonic counterparts to the left:

(26) a. *ruga* 'stack' *rügä* 'rye' *(*rugä, *rüga)*
b. *sɨna* 'word' *sinä* 'you' *(*sɨnä)*
(K&P)

Seto suffixes obligatorily harmonize, as shown in the following alternations:

(27)

a.	*pan-daq*	'put-INF'	*müü-däq*	'sell-INF'
b.	*nõsə-sə*	'rise-3'	*elä-se*	'live-3'
c.	*opp:a-ji-lə*	'teacher-PL-ALL'	*rebäs-i-le*	'fox-PL-ALL'
d.	*saa:-ma*	'get-INF'	*jää:-mä*	'stay-INF'
e.	*naĺ̗a-tta-nuq*	'joked-CAUS-PST.PTCP'	*nälü-ttä-nüq*	'starve-CAUS-PST.PTCP'
f.	*tütt:re-kkene*	'daughter-DIMINUITIVE'	*maama-kkənə*	'mom-DIMINUITIVE'
g.	*klɨbisə-ma*	'rattle-INF'	*libise-mä*	'flutter-INF'

(K&P)

Seto is not a typological anomaly: it is one of the fullest examples of a pattern of harmony systems demonstrated across the region (Kiparsky and Pajusalu, 2003). It shares the basic phenomena of backness harmony with transparency with Finnish, Votic, Veps, and most major non-Tallinn Estonian dialects, and shares the especially problematic /i/–/ɨ/ pair with Votic and N. Tarto Estonian.

3.2 What makes Seto hard to analyze?

While I am aware of no way to determine a pretheoretically correct feature system for the inventory of a language, one fact about Seto seems fairly difficult to contest: the two members of each of the pairs /i/–/ɨ/, /e/–/õ/, and /o/–/ö/ seem to differ from one another in approximately the same way as in our normal alternating pairs /a/–/ä/ and /u/–/ü/. I call the feature that distinguishes these pairs BACK. I can find no evidence that would support the introduction of a new feature, distinct from this one, for only the first three pairs, and thus hypothesize that the three neutral vowels are each paired with counterparts differing only in harmonic feature value. This fact, the presence of paired neutral vowels, causes considerable difficulty for the generative analysis of harmony in this language, as I discuss in the following sections:

3.2.1 It's not neutralization

Seto exhibits a pair of phenomena which together demonstrate that transparent vowels are more than just the byproduct of a neutralization process: the paired vowels /ɨ/ and /i/ are both licit in initial syllables, and there is no evidence for a complementary distribution of any kind. A near minimal pair is shown here:

(28) *klɨbisə-ma* 'rattle-INF'

(29) *libise-mä* 'flutter-INF'

The two are contrastive in this position. However, /i/ is also transparent in this position, since it can occur adjacent to back vowels. Since we know that /i/ must be contrasted from its back counterpart /ɨ/ at any phonological representation, we know that the harmony system must genuinely tolerate sequences like /i.a/ or /i.u/ even underlyingly.

This in turn means that no analysis can explain transparent vowels by claiming that they are harmonic at some level of representation before neutralization changes them into a potentially-disharmonic fixed surface form. This is an analytical approach that is often proposed in order to be able to maintain that harmony is strictly local and is used by, among others, Walker (1998) and Baković and Wilson (2000) in constraint-based frameworks, and Bach (1968) and Clements (1976) in earlier phonological traditions. A schematic example of this kind of two step approach is shown in (30). In the first phase of the derivation, words are made to be fully harmonic, but in the second a neutralization process changes the non-surfacing /ɨ/ into an [i] where it appears:

(30) a. /babi/ → /babɨ/ → [babi]
b. /bäbi/ → /bäbi/ → [bäbi]

In order to account for Seto, a grammar must be able to generate both of the patterns /i.a/ and /ɨ.a/, which neutralization systems cannot readily do. Typical neutralization grammars do not have a way of tolerating local disharmony at the stage of the derivation where harmony occurs: neutralization is meant to be the sole explanation of surface disharmony. So we must assume that the two vowels in both words harmonize with one another.

To generate /i.a/ in a system with progressive local harmony, it is necessary to start with an input that is specified as back in the initial syllable: if we assume that the initial syllable is a front vowel, than progressive harmony will force us to generate /i.ä/, which though attested is not our goal. Hence, we begin with the input /ɨ.a/:

(31)

/ɨ.a/	AGREE[BK]	IDENT[BK]	*ɨ
a. ☞ ɨ.a			*
b. i.a	*	*	
c. i.ä		**	

At the harmony stage of the grammar, then, the initial vowel remains back. To generate the observed surface form, grammar must force that back vowel to become front. With a different ranking of these constraints in the postlexical phonology, this is easily done:

(32)

/ɨ.a/	*ɨ	IDENT[BK]	AGREE[BK]
a. ɨ.a	*		
b. ☞ i.a		*	*
c. i.ä		**	

The grammar yields our desired form. If we now want to generate the output /ɨ.a/, we are in a bind. Starting with precisely the form that we want to generate, as above, does not yield the output we need. Since /ɨ/ is the best back counterpart in the inventory to /i/, the previous example required us to say that /ɨ/ neutralizes to /i/ postlexically. So if we want [ɨ] to surface, *we cannot have it surface by way of the phonological input /ɨ/.* While it may be possible to get around this by positing additional phonological symbols in the lexical grammar, and more complex markedness constraints (barring, say, an /ɨ/ that maps to [ɨ]) to properly manage the neutralization, it is hard to see a viable motivation for any of the stipulations involved.

Beyond these problems, the neutralization approach misses an easy generalization about transparent vowels. In Seto, as in most languages with back harmony, transparent vowels select for front affixes when they are alone in word stems. While trigger competition accomplishes this by supposing (as is already necessary) that transparent vowels are weak triggers, in a neutralization system this would mean find some means by which even though transparent vowels can become back at some stage of the derivation, underlyingly back transparent vowels are fronted before they can propagate harmony.

3.2.2 It's not underspecification

Another conventional approach to harmony, seen in Clements (1976), Kiparsky (1980), Archangeli and Pulleyblank (1994) and Ringen and Vago (1998), hypothesizes that transparent vowels are essentially unspecified or underspecified for the harmonic feature, and therefore do not need to participate in harmony. While this can work nicely in many languages, Seto makes it impossible to claim that the language does not need have a contrastive BACK feature for transparent vowels, since both have back counterparts. Maintaining such an analysis requires either abandoning the richness of the base assumption and banning input forms containing tokens of /i/ specified as [-BACK], or else allowing for and motivating the potentially troublesome operation of de-specifying specified front vowels. In addition, the peculiar harmonic alternation of /e/ and /ə/ after the otherwise transparent initial /e/, analyzed in section 3.5 below, suggests that there would need to be an additional implausible mechanism to force unspecified vowels to become specified as front *only if the nearest vowel to their right is a lax back vowel.*

This problem is not unique to Seto: Kimper describes a problem in Finnish which closely parallels the situation of Seto word-initial transparent vowels:

> For the medial [y:] in a word like [martty:ri] to be treated as transparent ([martty:ri-a], with a back suffix), it must be underlyingly unspecified along the front–back dimension. Otherwise, spreading backness from the initial [a] to the suffix vowel would run afoul of the inviolable line crossing prohibition. However, the distribution of [y] and [u] in Finnish — even in loanwords — is not predictable; the backness of a high rounded vowel is not determined by general (or even positional) well-formedness conditions, but is in fact arbitrary. This means that the [y] in words like [martty:ri] must be underlyingly specified as front. There is a fundamental incompatibility of these two requirements, and it is impossible to account for the patterns of transparency in Finnish loanwords while maintaining an inviolable prohibition on crossed association lines. (Kimper, 2011)

In the absence of either underspecification or neutralization as explanations for transparent vowels, it is necessary to look to approaches to vowel harmony which account for transparency more directly. In the following sections, I show that Kimper's Trigger Competition, which allows for non-locality (in the form of crossed association lines), accomplishes this for Seto.

3.3 Basic spreading

In this section, I demonstrate how Trigger Competition can be used to account for the simple suffixal alternations in Seto like these:

(33) a. *pan-daq* 'put-INF'

b. *müü-däq* ‘sell-INF’

(34) a. *nõsə-sə* ‘rise-3’

b. *elä-se* ‘they-3’

The parameters that I used for Seto, shown in Table 3 and subsequent tableaux, are somewhat arbitrary, and there are many generally similar such sets of values that could be used. I anchor the weight of the main stem identity constraint at -1, and the default trigger strength multiplier at 5, and set the remaining values to be clean decimal values that accommodate these.

Parameter	Value
k (distance multiplier)	0.4
$x[i]$ (trigger strength multiplier)	0.2
$x[e_{\acute{\sigma}}]$ (trigger strength multiplier)	1
$x[default]$ (trigger strength multiplier)	5

Table 3: Trigger Competition parameters for Seto.

I demonstrate this with the derivation of the harmonically straightforward [pan-daq] (‘put-INF’) from the (naively) hypothesized underlying form /pan+däq/. Harmony spreads onto the affix in one step, after which no further improvement is possible:

(35)

+ − /pan+däq/	STID[BK] −1	SPR[±BK] +.75	$\mathcal{H}$
a. + − pan+däq	0	0	0
☞ *b.* + pan+daq	0	5	3.5

3.4 Neutral vowels

3.4.1 Transparent vowels

In Trigger Competition, the same mechanism is used to account for word-initial transparency, where a word-initial vowel fails to force alternation in subsequent vowels, as in (36a) and (37a), and non-initial transparency, where a vowel in a harmonizing position fails to alternate but allows harmony from previous vowels to propagate past it, as in (38a).

(36) a. *teeda* ‘grandpa’

b. *esä* ‘father’

(37) a. *ilma* ‘without-GEN’

b. *silmä* ‘eye-GEN’

(38) a. *opp:a-ji-lə* ‘teacher-PL-ALL’

b. *rebäs-i-le* ‘foxPL-ALL’

The following example demonstrates how transparent vowels work in this system. In [opp:a-ji-lə], the final suffix harmonizes with back-harmonic stem, ignoring the intermediate front vowel /i/. In the first cycle of derivation, we add one affix and harmonize:

(39)

+ + − /opp:a+ji/	*{ö, ɨ} -20	StID[Bk] −1	Spr[±Bk] +.75	$\mathcal{H}$
a. opp:a+ji (+ + −)	0	0	0	0
☞ b. opp:a+ji (+ −)	0	0	5	3.75
c. opp:a+jɨ (+ +)	1	0	5	−16.25

Harmony occurs between the first two vowels but ignores the last: spreading harmony onto /i/ would create an illegal non-initial /ɨ/.[5] At this point, no further optimization is possible, and the derivation converges. In the next cycle, the first affix becomes part of the stem, a second affix is added, and optimization begins again.

(40)

+ − − /opp:aji+le/	*{ö, ɨ} -20	StID[Bk] −1	Spr[±Bk] +.75	$\mathcal{H}$
a. opp:aji+le (+ − −)	0	0	5	3.75
b. opp:aji+le (+ −)	0	0	5+ .2 = 5.2	3.9
☞ c. opp:aji+lə (+ −)	0	0	5+ 5 × .4 = 7	5.25
d. opp:ajɨ+le (+ −)	1	1	5+ 5 = 10	−13.5

Once again, only a single change occurs before the derivation converges: spreading onto the new affix. /a/ still cannot spread directly onto /i/, and so we see a neutral vowel. Since /a/ is a much better trigger of harmony than /i/ and can thus spread past it in spite of the distance penalty, that neutral vowel is transparent. Were further alternating affixes added to the end of the word, the grammar would continue to spread [+Back] onto those affixes from the final /ə/.

[5]Word initially, /ɨ/ and /ö/ can still surface because they are protected by a highly weighted stressed or initial or stressed syllable faithfulness constraint, which is not shown here.

In order to account for the different behaviors of /e/ in initial and non-initial positions, it is necessary to hypothesize that stressed tokens of /e/ are better cued for backness than unstressed tokens, and are thereby weaker triggers, and that the effect of stress on trigger strength is only discernible for this one vowel. I do not provide a complete justification for this claim here, but one unique property of /e/ is clearly relevant: stressed /e/ contrasts with the robustly back /õ/, while the back counterpart of unstressed /e/ is the much closer and more confusable lax vowel /ə/.

The parameters used here, shown in Table 3 above, ensure that /e/ is neutral initially, but participates normally in other positions. In the first tableau, we see that word-initial /e/ cannot force harmony on a normally alternating non-initial vowel:

(41)

− + (each linked) /e.a/	StID[Bk] −1	Spr[±Bk] +.75	$\mathcal{H}$
a. ☞ − + (each linked) e.a	0	0	0
b. − (linked to both) e̲.ä	1	1	−.25

Since /i/ has an even lower trigger strength value than /e/, it behaves similarly, and stems like [i.a] are acceptable from the point of view of harmony.

The stronger non-initial /e/, however, is able to propagate harmony from a preceding host, as shown in the following two-step derivation. In the first tableau, the /e/ comes to share a harmonic feature value with the initial syllable:

(42)

− − + (each linked) /ä.e.a/	StID[Bk] −1	Spr[±Bk] +.75	$\mathcal{H}$
a. − − + (each linked) ä.e.a	0	0	3.5
b. ☞ − (linked to ä, e) + (linked to a) ä̲.e.a	0	5	3.75
c. − (linked to ä) − (linked to e, ä) ä.e.ä	1	5	2.75

In the second step, /e/ propagates harmony on to the final vowel:

(43)

− + /ä . e̲ . a/	StID[Bk] −1	Spr[±Bk] +.75	$\mathcal{H}$
a. ☞ − ä . e̲ . ä	1	5+5=10	6.5
b. − + ä . e . a	0	5	3.75
c. − − ä . e̲ . ä	1	5	2.75

Finally, /a/ is still capable of spreading harmony normally onto non-initial /e/:

(44)

+ − /a . e/	StID[Bk] −1	Spr[±Bk] +.75	$\mathcal{H}$
a. + − a . e	0	0	0
b. ☞ + a̲ . ə	1	5	2.75

3.4.2 Transparent stems

Stems like /pet/ or /ihft́/ which contain only transparent vowels categorically select front suffixes:

(45) a. *pet-mä* 'deceive-INF'
     b. **pet-ma*

(46) a. *ihft́-mä* 'poison'[6]
     b. **ihft́-ma*

Accounting for this requires two things: I assume that even though word initial /e/ and /i/ are not strong enough triggers to overcome the faithfulness cost of triggering harmony on other stem vowels, they *do* have positive trigger strength values. I also assume that any constraints that protect the [Back] values of segments in affixes are weighted so low as to have no effect on the grammar. Otherwise, it would be necessary to claim either that transparent vowels become strong triggers when they are alone in a word, for which there is no clear perceptual motivation, or else that affixes are always underlyingly front, which would violate richness of the base. Given these two assumptions, it is better to attain a tiny reward by spreading front harmony onto the affix than it is to leave the affix back:

[6] The diacritic represents palatalization.

(47)

/i+a/	StID[Bk] −1	Spr[±Bk] +.75	$\mathcal{H}$
a. i+a	0	0	0
☞ *b.* i̲+ä	0	0.2	0.15

3.4.3 Opaque vowels

In Seto, /o/ is opaque: it can occur following front vowels, but any non-transparent vowels that follow /o/ must be back:

(48) *läh:ko-lə* 'near-ALL' [*läh:ko-le]

(49) *Höödo-kkənə* 'Teddy' ('Höödo-DIMINUTIVE') [*Höödo-kkene]

Since the constraint *ö blocks /o/ from alternating non-initially in this grammar, /o/ behaves as a neutral vowel. When a neutral vowel has a trigger strength that is similar to that of the surrounding vowels (here, 5), that neutral vowel will be opaque. This is demonstrated here for /o/ with [lähko+lə], derived from the hypothesized underlying /lähko+le/.

(50)

/lähko+le/	*{ö, ɨ} -20	StID[Bk] −1	Spr[±Bk] +.75	$\mathcal{H}$
a. lähko+le	0	0	0	0
b. lä̲hkö+le	1	1	5	−17.25
☞ *c.* lähko̲+lə	0	0	5	3.75
d. lä̲hko+le	0	0	5 × .4 = 2	1.5

In the first round of evaluation, the affix is added, and three kinds of harmony are considered. In candidate *b*, spreading from the initial front vowel leads to the creation of a non-initial [ö], which is heavily penalized. In candidate *d*, spreading from the initial syllable bypasses the neutral vowel, creating a viable candidate in which /o/ is treated as transparent. However, since /o/ is just as good a trigger as /ä/, it is able to avoid the distance penalty and accrue a greater reward than /ä/ for spreading its harmonic feature value onto the final /e/, and candidate *c* wins. At this point, no further improvement is possible according to the present constraints, and the derivation converges:

(51)

/lähko+lə/	*{ö, ɨ} -20	StID[Bk] −1	Spr[±Bk] +.75	$\mathcal{H}$
a. lähko+lə	0	0	0	0
b. lähkö+lə	1	1	5	−17.25
☞ *c.* lähko+lə	0	0	5	3.75
d. lähko+le	0	0	5 × .4 = 2	1.5

3.5 /e/–/ə/ harmony and translucent /e/

According to K&P, word initial /e/ is transparent, and can occur before non-opaque back vowels, as in (52). Furthermore non-initial /ə/ is a non-opaque back vowel, as in (53). However, the particular combination of /e/ in word-initial position and /ə/ in the second syllable position is not permitted, as in (54).

(52) *teeda* 'grandpa'

(53) *killõ* 'shrill'

(54) **kellõ*

This could easily be addressed directly using a stipulative constraint like the following:

> ***e.ə:** Assign one violation to any token of /ə/ that follows a token of /e/ at a distance of at most one syllable.

Before attempting this, it is worth looking into the possibility that this phenomenon can be accounted for without language specific mechanisms. A first attempt to do this might take advantage of the phenomenon's resemblance with parasitic harmony, a phenomena in which harmony only occurs between similar segments. Kimper addresses this in cases like Yawelmani. However, it is possible to account for this case without any direct reference to the shared phonological height of /e/ and /ə/ or any novel mechanisms or constraints.

Both a direct approach and a parasitic harmony approach miss the generalization that Trigger Competition allows us to capture. /e/ is a stronger trigger of backness harmony—a less transparent vowel—than /i/, both within Seto and crosslinguistically (Anderson, 1980). Combining this with the proposition that the laxing /ə/–/õ/ phoneme is not as well protected by faithfulness as are the rest of the harmonizing vowels yields the observed behavior.

The former can be accomplished using trigger strength multipliers, and the latter can be accomplished by the introduction of a more general new constraint which protects the backness values of tense vowels, but not lax vowels. This additional constraint is specified as follows:

(55) **IO-Ident-Stem[Back]$_{\text{tense}}$**: Assign one violation for any surface vowel which differs from its input correspondent in its specification for [±Back] and whose input correspondent is [+Tense].

Combining this with weights shown in the tableaux that follow yields the desired grammar. /e/ can force harmony on /ə/:

(56)

− + /e . ə/	StID[Bk] −0.75	Spr[±Bk] +0.75	StID[Bk]$_{\text{tns}}$ -0.25	*[-Tense] −0.1	$\mathcal{H}$
a. − + e . ə	0	0	0	1	−0.1
b. ☞ − e̲ . e	1	1	0	0	0

However the other transparent vowel, /i/, cannot force harmony on /ə/:

(57)

− + /i . ə/	StID[Bk] −0.75	Spr[±Bk] +0.75	StID[Bk]$_{\text{tns}}$ -0.25	*[-Tense] −0.1	$\mathcal{H}$
a. ☞ − + i . ə	0	0	0	1	−0.1
b. − i̲ . e	1	0.2	0	0	−0.6

And word-initial /e/ cannot force harmony on any other alternating segment, because doing so would incur an IO-Ident-Stem[Back]$_{\text{tense}}$ violation:

(58)

− + /e . a/	StID[Bk] −0.75	Spr[±Bk] +0.75	StID[Bk]$_{\text{tns}}$ -0.25	*[-Tense] −0.1	$\mathcal{H}$
a. ☞ − + e . a	0	0	0	0	0
b. − e̲ . ä	1	1	1	0	−0.25

The constraint weights here still generate the phenomena discussed in previous sections. Any candidate that received a violation of StID in the previous tableaux received a violation count of 1 on a constraint with weight −1, yielding a contribution to its total score of −1. Such a candidate now receives a violation count of 1 on StID and a violation count of 1 on StID$_{tns}$, with weights −0.75 and −0.25, respectively, yielding a contribution to its total score of −1 by the calculation $(1 \times [-0.75]) + (1 \times [-0.25]) = -1$. As a result, the grammar's treatment of tense vowels remains unchanged, and the only new predictions are confined to underlying forms containing /ə/.

3.6 Schwa

I have so far assumed that the grammar has some means of enforcing /õ/–/ə/ alternation. In this section I show that this can be accomplished without interacting adversely with harmony. My account assumes that /ə/ is specified as [+Back][7], and is differentiated from /õ/ by only one *contrastive* feature. I consider this contrast to involve [Tense], with /õ/ and all other harmonic vowels specified as [+Tense], and /ə/ specified as [-Tense]. Given this, we need only add two low-weighted constraints in order to account for the alternation. *õ straightforwardly establishes the slight increase in markedness associated with the tense variant, and *[-Tense]$_{\acute{\sigma}}$ expresses a preference for tense vowels in stressed (i.e., word initial) syllables:

(59) ***[-Tense]$_{\acute{\sigma}}$:** Assign one violation for each segment which is specified with the feature value [-Tense] in the output and which carries primary stress in the output.

It is reasonable to assume for symmetry's sake that a constraint Ident[Tense] exists in the grammar, but it is not weighted highly enough to ever impact the output in Seto.

The result of adding these two constraints is a grammar in which, regardless of harmony or any other phenomenon, mid non-round non-front vowels will be tense word-initially and lax elsewhere, regardless of underlying specification for [Tense]. This repair will not change the distribution of the harmonic feature value. However, since tensing or laxing these vowels to correct their distributions is a harmonically improving change, it is guaranteed to occur where it is warranted. This process takes place as shown in the following two examples:

(60)

(+ +) /ə . a/	...	*[-Tense]$_{\acute{\sigma}}$ -0.2	*õ −0.1	$\mathcal{H}$
a. (+ +) ə . a	...	1	0	−0.2
☞ *b.* (+ +) õ . a	...	0	1	−0.1

(61)

(+ +) /a . õ/	...	*[-Tense]$_{\acute{\sigma}}$ -0.2	*õ −0.1	$\mathcal{H}$
a. (+ +) a . õ	...	0	1	−0.1
☞ *b.* (+ +) a . ə	...	0	0	−0.0

[7] If this assumption turns out to be problematic for other reasons, this analysis can be recast without changing its essential features as long as the schwa still has the harmonic feature value opposite that of /e/—constraints referring to back need only be changed to refer to front with reversed polarity.

This repair also will not induce or allow the laxing of any other vowels. Any attempt to assign [-Tense] to some vowel other than /õ/ (or /e/, if motivated by harmony) will result in a segment that is banned by the markedness constraints that define the inventory. While [Tense] is not protected by an active faithfulness constraint, other inventory defining features are, so any attempt to transform any vowel other than /e/ or /õ/ into /ə/ will require changing that vowel's specification for height or roundness, and will fail.

3.7 Directionality: why only rightward?

Seto vowel harmony shows no evidence of leftward spreading: suffixes never force harmonic alternation in the stem or on preceding suffixes, and opaque vowels force harmony on the segments to their right, but not to their left. Kimper never directly accounts for directionality effects in harmony, but suggests that it may be necessary to do so to account for some languages:

> Finally, the issue of directionality in harmony has been set aside for the present discussion – however, Mullin (2010)[8] provides evidence that directional restrictions behave in many ways as though they are in fact influences on a trigger's strength (for example, overcoming blocking in one direction, but not the other).

I show that in order to account for directionality using a constraint like Spread in Serial Harmonic Grammar, it is necessary to take a hybrid approach. The grammar must be able to categorically ban leftward spreading—which I accomplish using Kimper's suggested mechanism of trigger strength multipliers—but it must also be able to force spreading from triggers at the left of the word to occur derivationally prior to spreading from triggers further to the right.

For an example of the problem of leftward spreading, I return to earlier derivation of the sequence /opːa+ji+le/, this time without imposing any directional restrictions on the candidates. When the first affix is added, the optimal output is one in which the /o/ and /a/ are linked regardless of any directionality limitations. However, once the second affix is added, a pathology emerges (deliberately omitted from earlier tableaux) if we permit leftward spreading:

[8] Revised and expanded as Mullin (2011). –SB

(62)

+ − − /opp:aji+le/	*{ö, ɨ} -20	StID[Bk] −1	Spr[±Bk] +.75	$\mathcal{H}$
a. opp:aji+le	0	0	5	3.75
b. opp:aji+le	0	0	5 + .1 = 5.1	3.825
☹ *c.* opp:aji+lə	0	0	5 + 5 × .4 = 7	5.25
d. opp:ajɨ+le	1	1	5 + 5 = 10	−13.5
☞ *e.* opp:aji+le	0	0	5 + 5 = 10	7.5

The final candidate wins, in which leftward spreading permits the affix to profitably harmonize with the stem-final transparent vowel, rather than with the alternating vowel. At this point, all the segments are harmonically associated and no further improvement is possible within a single step, and the ungrammatical output [oppːajile] results.

This reflects a subtle but substantial problem: in an ABC sequence where B is transparent, and A and C are disharmonic, it is always better for either AB or BC to share a feature value (whichever involves no faithfulness violation) than it is for one of A or C to have to change for the two of them share a feature value. This is demonstrated here:

(63)

+ − − /klɨbise/	StID[Bk] −1	Spr[±Bk] +.75	$\mathcal{H}$
a. klɨbisə	0	0	0
☹ *b.* klɨbisə	1	5 × .4 = 2	0.5
☞ *c.* klɨbise	0	5	3.75
d. klɨbise	0	0.2	0.15

Candidate *b* reflects the expected behavior in Seto, while candidate *c* is favored by every constraint. The easiest solution to this, at least for Seto, is to introduce an explicit notion of directionality.

3.7.1 Parameterizing directionality

I attempt to adapt Mullin's (2011) proposal to Kimper's system by introducing the multipliers l and r, which modify the reward assigned by SPREAD depending on whether the target is to the left of the right of the trigger at each instance of spreading. Since there is no evidence of leftward spreading in Seto, it is safe to set the values of these parameters to 0 and 1 respectively, preventing any reward from being assigned for leftward spreading. This yields this updated version of the above tableau:

(64)

+ − − /opp:aji+le/	*{ö, ɨ} -20	STID[BK] −1	SPR[±BK] +.75	$\mathcal{H}$
a. + − − opp:aji+le	0	0	5	3.75
b. + − opp:aj<u>i</u>+le	0	0	5 + .1 = 5.1	3.825
☞ *c.* + − opp:<u>a</u>ji+lə	0	0	5 + 5 × .4 = 7	5.25
d. + − opp:<u>a</u>jɨ+le	1	1	5 + 5 = 10	−13.5
e. + − opp:aji+l<u>e</u>	0	0	5 + 5 × **0** = 5	3.75

Here, the l multiplier eliminates the additional reward earned by candidate e, such that it is no longer optimal, while the remaining candidates are only affected by the r multiplier. Since the r multiplier is set to 1, it has no impact, and is not shown in the tableaux. This leaves the desired candidate c as the winner.

The typological effects of these proposed multipliers are largely independent of those of any other constraint or parameter, and they are summarized here:

(65) **Factorial Typology: Directionality Multipliers**

a. $l = 0 \quad r = 0$
No vowel harmony (English).

b. $l = 0 \quad r > 0$
Strictly rightward harmony (Seto).

c. $l = r \quad l, r > 0$
Non-directional harmony (bidirectional stem control or dominance harmony, described in Baković, 2000).

d. $l > r > 0$
Non-directional harmony, with the potential for vowels that are opaque to rightward spreading but transparent to leftward spreading (Dagbani and Southern Palestinian Arabic, Mullin, 2010).

This explicit directionality runs counter to claims by many, including Baković (2000), that directionality emerges naturally from morphological constituency relationships. The issue is certainly far from settled, and Hyman (2002) gives an extensive discussion of what evidence bears on directionality in vowel harmony. In any case, there are plenty of harmony systems which show patterns that appear to privilege one direction above the other stem-internally, including Khalkha Mongolian (Svantesson et al., 2005) or Jingulu (Pensalfini, 2003) as well as the two examples in Mullin. In the absence of any framework for harmony which captures our full set of observations without explicit left-to-right directionality, there is no choice but to stipulate it.

3.7.2 Left-first spreading and the HarmonizeFromLeft constraint

To capture Seto vowel harmony in a serial grammar, it is essential that harmony not only take place by way of operations that spread from a trigger on the left to a target on the right, but also that harmony proceed from left to right over the course of the derivation. The danger of left-to-right harmony that proceeds derivationally right-to-left is shown here.

In the following example, the input contains two non-initial front vowels and a word-initial back vowel. Under the existing constraints, even with purely leftward spreading, the best first step is to spread from the second syllable to the third, since that receives the same reward as spreading from the first to the second, but without incurring any new faithfulness violation.

(66)

+ − − /nose+se/	StID[Bk] −1	Spr[±Bk] +.75	$\mathcal{H}$
a. ☞ nose+se	0	5	3.75
b. ☹ nosə+se	1	5	2.75
c. nose+se	0	0	0
d. nose+sə	0	5 × .4 = 2	1.5

At this point the derivation converges: while it is still possible to initiate harmony from the initial syllable, doing so requires undoing the harmony between the subsequent syllables, and losing the associated reward. In a system which did not discriminate between directions of spreading, it would be possible to incur a further reward for spreading from the middle syllable to the initial syllable, yielding [nösese] as in *c* below; however, in the absence of that as a viable option, the winning output is disharmonic.

(67)

+ − /nose+se/	StID[Bk] −1	Spr[±Bk] +.75	$\mathcal{H}$
a. ☞ + − nose+se	0	5	3.75
b. ☹ + − nosə+se	1	5	2.75
c. − nöse+se	1	$5+0=5$	2.75
d. + − nose+sə	0	$5\times.4=2$	1.5

Sequences like [nosese] are unattested in Seto native words (c.f. *nosə-sə* 'rise-3'), and this problem carries over to any language with overtly directional harmony. Furthermore, there is no general way to solve it with featural markedness constraints: we cannot allow segmental markedness to affect the outcome of the above tableaux, since any change that caused us to prefer spreading from /u/ would only place us in an identical bind in stems with the feature values reversed, like /bübubu/.

A number of possible solutions present themselves:

One approach explicitly rewards candidates in which the initial or stressed syllable participates in harmony, either by way of a multiplier on Spread or by way of a separate positive constraint. This only pushes the problem along to more complex forms: as long as this reward can be attained by only a single instance of spreading, then the grammar will still accept forms like /bababäbä/ which are unattested in Seto, in which the first two syllables can harmonize with one another, and the last two with one another, with neither pair harmonizing with the other.

Another approach assigns an additional reward to any instance of spreading in which the trigger was associated with the initial syllable, in essence giving special privileges to the first harmony domain in the word. Again, this just pushes the problem into a slightly more specific set of cases: if the initial syllable is transparent, as in /bibubübü/, then either we predict (falsely) that it should spread and rule out real surface forms like [sina], or we don't create the special harmony domain at all, and the familiar problem of word-medial harmony emerges through forms like [bibübubu].

Another family of approaches, from which my solution is drawn, involves penalizing triggers that are not as far to the left as could be. The simplest version of such a constraint, which I will call Align(Trigger,L) is defined as follows:

(68) **Align(Trigger,L)**: Assigns as the violation count the number of segments that intervene between the current harmony trigger, if one exists, and the left edge of the word.

A constraint like this suffers from the crucial flaw that precludes the transmission of harmony across long words. In a word like /bababababä/, we would expect harmony to spread to the last syllable, but the grammar cannot produce this:

(69)

+ − /babababab ä/	StID[Bk] −1	Align(T,L) −1	Spr[±Bk] +.75	$\mathcal{H}$
a. ☞ + − babababab ä	0	0	$5 \times 3 = 15$	11.25
b. ☹ + bababab$\underline{a}$ba	1	3	$5 \times 4 = 20$	11
c. ☹ + b$\underline{a}$babababa	1	0	$5 + 5 \times .4^3 = 15.32$	10.49

Triggers far to the right like the one shown in *b* are heavily penalized by Align, and cannot beat the faithful candidate. Attempting to spread to the final syllable from an earlier trigger fails as well, as shown in *c*: while these triggers aren't penalized by Align, the distance penalty makes the reward so small as to make it impossible to overcome faithfulness within a stem. It is also not possible to avoid this problem by simply choosing a very low constraint weight for Align. In order to avoid the pathology that this constraint is meant to solve, even the difference in weight between zero violations and two (as in tableau 67) must be sufficient to sway the choice of trigger.

It is possible, though, to formulate a constraint like this which does not penalize long words unfairly. Such a constraint, rather than counting the distance between the trigger and the edge of the word, counts the distance between the trigger and the nearest harmony participant to its left. The constraint is formalized as follows:

(70) **HarmonizeFromLeft[±F]:** Assign as the violation count the number of consecutive nodes on the F tier associated with only one segment each which are to the immediate left of the left edge of any harmonic domain (any node on the F tier associated with multiple segments).

Under this definition, a word-initial trigger is assigned the exact same violation count—zero—as a non-initial trigger, provided that all the segments to the left of the non-initial trigger are participants in harmony. However, by penalizing non-initial triggers in words for which no harmony has yet occurred, the constraint promotes derivations in which harmony is propagated from left to right. With this constraint in place, the broken tableau above now produces the desired result: in the first step, the first two segments harmonize, and in the second step, harmony proceeds to the third segment:

(71)

+ − − /nose+se/	FromLeft −1.25	StID[Bk] −1	Spr[±Bk] +.75	$\mathcal{H}$
a. + − nosə+sə	1	0	5	2.5
☞ *b.* + − nosə+se	0	1	5	2.75
c. + − nose+sə	0	0	$5 \times .4 = 2$	1.5
d. + − − nose+se	0	0	0	0

(72)

+ − /nosə+se/	FromLeft −1.25	StID[Bk] −1	Spr[±Bk] +.75	$\mathcal{H}$
a. + − nosə+se	0	1	5	2.75
☞ *b.* + nosə+sə	0	1	$5 + 5 = 10$	6.5

Initial transparent vowels are not a problem: though sequences like /vii:ra+mä/ 'leave.a.wake-INF' require the grammar to incur a violation of HarmonizeFromLeft in order to enforce harmony between the second and third syllables, the difference in trigger strengths between transparent and non-transparent vowels is sufficient to ensure that harmony takes place as expected nonetheless:

(73)

− + − /vii:ra+mä/	FromLeft −1.25	StID[Bk] −1	Spr[±Bk] +.75	$\mathcal{H}$
☞ *a.* − + vii:ra+ma	1	0	5	2.5
b. − − vii:rä+mä	0	1	.2	−.85
c. − + vii:ra+mä	0	0	$.2 \times .4 = .08$	.06
d. − + − vii:ra+mä	0	0	0	0

This new constraint achieves the same purpose as a precedence constraint, of the sort used in Optimality Theory with Candidate Chains (OT-CC, McCarthy, 2007), in that it exerts pressure on the grammar to perform operations in a certain order. However, by referring to the structures that the grammar builds (autosegmental links) rather than the operations of the grammar themselves, HarmonizeFromLeft avoids the potential learning-theoretic pitfalls of constraints that refer to the derivation directly.

3.8 What can be a trigger?

Kimper's formalism requires that a vowel can only be a trigger if it is the *linearly closest segment to the target which is linked to its particular harmonic feature node* (12). What this means is that, if several segments have already harmonized and share a harmonic feature node, then only the rightmost of those segments is eligible to serve as a trigger for further harmony. Given the input (74), only the output (75) is possible, and not (76):

(74) n o̲ s ə s e (+ linked to o, ə; − linked to e)

(75) n o s ə̲ s ə (+ linked to o, ə, ə)

(76) * n o̲ s ə s ə (+ linked to o, ə, ə)

This restriction was proposed to account for a real phenomenon, but I see no clear a priori reason why it must be a part of the framework. I show in this section that such a restriction prevents the spread of harmony in a very common case, and cannot be maintained in full.

In the example nonce word input /bäbiba/ we expect harmony to force agreement between the first and last syllables, yielding [bäbibä]. I show here that harmony of this kind, involving spreading from a front vowel–transparent vowel sequence, is not possible without relaxing the conditions on triggers.

In the first step of the derivation of this example, we see that both relevant constraints favor agreement between the first and second syllables. Doing so yields a full reward for harmony without incurring any faithfulness violations:

(77)

− − + /bäbiba/	StID[Bk] −1	Spr[±Bk] +.75	$\mathcal{H}$
a. ☞ bä̲biba (− linked to ä, i; +)	0	5	3.75
b. bä̲bibä (− −)	1	5 × .4	.5
c. bäbiba (− − +)	0	0	0

At this point, the only viable option that remains is to spread harmony on to the final syllable. If we follow Kimper's restriction on triggers, then only /i/ can serve as a trigger for this operation. However, as we know from many examples elsewhere in the language, the sequence /i.a/ is allowed, and /i/ is not a strong enough trigger to overcome faithfulness and force harmony on subsequent stem vowels. Using a trigger strength consistent with this observation leaves us with competition between candidates *a* and *b*, with *c* excluded as a representationally impossible candidate. Under any viable constraint weighting that allows for transparency, *a* wins, and the derivation converges, leaving us with the illegal output [bäbiba]:

(78)

− + /bäbiba/	StID[Bk] −1	Spr[±Bk] +.75	$\mathcal{H}$
☞ *a.* − + bäbiba	0	5	3.75
b. − bäbi̲bä	1	5 + 0.2	2.9
☹ *c.* − (bä̲bibä)	1	5 + 5 × 0.4 = 7	4.25

The only way around this problem is to relax this restriction, and allow the grammar to generate candidate *c*, in which a non-domain-final vowel is treated as the harmony trigger. If this is done, and if the reward for spreading is assigned accordingly, then *c* wins, yielding the desired output.

Precisely this problem seems to emerge in Hungarian as well. Unfortunately, as with Seto, I have not found any attested uninflected roots in Hungarian with front-transparent-front or front-transparent-back sequences (the non-transparent front vowels /ö/ and /ü/ are relatively rare in Hungarian). However, under every analysis of Hungarian that I can find, (e.g., Gafos and Benus, 2006, Ringen and Vago, 1998, Booij, 1984) sequences like /ö.i.ü/ are possible, while /ö.i.u/ are acceptable only in loanwords, for which root-internal harmony is not enforced. This yields the same contradiction we see in Seto: these observations force us to claim that /i/ can propagate front harmony on its own if we follow Kimper's restriction, in spite of many examples of words like *csipa* 'gum' in which it clearly does not.

Kimper's restriction does have some empirical basis, and fixing this problem is not as simple as merely eliminating the restriction altogether: he documents a case of *icy targets* in Khalkha Mongolian rounding harmony (Svantesson et al., 2005). In this harmony system, we see exactly the phenomenon represented by candidate *a* in the previous tableau: vowels that allow harmony for the opposite feature value to spread past them, but which do not allow harmony for their own feature value to spread past them, instead blocking the harmony and allowing vowels of either harmonic feature value to follow. This is no comparably straightforward way to generate such a system under Trigger Competition without somehow restricting candidates like *c*.

Because of this, I propose another language specific weighting factor for Spread: in instances of spreading where the trigger is not the closest member of its harmonic domain to the target, the reward associated with spreading is multiplied by a factor k_{linked} once for each linked segment it must spread past, in addition to being multiplied by the already defined general distance factor k. For languages like Seto and Hungarian, this factor can be one, preventing it from affecting harmony. For languages like Khalkha, it can be zero, and serve as a categorical ban on that kind of spreading. Judging from these two examples I see no reason to hypothesize that the parameter can take on any value other than one and zero, but I also do not have sufficient evidence to claim that a language with an intermediate value—in which targets can be icy only after weak triggers—does not exist.

3.9 How is trigger strength determined?

One of the greater sources of complexity in Kimper's new framework is the proposal that each vowel has two related but separately specified grammatical properties: segmental markedness, specified by markedness constraints, and trigger strength, specified by the trigger strength factors, neither of which is specified clearly enough to be extracted from the data mechanically. While I do not have enough crosslinguistic evidence to robustly claim that these two parameters are related, I show that for Seto, using the same values for both only simplifies the grammar.

I consider the following proposal, and show that it is adequate for Seto:

(79) **Markedness as strength:** The trigger strength multiplier associated with a segment in a given position is equal to the total violation score of all the context-free markedness constraints penalizing that segment.

For Seto, it is necessary to break the feature values of some segments down by stress: /e/ clearly behaves differently in stressed (word-initial) positions than it does elsewhere. For /ö/ and /ɨ/, the case is somewhat less clear. Simply assigning these two vowels high values in all positions would ensure that they are correctly neutralized non-initially while making them strong harmony triggers. However, there is reason to believe that they are not *stronger* harmony triggers word-initially than are the other alternating values: if /ö/ in particular were substantially stronger, we would expect the front vowel /ö/ to spread front harmony past opaque segments like /o/, which does not occur.

The absolute size of these values is entirely arbitrary. However, the relative scale defined by these values is necessary to account for Seto, and once the constraint weights are fixed, the values are fairly tightly constrained. While I use segmental markedness constraints that each penalize exactly one vowel type, it is likely that one could build a working set of scores by replacing these specific constraints with carefully weighted general constraints like *[+FRONT, +ROUND] or *[-TENSE]$_{\acute{\sigma}}$. Using the constraint values shown elsewhere in this paper, I selected the parameter values shown in Table 4. These parameters generate the correct output for every example shown in this paper, and several dozen additional diagnostic inputs.

I demonstrate the operation of a grammar with this merger below. In the interest of saving space, all fifteen of the segmental markedness constraints are represented in a single column of the tableau, labeled SEGMARK. The violation score for that column is the sum of all the segmental markedness violations incurred by the word.

Parameter	Value
k (distance multiplier)	0.4
k_{linked} (linked distance multiplier)	1
l (directional multiplier)	0
r (directional multiplier)	1
Markedness/strength:	
ɨ$_{\breve{\sigma}}$, ö$_{\breve{\sigma}}$	20
õ	5.1
ɨ$_{\acute{\sigma}}$, ö$_{\acute{\sigma}}$, ə$_{\acute{\sigma}}$, ü, ä, ə, u, a, o	5
e$_{\breve{\sigma}}$	4.9
e$_{\acute{\sigma}}$	1
i	0.2

Table 4: Parameter values for Seto under the Markedness as Strength hypothesis.

I demonstrate this here with the test case [oppːa+ji+lə]. In the first step, the non-initial /a/ is linked to the initial /o/, leaving the segments of the word unchanged. The alternative outputs *c* and *d*, in which the affixal /i/ is linked to one of the stem segments and becomes back, dramatically increase the word's SegMark violation score, and are ruled out:

(80)

+ + − / opp : a+j i /	FromLeft −1.25	SegMark −1	StID[Bk] −1	Spr[±Bk] +.75	$\mathcal{H}$
a. + + − opp : a+j i	0	5 + 5 +.2 = 10.2	0	0	−10.2
☞ *b.* + − o̲pp : a+j i	0	5 + 5 +.2 = 10.2	0	5	−6.45
c. + + opp : a̲+j ɨ	1	5 + 5 +20 = 30	0	5	−27.5
d. + + o̲pp : a+j ɨ	0	5 + 5 +20 = 30	0	5 × .4 = 2	−28.5

At this point the derivation converges, and the next affix is added. As expected, and as in the previous derivations of this form in 3.4.1 and 3.7, harmony skips the first affix to spread to the second, replacing an /e/ with a slightly more marked /ə/:

(81)

+ − − /opp:aj i+le/	FromLeft −1.25	SegMark −1	StID[Bk] −1	Spr[±Bk] +.75	$\mathcal{H}$
a. + − − opp:aj i+le	0	5 + 5 + .2 +4.9 = 15.1	0	5	−11.35
b. + − opp:aj i̲+le	0	5 + 5 + .2 +4.9 = 15.1	0	5 + .1 = 5.1	−11.275
☞ *c.* + − opp:a̲j i+lə	0	5 + 5 + .2 +5 = 15.2	0	5 + 5 × .4 = 7	−9.95
d. + − opp:a̲j ɨ+le	0	5 + 5 + 20 +4.9 = 34.9	1	5 + 5 = 10	−28.4

3.9.1 Regulating the inventory

In this section, I describe an unusual but not pathological set of interactions between the Spread constraint and the inventory-defining markedness constraints which emerges under the *markedness as strength* hypothesis. An analysis like this which proposes that segments can rewarded for the very same properties that make them marked can potentially interfere with the mechanisms that control the segmental inventories of a language. I show here that, because of the way that Trigger Competition rewards harmony triggers, this does not ultimately prevent either the systematic limitation of inventories or the systematic application of vowel harmony.

Under the weighting shown here, out-of-inventory segments are thoroughly banned, and do not surface or interfere with harmony. In the following examples, I assume that the vowel feature system used here is capable of describing nasalized vowels, but that these are disallowed by the phonology of Seto. Further, I assume that they are banned by a segmental markedness constraint with a very high weight, arbitrarily chosen here to be 500. Since nasal vowels do not surface, I assume that no highly weighted faithfulness constraint protects the nasality specification of vowels.

In the following tableau, I show that under this weighting, it is far better to correct an out-of-inventory segment to bring it into the inventory than to allow that segment to participate in harmony:

(82)

+ − − /o . ã̃ . ä/	FromLeft −1.25	StID[Bk] −1	SegMark −1	Spr[±Bk] +.75	$\mathcal{H}$
☞ a. + − − o . ä . ä	0	0	5 + 5 +5 = 10	0	−10
b. + − o . ã̲̃ . ä	1	0	5 + 500 +5 = 510	500	−136.25
c. + − o̲ . ã . ä	0	1	5 + 500 +5 = 510	5	−507.25

The output to the first step of the derivation, /o.ä.ä/, is disharmonic, but the derivation has not yet converged. In the remaining steps, the initial back vowel will spread harmony onto the succeeding syllables, creating the grammatical output [o.a.a].

Somewhat counterintuitively, if we hypothesize a language in which the harmony-enforcing constraint, Spread, is weighted more highly than markedness, then we get a language that does not show normal categorical harmony on the surface. In this language, harmony occurs normally with licit inputs, but out-of-inventory segments in the inputs can create various kinds of exceptions. Crucially, though, even such a grammar does not allow out-of-inventory segments to surface.

In such a language, the derivation for an input with an illicit segment begins with harmony, with spreading commencing from the illicit trigger:

(83)

+ − − /o . ã̃ . ä/	Spr[±Bk] +2	FromLeft −1.25	StID[Bk] −1	SegMark −1	$\mathcal{H}$
a. + − − o . ä . ä	0	0	0	5 + 5 +5 = 10	−10
☞ b. + − o . ã̲̃ . ä	500	1	0	5 + 500 +5 = 510	488.75
c. + − o̲ . ã . ä	5	0	1	5 + 500 +5 = 510	−501

However, once all spreading from that segment is complete, it becomes optimal to correct the inventory. Since the trigger strength reward in Trigger Competition reflects the strength of the trigger that initially caused harmony, and not the strength of that segment in the final output, correcting the inventory does not diminish the large reward associated with harmony:

(84)

+ − /o . ä̃ . ä/	Spr[±Bk] +2	FromLeft −1.25	StID[Bk] −1	SegMark −1	$\mathcal{H}$
a. + − o . ä̃ . ä	500	1	0	5 + 500 +5 = 510	488.75
☞ *b.* + − o . ä . ä	500	1	0	5 + 5 +5 = 15	983.75

At this point, though, the existing harmony links are so robustly rewarded that no normal harmony process can break them, and it is not profitable to produce any more links. This yields a disharmonic [o.ä] sequence. However, in this particular case, leftward spreading is worthwhile in order to resolve a FromLeft violation, even though it is not rewarded by Spread. This yields a harmonic result, but one in which harmony is not initiated by the initial syllable:

(85)

+ − /o . ä . ä/	Spr[±Bk] +2	FromLeft −1.25	StID[Bk] −1	SegMark −1	$\mathcal{H}$
a. + − o . ä . ä	500	1	0	5 + 5 +5 = 15	983.75
b. + − o̲ . a . ä	5	0	1	5 + 5 +5 = 15	−6
☞ *c.* − ö . ä̲ . ä	500 + 0 = 500	0	1	5 + 5 +5 = 30	984

3.9.2 Conclusions

This approach simplifies the grammar for Seto, but may not be principled in every case. The system presented here has no mechanism for differentiating between trigger strength values for different feature dimensions, but since Kimper suggests that the trigger strength values have to do with the strength of the cues contrasting the two values of the harmonic feature, it is reasonable to imagine that such a mechanism might need be introduced for a language like Khalkha Mongolian with robust harmony systems along multiple feature dimensions.

This is only a first step in the direction of formalizing the notion of trigger strength. Kimper provides no direction to the reader attempting to build an analysis using trigger strength, so such a proposal is of value if only as a viable null hypothesis for future research.

3.10 Remaining issues in Seto vowel harmony

Two classes of clitics participate in harmony. Most clitics participate fully in harmony, and participation is largely predictable on the basis of syntactic function. However, the generalization emerges that clitics with the vowels /ə/ or /e/ always participate. It might be possible to enforce this pattern by claiming that some faithfulness constraint that generally protects clitics from harmony does not protect lax vowels, or even potentially vowels harmonically paired with lax vowels.

A related pattern appears elsewhere in the morphology. Suffixes beginning with /k/ generally do not participate in harmony, except if the vowel following the /k/ is /ə/ or /e/, in which case they categorically do participate. These suffixes may be idiosyncratically treated as clitics in the morphophonology, allowing them to inherit the behavior just described. Without a more thorough understanding of the typology and morphosyntax of these affixes and clitics, I do not attempt an analysis.

Seto shows two other types of lexical exception to harmony that I have not yet addressed, both involving stress. Certain affixes carry their own initial stress, and are reliably non-alternating:

(86) -*ˈidə* ‘GEN.PL’
   a. *ilos-ˈidə*
   b. *verev-ˈidə*

(87) -*ˈminə* ‘NMLZ’
   a. *tege-ˈminə* ‘doing’
   b. *näge-ˈminə* ‘seeing’

I follow Kiparsky and Pajusalu’s (2001) claim that these affixes are represented phonologically as separate words, and as such, are not attached to their semantic/syntactic host words until after harmony computations take place.

A few non-derived words also carry an audible non-initial stress, and show a break in harmony immediately before the stressed syllable:

(88) *ˈineˈminə* ‘human being’

The hypothesis here is the same. These anomalous words are stored in the lexicon similarly to the English *haberdasher*: they consist of multiple phonological words, despite admitting no syntactic or semantic analysis.

4 General Conclusions

This paper accomplishes three things: it describes an implementation of the Trigger Competition framework for vowel harmony, complete with a new treatment of directionality effects and a repaired treatment of word-medial transparent vowels. It presents the case of Hungarian, in which Trigger Competition directly captures an interaction between distance and trigger quality which has required complex and stipulative treatments in earlier frameworks. It also presents the case of Seto vowel harmony, a previously undescribed system, showing that the framework can be made to capture the complexity of a full harmony system, and that it makes correct predictions about the behavior of otherwise problematic paired neutral vowels.

The constraint-based grammar for harmony that emerges in Trigger Competition is quite complex, with parameters more or less directly capturing certain facets of harmony. Were there any simpler system of constraints which could capture or even reasonably approximate all the observed data, this would be cause for concern. However, I hold that the phenomena behind vowel harmony, broadly viewed, are not simple, and that no simple analysis can succeed. Given the preponderance of evidence for a constraint-based approach to phonology, and the failure of any substantially simpler approach, Trigger Competition appears to be the most viable framework for vowel harmony yet proposed.

This paper leaves a clear direction for several possible future studies: case studies of other languages would clarify a number of issues, especially those surrounding the selection of trigger strengths, and a close look at a non-Finno-Ugric language could help to determine the typological status of the phenomena discussed here. A study of languages with unconventional blocking effects or other kinds of stem-internal disharmony could shed light on directionality and the status of the FromLeft constraint. Finally, it would be valuable to investigate possible stochastic implementations of this theory (possibly following Jesney, 2007), especially with an eye towards modeling the precise patterns of variation shown in Hungarian vacillation.

References

Lloyd Anderson. Using asymmetrical and gradient data in the study of vowel harmony. *Issues in Vowel Harmony. Amsterdam: John Benjamins*, pages 271–340, 1980.

Diana Archangeli and Douglas Pulleyblank. *Grounded phonology*. MIT Press, 1994.

Emmon Bach. Two proposals concerning the simplicity metric in phonology. *Glossa*, 2(2): 128–149, 1968.

Eric Baković. *Harmony, Dominance and Control*. PhD thesis, Rutgers University, 2000.

Eric Baković and Colin Wilson. Transparency, strict locality, and targeted constraints. In *Proceedings of the West Coast Conference on Formal Linguistics*, volume 19, page 43. Linguistics Department, Stanford University, 2000.

Max Bane, Jason Riggle, and Morgan Sonderegger. The VC dimension of constraint-based grammars. *Lingua*, 120(5):1194–1208, 2010.

Jill Beckman. *Positional faithfulness*. PhD thesis, University of Massachusetts, Amherst, 1998.

Stefan Benus, Adamantios Gafos, and Louis Goldstein. Phonetics and phonology of transparent vowels in Hungarian. In *Proceedings of the 29th Annual meeting of the Berkeley Linguistic Society*, pages 485–497, 2003.

Geert Booij. Neutral vowels and the autosegmental analysis of Hungarian vowel harmony. *Linguistics*, 22(5):629–642, 1984.

George N. Clements. Neutral vowels in Hungarian vowel harmony: An autosegmental interpretation. In *Proceedings of the Seventh Annual Meeting of North Eastern Linguistic Society*, pages 49–64, 1976.

Adamantios Gafos and Stefan Benus. Dynamics of phonological cognition. *Cognitive Science*, 30(5):905–943, 2006.

John Goldsmith. *Autosegmental Phonology*. Garland Publishing, 1976.

Bruce Hayes and Zsuzsa C. Londe. Stochastic phonological knowledge: The case of Hungarian vowel harmony. *Phonology*, 23(01):59–104, 2006.

Larry Hyman. Is there a right-to-left bias in vowel harmony. In *9th International Phonology Meeting, Vienna*, 2002.

Karen Jesney. The locus of variation in weighted constraint grammars. Poster presented at the Workshop on Variation, Gradience and Frequency in Phonology, Stanford University, 2007.

Wendell A. Kimper. *Competing triggers: Transparency and opacity in vowel harmony*. PhD thesis, University of Massachusetts Amherst, 2011.

Paul Kiparsky. Remarks on the metrical structure of the syllable. *Phonologica*, 1980.

Paul Kiparsky. Opacity and cyclicity. *The Linguistic Review*, 17, 2000.

Paul Kiparsky and Karl Pajusalu. Seto vowel harmony and the typology of disharmony. Unpublished ms., Stanford University, 2001.

Paul Kiparsky and Karl Pajusalu. Towards a typology of disharmony. *The Linguistic Review*, 20, 2003.

John J. McCarthy. *Hidden generalizations: Phonological opacity in Optimality Theory*. Equinox, 2007.

Kevin Mullin. Strength in harmony. *Ms., University of Massachusetts, Amherst*, 2010.

Kevin Mullin. Strength in harmony systems: Trigger and directional asymmetries. *Ms., University of Massachusetts Amherst*, 2011.

Joe Pater. Serial harmonic grammar and Berber syllabification. In *Prosody Matters: Essays in Honor of Lisa Selkirk*. London: Equinox Publishing, 2010.

Joe Pater, Paul Boersma, and Andries Coetzee. Lexically conditioned variation in harmonic grammar. *OCP-5. Université de Toulouse-Le Mirail*, 2008.

Rob Pensalfini. *A Grammar of Jingulu: An Aboriginal language of the Northern Territory*. Pacific linguistics, Research School of Pacific and Asian Studies, Australian National University, 2003.

Catherine Ringen and Robert Vago. Hungarian vowel harmony in optimality theory. *Phonology*, 15(3):393–416, 1998.

Jan-Olof Svantesson, Anna Tsendina, Anastasia Mukhanova Karlsson, and Vivan Franzen. *The phonology of Mongolian*. Oxford University Press, 2005.

Rachel Walker. *Nasalization, neutral segments, and opacity effects*. PhD thesis, University of California, Santa Cruz, 1998.

Rachel Walker. Consonantal correspondence. In *Proceedings of the Workshop on the Lexicon in Phonetics and Phonology*, 2001.