Relating application frequency to morphological structure: the case of Tommo So vowel harmony

August 2014
Laura McPherson and Bruce Hayes
Dartmouth College and UCLA

Abstract

We describe three vowel harmony processes of Tommo So (Dogon, Mali) and their interaction with morphological structure. The verbal suffixes of Tommo So occur in a strict linear order, establishing a Kiparskian hierarchy of distance from the root. This distance is respected by all three harmony processes; they “peter out”, applying with lower frequency as distance from the root increases. The function relating application rate to distance is well fitted by families of sigmoid curves, declining in frequency from one to zero. We show that, assuming appropriate constraints, such functions are a direct consequence of the theory of Harmonic Grammar. The crucially conflicting constraints are FAITHFULNESS (violated just once by harmonized candidates) and a scalar version of AGREE (violated 1-7 times, based on closeness of the target to the root). We show that our model achieves a close fit to the data while a variety of alternative models fail to do so.
1. Introduction

In Tommo So, a Dogon language of Mali, there are three vowel harmony processes, all of them optional. Our focus is on the fact that for all three processes, application frequency interacts with morphological structure: intuitively, harmony “peters out”, in the sense that it applies with gradually diminishing frequency in outer layers of the morphology (where layers are defined by the possibilities of affix order, following Kiparsky 1982). Each process peters out at a different rate: Low Harmony gives out earliest (applying only within the innermost morphological layers), Backness Harmony next, and ATR Harmony last.

We propose a formal model of the Tommo So harmony pattern, drawing on two essential aspects of phonological theory: Harmonic Grammar (Legendre, Miyata, and Smolensky 1990, Smolensky and Legendre 2006, Pater 2009, Jesney 2010, Potts et al. 2010, Jesney and Tessier 2011), and the theory of scalar constraints (Prince and Smolensky 1993; Frisch, Broe, and Pierrehumbert 1997, 2004; Flemming 2001). We show that this model achieves a far better fit to the data than a number of alternatives, notably the widely-cited inverted exponential model of Guy (1991).

Our paper is structured as follows: In §2 we review a fundamental generalization, due to Kiparsky, that establishes the theoretical context of our work. §3 gives basic background (vowel inventory and verbal morphology of Tommo So), then describes the pattern of affix ordering used to motivate the morphological layers. In §4, we define and exemplify the three vowel harmony processes, and in §5 give the quantitative data that we will attempt to model. §6 applies Guy’s multiplicative model to our data and demonstrates that it provides a poor fit. §7 is the main analytic section: we propose a set of constraints and deploy them in a Harmonic Grammar framework, achieving a greatly improved fit to the data. We also show that our approach is highly restrictive: only certain frequency patterns can be generated, and these include patterns that match closely to our empirical findings. §8 reviews other linguistic phenomena that show similar quantitative patterns, while §9 addresses theories that can describe our data but are not restrictive. §10 covers residual issues, and §11 concludes.

2. The theoretical context: affix order and phonological process applicability

Our work pursues a fundamental insight from Kiparsky (1982): affix ordering and the applicability of phonological processes tend to be closely correlated. Specifically, the affixes that occur “closer to the root” (as diagnosed by ordering tests) are the ones that characteristically trigger or undergo more phonological processes. Kiparsky suggested that this correlation is “a general property of languages” (p. 11), and in light of this proposed a Strong Domain Hypothesis (1984:142): “at lower levels of the lexicon and in the postlexical phonology rules may be ‘turned off’ but no new ones may be added.” Our Tommo So data follow this pattern, but in a gradient way (“turning off” is gradual); thus we will be treating the Kiparskian correlation in somewhat different theoretical terms.

We emphasize that the concept of root closeness on which Kiparsky relies is abstract: root closeness is not the literal distance seen within individual forms; rather it is calculated by first examining the morphology as a system. To give Kiparsky’s example (1982:11): English has two
negative prefixes, non- and in-, of which the latter may occur as il- by assimilation, as in illegible. When the two prefixes cooccur, non- may precede in-, but not vice versa: non-illegible but not *in-non-legible. This fact is reflected in the phonology: the process that assimilates /n/ to [l] before [l] is applicable to in- (illegal, illegible) but not to non- (*nol-legible). The essential correlation is between morphological distance — as reflected in the general affix ordering principles — and phonological process application. It remains true that linearly, the non- of non-legible is just as close to the root (i.e., adjacent) as the il- of illegible; the essential criterion of morphological distance is inferred from examination of the morphology as a whole.

We will see that Tommo So constitutes a far more elaborate case than English of the correlation of “closeness” as diagnosed by affix order and phonology.

3. Background on Tommo So

3.1 Language and data sources

Tommo So is spoken by about 60,000 people living on the Bandiagara Escarpment in the Mopti Region of Mali (Hochstetler et al. 2004). It is documented in a reference grammar by McPherson (2013), which gives background on the phonemic system and morphology. The data for this article were gathered by McPherson in Mali during a total of 14 months of fieldwork (2008-2012). There were four primary consultants, all from the commune of Tédié; their speech appears to be relatively uniform, and it is reasonable to consider the data as reflecting one single dialect of Tommo So.

There exists no large corpus of Tommo So language material. The data we used were obtained by combing through the entirety of McPherson’s field materials (consisting of an extensive lexicon, example sentences, and a variety of narratives, traditional stories, and conversations) for words containing any of the suffixes under consideration (§3.3); if the form contained no suffix vowel (due to vowel hiatus resolution, (11)), it was deemed uninformative and not included. We mined this data corpus (2818 forms) for the statistical generalizations given below. While the corpus is hardly a random sample, there is no particular reason to expect that the words obtained were somehow biased concerning vowel harmony; most were elicited for other reasons, such as analysis of the morphology or of the tonal system. Stylistically, the corpus is fairly uniform: 78% of the forms come from elicited material. The full corpus is available in the online Supplemental Materials.1

3.2 The vowel system

The Tommo So vowel inventory is given in (1):

---

1 For the submitted version of this paper, these materials may be found at https://sites.google.com/site/linguisticsupplements/.
(1) *Tommo So* vowels

<table>
<thead>
<tr>
<th></th>
<th>i</th>
<th>(i)</th>
<th>u</th>
<th>iː</th>
<th>uː</th>
</tr>
</thead>
<tbody>
<tr>
<td>e</td>
<td>o</td>
<td>eː</td>
<td>oː</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ε</td>
<td>ω</td>
<td>εː</td>
<td>ωː</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>aː</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The seven vowels [i e ɛ a ɔ o u] form a contrasting set, for which minimal and near-minimal sets are given in (2a). The vowels may also occur long, for which examples are given in (2b). Long and short vowels of the same quality behave identically in harmony. Sound files illustrating the vowels may be found in the Supplemental Materials.

(2) Examples of *Tommo So* vowels

<table>
<thead>
<tr>
<th></th>
<th>a.  /i/</th>
<th>[bɪl] ²</th>
<th>‘ladder’</th>
<th>b. /iː/</th>
<th>[gɪːr]</th>
<th>‘talisman to stop bleeding or diarrhea’</th>
</tr>
</thead>
<tbody>
<tr>
<td>/e/</td>
<td>[bɛl]</td>
<td>‘grass’</td>
<td>/eː/</td>
<td>[dɛːlɛ]</td>
<td>‘goat’s waddle’</td>
<td></td>
</tr>
<tr>
<td>/e/</td>
<td>[bɛl]</td>
<td>‘animal’</td>
<td>/eː/</td>
<td>[dɛːlɛ]</td>
<td>‘bring’</td>
<td></td>
</tr>
<tr>
<td>/a/</td>
<td>[kidɛ bʊl]</td>
<td>‘gathered thing’</td>
<td>/aː/</td>
<td>[dɔːlɛ]</td>
<td>‘sweep a little’</td>
<td></td>
</tr>
<tr>
<td>/ɔ/</td>
<td>[anə bol]</td>
<td>‘mouth sore’</td>
<td>/ɔː/</td>
<td>[dɔːlɔ]</td>
<td>‘rooster’s waddle’</td>
<td></td>
</tr>
<tr>
<td>/o/</td>
<td>[bɔl]</td>
<td>‘sweep up (village)’</td>
<td>/ɔː/</td>
<td>[dɔːlɔ]</td>
<td>‘foot chain’</td>
<td></td>
</tr>
<tr>
<td>/u/</td>
<td>[bʊl]</td>
<td>‘small pox’</td>
<td>/ʊː/</td>
<td>[dʊːlɔ]</td>
<td>‘twin’</td>
<td></td>
</tr>
</tbody>
</table>

In addition, there often appears a reduced vowel, [i], which is limited to medial syllables (initial and final high vowels are not reduced). In this environment, there is no phonemic contrast among the high vowels [i u i]; what one hears is often phonetically intermediate. A rough characterization of the facts is as follows. In faster speech [i] normally appears, whereas in slower speech the output tends to be closer to [i] or [u]. The latter choice is determined in part by whether the reduced vowel is in a root or in a suffix. In roots, the determining factor is normally the backness of the preceding vowel ([i e ɛ] tend to favor [i]-like qualities; [u o ɔ] favor [u]-like qualities). When the root vowel is [a], the quality of the reduced vowel is influenced by the place of articulation of neighboring consonants (labials tend to prefer [u], coronals [i]). In the three suffixes that include a reduced vowel (Reversive, Transitive, and Mediopassive; §3.3), the reduced vowel tends to surface in slower speech as [i], though both [i] and coarticulatorily-induced [u] are also observed.

Phonologically, vowel reduction is important because reduced vowels are transparent to harmony. As far as the underlying representations of reduced vowels, this is indeterminate (since their backness and rounding are not contrastive); somewhat arbitrarily we will depict UR’s with their most typical surface vowel.

---

² Tone is phonemic, contrasting High, Low, and toneless; see McPherson (2013, ch. 4).
We adopt the following, fairly standard feature assignments for the Tommo So vowels; 0 indicates underspecification and [ATR] denotes Advanced Tongue Root (Halle and Stevens 1969). The feature [reduced] is employed as an ad hoc stand-in; in a fully formalized theory reduced vowels would be identified by their weak position in metrical structure.

\begin{tabular}{|c|c|c|c|c|}
\hline
\hline
i & + & - & - & 0 & - \\
\hline
e & - & - & - & + & - \\
\hline
ɛ & - & - & - & - & - \\
\hline
a & - & + & 0 & 0 & - \\
\hline
ɔ & - & - & + & - & - \\
\hline
o & - & - & + & - & - \\
\hline
u & + & - & + & 0 & - \\
\hline
i & + & - & 0 & 0 & + \\
\hline
\end{tabular}

The feature [round] is entirely predictable in its distribution, since vowels are [+round] when [+back] and otherwise [−round]; for brevity we omit the straightforward rules or constraints that would be needed to fill in this value. It is sensible to treat [back] as the phonologically active feature; [a] turns out to be a non-trigger of Backness Harmony (see §4.3 below), and this may plausibly be related to the fact that it is phonetically neither front nor back. We leave [ATR] unspecified in nonmid vowels simply as an expression of agnosticism; we have no phonological or phonetic evidence to justify a classification.

### 3.3 Tommo So verbal morphology

In this section we give an overview of Tommo So verbal morphology, focusing on the key suffixes that demonstrate the affiliation of affix ordering and harmony application. For fuller discussion of these suffixes, see McPherson (2013:§11-12). For simplicity at this preliminary stage we give examples that happen not to involve vowel harmony.

#### 3.3.1 Derivational suffixes

The Factitive, which we treat as underlying /-ndɛ/, is used for two main purposes: it derives transitive verbs from intransitive (often with causative meaning), as in [dzímɛ] ‘be injured’ ~ [dzímɛ-ndɛ] ‘injure’. It can also be used to derive inchoative verbs from adjectives (always in conjunction with the Mediopassive suffix), as in [pálά] ‘long’ ~ [pálɑ-nd-ijɛ]³ ‘become long’.

The Reversive suffix is /-ilɛ/, as in [dèbɛ] ‘get stuck’ ~ [dèb-ilɛ] ‘get unstuck’.

³ Underlyingly /pálɑ-nd-ijɛ/; for the hiatus resolution seen in this and other examples see §4.2.
Transitive /-ɪrɛ́/ denotes that the subject is performing the action of the verb to or on someone else, as in [tɛ́mbɛ́] ‘find oneself in a situation’ ~ [tɛ́mb-ɪrɛ́] ‘make somebody find something’.

Mediopassive /-ɪjɛ́/ denotes that the subject is performing the action on herself; thus [pɛ́ndɛ́] ‘spread out (objects)’ ~ [pɛ́nd-ɪjɛ́] ‘spread selves out (said of a group)’.

The Causative suffix is /-mɔ́/; an example is [sɛ́mɛ́] ‘slaughter’ ~ [sɛ́mɛ́-mɔ́] ‘make slaughter’.

### 3.3.2 Inflectional suffixes

Only one inflectional suffix undergoes vowel harmony: the Defocalized Perfective, which is the version of the perfective employed when some element in the clause other than the verb is focused. This suffix has two allomorphs, /-ɪ/ and /-ɛ/. Their distribution is somewhat complex, and the harmonic behavior of the two is somewhat different. We will only discuss the /-ɪ/ allomorph in this paper, since the /-ɛ/ allomorph never cooccurs with other suffixes, making it impossible to justify a morphological layer on the basis of affix ordering. For further discussion of the two allomorphs, see McPherson (2013: §12.4). An example of the /-ɪ/ allomorph is [nɔ́ːlɔ́] ‘mix’ ~ [nɔ́ːl-ɪ] ‘mixed’.

There are a fair number of other inflectional suffixes, none of which alternate by harmony. They include /-ɛːlɛ́/ ‘negative imperfective’, /-dɛ́/ ‘affirmative imperfective’, /-aː/ ‘perfective nonfinal’, /-ɛː/ ‘imperfective nonfinal’, /-lɪ/ ‘negative perfective’, /-gʊ/ ‘negative imperative’, and /-mɔ́/ ‘hortative’.

### 3.4 The system of vowel phonemes in suffixes

We see above that a wide variety of contrasting vowel qualities may appear in Tommo So verbal suffixes: specifically, there are examples of /i/, /u/, /a/, /e/, /ɛ/, /ɔ/, and /ɔ́/ (we assume the absence of /o/ is an accidental gap). This would likely defeat, or at least complicate, any effort to derive suffix vowel qualities by some sort of default insertion processes, and in what follows we will assume fully specified underlying suffix vowels, noting alternatives where feasible.

### 3.5 Suffix ordering

As noted in the introduction, the purpose of this work is to present and analyze a Kiparskian correlation between affix ordering and applicability of phonology (vowel harmony). In this section, we set out the relevant principles of affix ordering in Tommo So. We extend Kiparsky’s simple ordering test (recall non-in-X vs. *in-non-X) to the rather more elaborate ordering pattern found in Tommo So. What emerges is a whole chain of pairwise orderings, diagnosing a system of morphological levels or layers.
To begin, we find that the Factitive often precedes the Mediopassive, as in [pálá-nd-íjé] ‘long-FACTITIVE-MEDIOPASSIVE’ = ‘become long’. The opposite order is never found; i.e. *[X-íjé-ndé] is ungrammatical in Tommo So. We can do the same test with other pairs and get the same result: only one order is possible. For brevity, we summarize the results of this test in (4):

(4) a. **Factitive before Mediopassive**

írë-nd-íjé ‘better-FACTITIVE-MEDIOPASSIVE’ = ‘get better’

b. **Reversive before Mediopassive**

mènn-il-íjé ‘fold-REVERSIVE-MEDIOPASSIVE’ = ‘become unfolded’

c. **Transitive before Mediopassive**

só-ír-íjé ‘sweat-TRANSITIVE-MEDIOPASSIVE’ = ‘sweat’

d. **Mediopassive before Causative**

jùb-íjé-mó ‘spill-MEDIOPASSIVE-CAUSATIVE’ = ‘spill’

e. **Causative before Perfective**

èbè-m-i ‘buy-CAUSATIVE-PERFECTIVE’ = ‘X made somebody buy’

Moreover, the combinations we might expect on grounds of transitivity (like Factitive before Causative) are generally attested, a point we will not document here. Naturally, in some words there are multiple suffixes present, and these reflect the same ordering generalizations as the two-suffix forms; e.g. àmá-nd-íjé-m-ì ‘rancid-FACTITIVE-MEDIOPASSIVE-CAUSATIVE-PERFECTIVE’.

To test our affix ordering principles, we used the entire database, extracting from it all words in which two or more verbal affixes attach to the same root. The evidence thus obtained is fairly clear. In particular, there are no cases whatsoever in which two affixes occur in the opposite order to what is given above in (4); ordering is entirely consistent. However, there is one set of affixes (Factitive, Reversive, Transitive) that never cooccur and thus cannot be assessed for linear order. The results of our ordering study are given in (5) in the form of a Hasse diagram.

---

4 A caveat: Transitive and Mediopassive seldom occur; but the attested cases are ordered as shown.
Affix ordering in Tommo So verbs

Root
   Factitive  Reversive  Transitive
      Mediopassive
          Causative
              Perfective

For purposes of the analysis to follow, we elaborate these empirical findings into a system of morphological layers, roughly along the lines envisioned in the theory of Lexical Phonology and Morphology (Kiparsky 1982 et seq.).\(^5\) We label our layers as in (6).

Morphological layers assumed for Tommo So

1. Root
2. Factitive  /-ndé/ (derivation)
3. Reversive  /-iile/ (derivation)
4. Transitive  /-iré/ (derivation)
5. Mediopassive  /-iże/ (derivation)
6. Causative  /-mɔ/ (derivation)
7. Defocalized Perfective  /-i/ (inflection)

These reflect the ordering observations summarized in (5), but go further in placing Factitive, Reversive, and Transitive in separate layers. The placement of Transitive in a layer “outside” Factitive and Reversive can be defended on the grounds that the Transitive is more productive and semantically transparent than the latter two suffixes; it is a characteristic of most level-ordered morphological systems that productive and transparent affixes gravitate to outer layers (see, e.g., Katamba 2004). The same reasoning might also serve to justify the placement of Reversive outside Factitive, although the difference in productivity is not as clear in this case.\(^6\)

4. The vowel harmony pattern

With the principles of affix ordering in place, we can turn now to the other side of the Kiparskian correlation: the applicability of phonological processes, here vowel harmony.

\(^5\)For a summary of arguments against defining levels based on affix ordering, see Kaisse and McMahon (2011). In this paper, we take the Kiparskian correlation between affix order and phonological process application as our starting point, but note that the “morphological distance” defined by affix ordering may represent not only lexical strata, in the sense of Lexical Phonology and Morphology, but also a more abstract notion of “morphological cohesion”, as in the work of Hay (2002 et seq.).

\(^6\)The absence of forms with Reversive outside Factitive may indeed be an accidental gap; the existing Factitive forms of Tommo So appear to be semantically incompatible with Reversive.
We give the facts of Tommo So vowel harmony first in rule-based phonology (Chomsky and Halle 1968 = SPE), employed for the sake of descriptive precision. In this framework, Tommo So would be considered to have three separate vowel harmony rules, one each for the features [low], [back], and [ATR]. Since harmony is frequently optional, we will give examples for each harmony process of both application of harmony and non-application.

4.1 Verb roots

As in other languages (see e.g. Kiparsky 1973:36 on Finnish, Clements and Sezer 1982:222-231 on Turkish), the vowel harmony pattern is slightly different as it applies within roots vs. in affixes, so we discuss the two separately. Harmony is also slightly different for verbs vs. nouns, and we will be focusing on verbs here. For roots, we give data based on a corpus of verbal roots (all of the affixes we discuss are verbal suffixes), which appears as Table 1.

**Table 1: Sequences of vowels appearing in Tommo So verb roots**

<table>
<thead>
<tr>
<th></th>
<th>i</th>
<th>e</th>
<th>ɛ</th>
<th>a</th>
<th>ɔ</th>
<th>o</th>
<th>u</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>9</td>
<td>39</td>
<td>56</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e</td>
<td>37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e</td>
<td>4</td>
<td>79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>2</td>
<td>4</td>
<td>151</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>o</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>u</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>46</td>
<td>8</td>
</tr>
</tbody>
</table>

Key to shading: Black = ATR disharmonic, dark gray = Backness disharmonic, light gray = Low disharmonic

The rows of Table 1 are labeled with the first vowel of a root, the columns with the second or third vowels (there are no roots longer than three syllables); thus the “9” in the upper leftmost cell indicates that there are nine cases in the corpus in which a root has [i] as its first vowel and another [i] as its second or third vowel.

The boxed regions of the table, showing non-initial high vowels, require comment. It is a firm phonological constraint of Tommo So that no verb root may end in a high vowel; thus all of the high vowel counts in the boxes represent medial high vowels in trisyllabic roots. As such, they are in the context for vowel reduction, and the observed distribution between [i] and [u] reflects the allophonic variation among reduced vowels noted in §3.2.

It is evident from Table 1 that many of the logical possible sequences of vowels are absent or severely underrepresented. These gaps are largely the result of vowel harmony and will be accounted for below as we go through the various harmony processes.
4.2 Low Harmony

The Low Harmony pattern is a bit complex; for precision we first state it as an SPE-style rule in (7).

(7) Low Harmony stated as a rule

\[
\begin{bmatrix}
\text{V}_{-\text{reduced}}
\end{bmatrix} \rightarrow [\alpha_{\text{low}}] / \# C_0 \begin{bmatrix}
\text{V}_{\alpha_{\text{low}}}
\end{bmatrix} X ___
\]

“A nonreduced vowel takes on the same value of [low] as the initial vowel.”

Here, X stands for any sequence, meaning that the rule can apply non-locally, affecting all the non-initial vowels of the word.

The surface pattern implied by (7) is as follows. [a] can never be followed by a mid vowel ([−high, −low]), nor can [a] ever come after an initial vowel other than [a]. Initial high vowels trigger Low Harmony, since in that position they are not reduced (§3.2). Thus, an initial high is never followed by [a]. But high vowels may follow initial [a], either because they are medial and therefore reduced, or because they occur in the Defocalized Perfective suffix and are therefore outside the domain of Low Harmony; see (12).

The effects of Low Harmony can be seen clearly in our verb root corpus. In Table 1, the cells that would be excluded by Low Harmony are those with light gray shading, forming a gapped cross seven cells high by five wide. Some representative data illustrating the pattern, as well as exceptions, are given in (8).

(8) Low Harmony in roots

a. Regular forms

[ámá] ‘be fattened’
[dâmbá] ‘push’
[ádîbá] ‘think’
[dènné] ‘look for’
[sùmmó] ‘dilute’

---

7 We gloss over a fairly major, though orthogonal issue concerning the relationship of harmony to prosodic structure. A precedent for our claim that initial, prosodically-prominent vowels can trigger harmony non-locally is given in Ringen and Heinämäki’s (1999:316) account of Finnish.

8 Reviewers ask if harmony is statistically more reliable in local contexts, i.e. when the trigger and target are in adjacent syllables. It appears that there is a modest effect of this kind, notably, disyllabic stems have only about a 1% harmony exception rate; whereas for the final vowels of trisyllabic stems the exception rate is about 7%. In principle we could complicate our analyses to include the local/nonlocal distinction, but trisyllabic stems are so rare (about 10% of the total) that it would make little difference to the outcome.
b. (Rare) exceptional forms

[saːdɛ] ‘die without being slaughtered’
[jàmĩndɛ] ‘rub soap between hands’

The words given in (8) represent two out of the total of only four exceptions in roots.

Consider next the behavior of Low Harmony in suffixes. Because of the “petering out” effect that is the main focus of this article, only two suffixes are affected by Low Harmony, namely the ones that form the innermost layers of the morphology ((6)).

The Factitive suffix surfaces as [-ndá] when it follows a low-vowel root, but as any of [-ndē], [-ndě], [-ndó], or [-ndɔ] in other contexts, depending on other vowel harmony processes. We treat the Factitive as underlying /-ndɛ/, the value that surfaces when no vowel harmony process applies.

(9) Low Harmony in Factitive forms

a. /dzάː-ndɛ/ → [dzάː-ndá] ‘meal-FACTITIVE’ = ‘cook’
b. /dągá-ndɛ/ → [dągá-ndá] ‘be.good-FACTITIVE’ = ‘fix’

The application of Low Harmony for Factitive forms is about 85%. By this we mean, of all cases in the data where a low-vowel root precedes the Factitive suffix, 85% surface with a lowered Factitive vowel. An example of non-application is [dzάː-ndɛ], which is the very same word as (9a), uttered by the same speaker on a different occasion.

We treat the Reversive suffix as underlying /-ílɛ/. For this suffix, Low Harmony is actually the exception, not the norm; the application rate is about 20%. When Low Harmony applies, the Reversive surfaces as [-ilá]; otherwise, it appears after [a] as its underlying form [-ilɛ], or else as [-ilɛ], [-ilɔ], or [-ilɔ] where other vowel harmony processes are applicable. It can be observed that application of Low Harmony to this suffix is non-local, skipping over the reduced suffix-initial vowel.

(10) Low Harmony in Reversive forms

a. Application

/pánda-ilɛ/ → [pánda-ilá] ‘widow-REVERSIVE’ = ‘marry a widow’
/mánda-ilɛ/ → [mánda-ilá] ‘seal-REVERSIVE’ = ‘unseal’

---

9 This word is a loanword, Fulfulde saxduđe ‘to die (said of animals)’ (Osborn et al. 1993).
b. Exceptional non-application

/pándá-ilé/ → [pánd-ilé] (in free variation with (10a))
/jàmbá-ilé/ → [jàmb-ilé] ‘cover-REVERSIVE’ = ‘uncover’

The forms in (10) illustrate another general phonological process of Tommo So, namely the resolution of hiatus. Normally, when suffixation creates a sequence of two vowels, the first of the two is deleted (McPherson 2013, §3.7.3). The process could thus be described with a very simple SPE-style rule, as in (11).

(11) Hiatus Resolution stated as a rule

\[ V \rightarrow \emptyset / \_\_ V \]

Hiatus Resolution can be seen in (10a), where the final /a/ deletes before suffix-initial /i/. The form in (10b) is an irregular case, with hiatus resolved by loss of the second vowel. For more discussion of Hiatus Resolution, see §10.3 below.

The remaining suffixes fall outside the domain where Low Harmony applies. To demonstrate this, we give examples of non-application for these suffixes.

(12) Non-application of Low Harmony to the remaining suffixes

a. /jàmbâ-ìrɛ/ → [jàmbirɛ] (*[jàmbirá]) ‘cover-TRANSITIVE’ = ‘cover’
b. /jàmbá-ìjɛ/ → [jàmbijɛ] (*[jàmbijá]) ‘cover-MEDIOPASSIVE’ = ‘cover oneself’
c. /káná-mɔ́/ → (no change; *[kánámá]) ‘do-CAUSATIVE’ = ‘cause to do’
d. /káná-ì/ → [kànì] (*[kànà]) ‘do-PERFECTIVE’ = ‘did’

4.2.1 Excursus: defining application rate

Since application rates of harmony processes are the key data in this article, we take a moment to define them, using examples from the data just presented. There are four relevant data categories.

(13) Categories defining application rates

c. Vacuous application. Process is applicable, but its conditions were already met in the input, so no change is made. Example: /jɛ-ndɛ/ → [jɛ-ndɛ] ‘see-FACTITIVE’ = ‘look at’.

\[10\] The Defocalized Perfective form is characterized by an all-L grammatical tone pattern (McPherson 2013:§12.4.1).
In the output, the [−low] of the suffix matches the [−low] in the root vowel, satisfying (7), but it did so already in the input form.

d. **Inapplicability.** Process is simply not applicable. This never arises for Low Harmony, since it is triggered by any vowel. A legitimate example arises below for Backness Harmony: by our statement (15) below it is not triggered by the low vowel [a], so any disyllabic form whose first vowel is [a] would be a case of inapplicability.

In what follows, we will define application rate by the formula in (14):

(14) **Defining application rate**

\[
\frac{\text{cases of nonvacuous application}}{\text{(cases of nonvacuous application + cases of nonapplication)}}
\]

The numerator counts forms in which application would produce an observable change, and the denominator counts all cases of potential applicability. This definition is the sensible one to use here, as it matches the criterion of adequacy for a constraint-based analysis (see §7 below). In cases of vacuous application or inapplicability, the faithful candidate incurs neither Markedness nor Faithfulness violations and will always win, making it uninformative about ranking or weighting.

4.3 **Backness Harmony**

In this section we provide data supporting a Backness Harmony pattern that could be described with the following rule:

(15) **Backness Harmony stated as a rule**

\[
\begin{array}{c}
\left[ \text{V} \leftarrow \text{reduced} \right] \rightarrow [\alpha_{\text{back}}] / \# C_0 \left[ \text{V} \right]_{\alpha_{\text{back}}} X \end{array}
\]

“A nonreduced vowel takes on the same value of [back] as the initial vowel.”

We first cover some details of the pattern. Consider first the behavior of [a]. Under the feature system we assume (see (3)), [a] cannot be a trigger of Backness Harmony, since it has no backness value that can be transmitted to the target vowel. Hence [a] is compatible with both front and back following vowels. Moreover, it is simply impossible to tell if [a] is an undergoer of Backness Harmony because there are no possible inputs; in roots, Low Harmony eliminates all non-initial [a] after backness-specified vowels, and there are no low-vowel suffixes in the layers of the morphology where harmony prevails. Second, Backness Harmony does not target reduced vowels (see §3.2 above) because, as noted above, their backness appears to be gradient and largely determined by coarticulation. For instance, in a form like [jàmìɲdʑɛ́] ‘rub soap between hands’, the [i] largely reflects phonetic assimilation to the following palatal [n].

Turning to the data, we first illustrate Backness Harmony with data from roots, as summarized above in Table 1. With very rare exceptions (about 1.7% of all forms), the front
vowels [i e e] may only cooccur with front vowels, and the back vowels [u o o] only with back vowels (see gray region of table). For the central vowel [a], the pattern is difficult to discern with root data alone, because Low Harmony forces vowels preceded by [a] to be [a], preempting Backness Harmony. In (16) we give examples of both normal harmonic roots, along with examples of the rare exceptions.

(16) **Backness Harmony in roots**

a. *Regular forms*

- [gijé] ‘harvest’
- [kéré] ‘bite’
- [dùgó] ‘casts spells’
- [bōgóló] ‘bellow’

b. *(Rare) exceptional forms*

- [g̃b̃d̃] ‘barely touch something’
- [k̃j̃] ‘be hoarse’

In the suffix data, we find a consistent pattern of “petering out” as we move morphologically away from the root, the theme of this article. We cover the suffixes in order below.

In the Factitive, Backness Harmony is virtually exceptionless; indeed our data corpus includes only one exceptional form, (17b). For completeness, our documentation here includes cases of vacuous as well as nonvacuous application.

(17) **Backness Harmony in Factitive forms**

a. *Application*

- /dɔ̃ː-nd̃/ → [dɔ̃ː-nd̃] ‘arrive-FACTITIVE’ = ‘move (sth) near’
- /gɔ̃ː-nd̃/ → [gɔ̃ː-nd̃] ‘go.out-FACTITIVE’ = ‘take (sth) out’
- /dùː-nd̃/ → [dùː-nd̃] ‘bottom-FACTITIVE’ = ‘put down’

b. *Non-application*

- /dzɔbɔ-nd̃/ → (same) ‘run-FACTITIVE’ = ‘make (horse) run’

---

11 The tonal alternation is due to Tonal Absorption (McPherson 2013:§4.3.1); Rise becomes Low before High.

12 This form has a noun base; we include it to show that [u] is a back harmony trigger; for discussion of ATR harmony with noun bases, see fn. 17. The High tone of the nominal must become Low to conform to verbal tonotactics (McPherson 2013:§4.1.3).
c. **Vacuous application**

- `/jɛ̀-ndé/ → [jɛ̀-ndé]` ‘see-FACTITIVE’ = ‘look at’
- `/dɛ̀-ndé/ → [dɛ̀-ndé]` ‘know-FACTITIVE’ = ‘introduce’
- `/dɛ́mɛ̀-ndé/ → [dɛ́mɛ̀-ndé]` ‘be.hurt-FACTITIVE’ = ‘hurt (sb)’

d. **Inapplicable**

- `/dɛ̀-ndé/ → [dɛ̀-ndá] or [dɛ̀-ndɛ̀]` ‘meal-FACTITIVE’ = ‘cook’

Form (17d) shows optional application of Low Harmony. It also forms part of the evidence that [a] is not a Backness Harmony trigger: in the free variant where Low Harmony does not apply, we get [ɛ̀] as the suffix vowel, which reflects the underlying form. In all other forms of (17), the suffix also harmonizes for the feature [ATR], to be discussed below.

The next morphological layer consists of the Reversive suffix. Here, Backness Harmony is only slightly less robust, applying to 90.9% of applicable forms. In (18) we give representative examples.

(18) **Backness Harmony in Reversive forms**

a. **Application**

- `/g̥ŋ̥-i̱l̥/ → [g̥ŋ̥-i̱l̥]` ‘fence.in-REVERSIVE’ = ‘unfence’
- `/t̥ŋ̥-i̱l̥/ → [t̥ŋ̥-i̱l̥]` ‘crumple-REVERSIVE’ = ‘uncrumple’
- `/m̥nung̥-i̱l̥/ → [m̥nung̥-i̱l̥]` ‘roll-REVERSIVE’ = ‘unroll (mat)’

b. **Non-application**

- `/m̥ndz̥-i̱l̥/ → [m̥ndz̥-i̱l̥]` ‘break-REVERSIVE’ = ‘break off’
- `/úm̥-i̱l̥/ → [úm̥-i̱l̥]` ‘breathe-REVERSIVE’ = ‘resuscitate’

c. **Vacuous application**

- `/d̥b̥-i̱l̥/ → [d̥b̥-i̱l̥]` ‘get.stuck-REVERSIVE’ = ‘get unstuck’
- `/n̥m̥b̥-i̱l̥/ → [n̥m̥b̥-i̱l̥]` ‘trim-REVERSIVE’ = ‘cut off branch’
- `/d̥ŋ̥-i̱l̥/ → [d̥ŋ̥-i̱l̥]` ‘tie-REVERSIVE’ = ‘untie’

d. **Inapplicable**

- `/t̥g̥-i̱l̥/ → [t̥g̥-i̱l̥]` ‘shoe-REVERSIVE’ = ‘take off one’s shoes’

The vowel transcribed as [i] in these Reversive forms is a reduced vowel (see §3.2), hence skipped over by harmony. Once again, we see ATR alternations as well as backness alternations.
At the next morphological layer, the Transitive, the application rate of Backness Harmony drops to 69%. Again, we see that [a] is not a trigger, that the reduced vowel [i] in the suffix is transparent, and that ATR alternations accompany backness alternations.

(19) Backness Harmony in Transitive forms

a. Application

\[
\begin{align*}
\text{/òg-iré/} & \rightarrow [òg-irê] \quad \text{‘hot-TRANSITIVE’ = ‘heat (sth)’}
\text{/dògó-iré/} & \rightarrow [dòg-irô] \quad \text{‘be.face.up-TRANSITIVE’ = ‘hold (sth) face up’}
\text{/túŋ-iré/} & \rightarrow [túŋ-irô] \quad \text{‘kneel-TRANSITIVE’ = ‘make (sb) kneel’}
\end{align*}
\]

b. Non-application

\[
\begin{align*}
\text{/òg-iré/} & \rightarrow [òg-iré] \quad \text{(same input form as above, same speaker)}
\text{/sònnúgó-iré/} & \rightarrow [sònnúg-iré] \quad \text{‘place on shoulders-TRANSITIVE’ = ‘put on somebody else’s shoulders’}
\end{align*}
\]

c. Vacuous application

\[
\begin{align*}
\text{/sè-iré/} & \rightarrow [sè-irê] \quad \text{‘adorn-TRANSITIVE’ = ‘adorn (sb)’}
\text{/tègè-iré/} & \rightarrow [tèg-irê] \quad \text{‘drip-TRANSITIVE’ = ‘make drip’}
\text{/dimbè/} & \rightarrow [dimb-irê] \quad \text{‘follow-TRANSITIVE’ = ‘make follow’}
\end{align*}
\]

d. Inapplicable

\[
\begin{align*}
\text{/tágá-iré/} & \rightarrow [tág-irê] \quad \text{‘shoe-TRANSITIVE’ = ‘put shoes on somebody’}
\end{align*}
\]

At the next morphological layer, the Mediopassive, the application rate of Backness Harmony drops to 44%.

(20) Backness Harmony in Mediopassive forms

a. Application

\[
\begin{align*}
\text{/tómá-ijé/} & \rightarrow [tóm-ijô] \quad \text{‘wind-MEDIOPASSIVE’ = ‘become wound up’}
\text{/tôŋnó-ijé/} & \rightarrow [tôŋn-ijô] \quad \text{‘curl.up-MEDIOPASSIVE’ = ‘be curled up’}
\text{/jùbó-ijé/} & \rightarrow [jùb-ijô] \quad \text{‘spill-MEDIOPASSIVE’ = ‘be spilled’}
\end{align*}
\]

13 /òg/ is an adjectival root. Either we can assume deadjectival derivation in this case, or we can posit the existence of a bound verbal root /ògô/, which obeys verb root phonotactics by ending in a harmonic vowel; the second vowel would be removed by Hiatus Resolution.

14 Indeed, the two free variants were uttered in the same session, probably about ten minutes apart.

15 Bound root /sè/; long vowels shorten rather than deleting prevocally.
b. Non-application

/tómá-ijé/ → [tóm-ijé] ‘wind.up-MEDIOPASSIVE’ = ‘be wound up’
/góró-ijé/ → [gór-ijé] ‘hat-MEDIOPASSIVE’ = ‘wear a hat’
/mùnnó-ijé/ → [mùnn-ijé] ‘roll-MEDIOPASSIVE’ = ‘be rolled up’

c. Vacuous application

/péndé-ijé/ → [pénd-ijé] ‘make.tight-MEDIOPASSIVE’ = ‘get crowded’
/dɔ̀bá-ijé/ → [dɔ̀bá-ijé] ‘hang-MEDIOPASSIVE’ = ‘be hanging’
/tímbé-ijé/ → [tímb-ijé] ‘stack-MEDIOPASSIVE’ = ‘become stacked’

d. Inapplicable

/káná-ijé/ → [kán-ijé] ‘do-MEDIOPASSIVE’ = ‘take place’

At the next morphological layer, the Causative, the application rate of Backness Harmony drops to just 18%.

(21) Backness Harmony in Causative forms

a. Application

/témé-mó/ → [témé-mé] ‘eat-CAUSATIVE’ = ‘make (sb) eat’
/biré-mó/ → [biré-mé] ‘work-CAUSATIVE’ = ‘make (sb) work’

b. Non-application

/kéré-mó/ → [kéré-mó] ‘bite-CAUSATIVE’ = ‘make (sb) bite’
/jèmé-mó/ → [jèmé-mó] ‘melt-CAUSATIVE’ = ‘make (sth) melt’
/sídé-mó/ → [sídé-mó] ‘pay-CAUSATIVE’ = ‘make (sb) pay’

c. Vacuous application

/dɔ̀bó-ómó/ → [dɔ̀bó-ómó] ‘run-CAUSATIVE’ = ‘make (sb) run’
/òbó-mó/ → [òbó-mó] ‘give-CAUSATIVE’ = ‘make (sb) give’
/nújó-mó/ → [nújó-mó] ‘sing-CAUSATIVE’ = ‘make (sb) sing’

d. Inapplicable

/káná-mó/ → [káná-mó] ‘do-CAUSATIVE’ = ‘make (sb) do’

The form [káná-mó] in (21d) should be compared with (20d), [kán-ijé]: the pair illustrates that [a] is not a Backness Harmony trigger and that after [a] roots the underlying backness value of the suffix surfaces; i.e. /-mó/ vs. /-ijé/. Since /-mó/ is underlyingly [+back], our examples of non-application involve front-vowel roots.
In the next morphological layer, the Defocalized Perfective, the Backness Harmony rate drops to 13.6%. Here are representative forms.

(22) Backness Harmony in Defocalized Perfective forms

a. Application
\[
/\text{bødó}-i/ \rightarrow [\text{bød-ù}] \quad \text{‘put-PERFECTIVE’ = ‘put’}
\]
\[
/\text{óbó}-i/ \rightarrow [\text{òb-ù}] \quad \text{‘give-PERFECTIVE’ = ‘gave’}
\]
\[
/\text{dzúngó}-i/ \rightarrow [\text{dzúng-ù}] \quad \text{‘nod-PERFECTIVE’ = ‘nodded’}
\]

b. Non-application
\[
/\text{bògɔ̀l}-i/ \rightarrow [\text{bògɔ̀l}] \quad \text{‘chatter-PERFECTIVE’ = ‘chattered’}
\]
\[
/\text{bødó}-i/ \rightarrow [\text{bød-ì}] \quad \text{‘put-PERFECTIVE’ = ‘put’}
\]
\[
/\text{dzúngó}-i/ \rightarrow [\text{dzúng-ì}] \quad \text{‘nod-PERFECTIVE’ = ‘nodded’}
\]

c. Vacuous application
\[
/\text{ségiré}-i/ \rightarrow [\text{sègir-ì}] \quad \text{‘meet-PERFECTIVE’ = ‘met’}
\]
\[
/\text{dzidzibé}-i/ \rightarrow [\text{dzidzib-ì}] \quad \text{‘shake-PERFECTIVE’ = ‘shook’}
\]

d. Inapplicable
\[
/\text{káná}-i/ \rightarrow [\text{kàn-ì}] \quad \text{‘do-PERFECTIVE’ = ‘did’}
\]

All remaining morphology has a Backness Harmony rate of zero; thus the suffixes (all inflectional; §3.3.2) have but a single surface realization.

4.4 ATR Harmony

The feature [ATR] is phonemic in Tommo So only for the class of mid vowels [e ɛ o ɔ]. It is these vowels that form both the trigger class and the target class for ATR Harmony, stated as a rule in (23).

(23) ATR Harmony stated as a rule

\[
\left[ \begin{array}{c} V \\ -\text{high} \\ -\text{low} \end{array} \right] \rightarrow [\alpha\text{ATR}] / \left[ \begin{array}{c} V \\ \text{low} \\ \text{low} \end{array} \right] \quad X \quad \text{“A mid vowel takes on the ATR value of the closest preceding mid vowel.”}
\]

In verb roots, ATR Harmony is entirely exceptionless: see Table 1, where the eight black boxes, each with frequency zero, represent the eight logically possible sequences of disagreeing mid vowels ([e e], [e ɔ], [e e], [e o], [o e], [o ɔ], [ɔ e], [ɔ o]). We give a few representative harmonic roots in (24).
(24) **ATR Harmony in roots (exceptionless)**

a. [ébé] ‘buy’
b. [gègédé] ‘(insects) bite off and eat’
c. [kɔrɔndɔ] ‘snore’
d. [kɔmmɔ] ‘crumple’
e. [gɔbɔdɛ] ‘barely touch something’

The form (24e) is a rare exception to Backness Harmony in roots; it nevertheless obeys ATR Harmony.

The high vowels [i] and [u] are not in general ATR triggers:\textsuperscript{16} in roots (Table 1) they may cooccur with either the [+ATR] vowels [e o] or the [−ATR] vowels [ɛ ɔ], in either order. This is shown in (25).

(25) **Free combination of [i] and [u] with both values of [ATR]**

a. [kɪdɛ] ‘discuss’
b. [pɪjɛ] ‘cry’
c. [udɔ] ‘build’
d. [tʊŋdzɔ] ‘slap wet laundry against a stone’

In principle, we would could test our claim of non-triggerhood with suffix data, but due to the phonotactic restriction on high vowels in verb roots (§4.1), no actual cases arise: all roots contain at least one non-high vowel; and a high vowel will never be the closest vowel to the suffix. For this reason, the rule is formulated with the closest mid vowel as trigger rather than the initial vowel (as in Low and Backness Harmony).

The low vowel [a] likewise does not trigger ATR Harmony — with the mid defocalized perfective allomorph (which is 100% ATR harmonic), it takes [−ɛ] (e.g. /bála-ɛ/ → [bá-ɛ] ‘sweep-PERFECTIVE’ = ‘swept’); see McPherson (2013; §12.4) for more data.

We now go through the layers of the morphology, as before, exemplifying applicability of ATR Harmony. However, the data are far simpler, because the harmony rate is always either 100% or zero.

For the Factitive, Reversive, Transitive, and Mediopassive, the harmony rate is 100%. The actual allomorph that surfaces depends as well on the other forms of harmony.

\textsuperscript{16} There is a small class of Factitive forms where they do seem to act like ATR triggers; see fn. 17.
There is a minor conundrum involving the Factitive suffix that we have not resolved: the very small amount of evidence we have (just four roots) indicates that high vowel roots trigger [+ATR] in the Factitive; for instance [dùː-ndó] ‘bottom-FACTITIVE’ = ‘put down’. All four roots are nominal or adjectival; as already noted, verb roots may not end in high vowels, which is why these examples are rare. The appearance of [+ATR] vowels in post-high Factitives is a puzzle, since high vowels are fully compatible with [−ATR] vowels in roots (see (25)). To solve this, we might add an additional minor phonological process specific to the Factitive, or perhaps adopt a special underlying representation for it (e.g. underspecified for [ATR], or multiple listed allomorphs). As this behavior does not appear to bear on our main point, we will not attempt to resolve it here.
b. *Vacuous application*

/tımbé-ıřé/ → [tımb-ıře] ‘fern TRANSITIVE’ = ‘make (sb) find’

/tıg-ıř/ → [tıg-ıř] ‘hot TRANSITIVE’ = ‘heat (sth)’

c. *Inapplicable*

/tıgá-ıřé/ → [tıg-ıřé] ‘shoe TRANSITIVE’ = ‘put shoes on somebody’

(29) *ATR Harmony in Mediopassive forms*

a. *Application*

/dıxelé-ıjé/ → [dıxel-ıjé] ‘hang MEDIOPASSIVE’ = ‘be hanging’

/tóŋŋó-ıjé/ → [tóŋŋ-ıjó] ‘crumple MEDIOPASSIVE’ = ‘become crumpled’

b. *Vacuous application*

/péngd-ıjé/ → [péngd-ıjé] ‘spread out MEDIOPASSIVE’ = ‘spread out (intr)’

/tómó-ıjé/ → [tóm-ıjó] ‘wind MEDIOPASSIVE’ = ‘become wound’

c. *Inapplicable*

/tıgá-ıjé/ → [tıg-ıjé] ‘shoe MEDIOPASSIVE’ = ‘put shoes on oneself’

For the Causative suffix, the harmony rate is zero. Hence it always surfaces as [−ATR]; usually [-mo] but sporadically [-me] due to Backness Harmony.

(30) *ATR Harmony in Mediopassive forms*

a. *Non-application*

/dıxŋŋe-mó/ → [dıxŋŋé-mó] ‘pick up CAUSATIVE’ = ‘make pick up’

/óbó-mó/ → [óbó-mó] ‘give CAUSATIVE’ = ‘make give’

b. *Vacuous application*


/nó-mó/ → [nó-mó] ‘drink CAUSATIVE’ = ‘make drink’

c. *Inapplicable*

/wálá-mó/ → [wálá-mó] ‘farm CAUSATIVE’ = ‘make farm’

4.5 *Summary of suffix harmony*

In Table 2 we summarize what has been said so far about suffix harmony. The second column gives our proposed underlying representation for each suffix, and the remaining columns
list the surface allomorphs of each suffix as they appear after the five types of verb root (recall that due to verbal phonotactics there are no verb roots with all high vowels). All surface forms assumed to have undergone some harmony process nonvacuously are labeled for the processes they undergo; LH = Low Harmony, BH = Backness Harmony, AH = ATR Harmony. In each cell with more than one allomorph, the more frequent one is listed first.

**Table 2: Suffix allomorphs**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Factitive</td>
<td>/-ndé/</td>
<td>-ndéAH</td>
<td>-ndé</td>
<td>-ndéLH</td>
<td>-ndéBH</td>
<td>-ndéAH</td>
</tr>
<tr>
<td>Reversive</td>
<td>/-ilé/</td>
<td>-iléAH</td>
<td>-ilé</td>
<td>-ilé</td>
<td>-iléBH</td>
<td>-iléAH</td>
</tr>
<tr>
<td>Transitive</td>
<td>/-íré/</td>
<td>-íréAH</td>
<td>-íré</td>
<td>-íré</td>
<td>-íréBH</td>
<td>-íréAH</td>
</tr>
<tr>
<td>Mediopassive</td>
<td>/-íjé/</td>
<td>-íjéAH</td>
<td>-íjé</td>
<td>-íjé</td>
<td>-íjéBH</td>
<td>-íjéAH</td>
</tr>
<tr>
<td>Causative</td>
<td>/-mó/</td>
<td>-móBH</td>
<td>-mó</td>
<td>-mó</td>
<td>-móBH</td>
<td>-móBH,AH</td>
</tr>
<tr>
<td>Defocalized</td>
<td>/-í/</td>
<td>-íBH</td>
<td>-í</td>
<td>-í</td>
<td>-íBH</td>
<td>-íBH</td>
</tr>
</tbody>
</table>

It can be seen that the Factitive and Reversive suffixes show a five-way alternation ([e ~ ɛ ~ a ~ ɔ ~ o]), since all three harmony processes apply to them. The Transitive and Mediopassive suffixes have only four allomorphs ([e ~ ɛ ~ ɔ ~ o]), since Low Harmony peters out before reaching them. The Causative has but two allomorphs ([ɛ ~ ɔ]), determined by Backness Harmony, since ATR Harmony peters out before reaching it. Lastly, the Defocalized Perfective likewise has only two allomorphs ([i ~ u]), determined by the only applicable harmony process, Backness Harmony.

5. **Frequency of application: the data**

In (31) we gather together the application rates reported above for all three harmony processes in all morphological contexts. The frequency values represent token frequency, calculated by counting each instance of a repeated underlying form as a separate case (for further discussion, see §10.2). Moreover, where a form has two harmonizing suffixes (none has three), it is counted twice, once for each suffix; this makes sense because what we are seeking to model

18 Both -ndé and -ilé occur zero times after [o] roots in the data corpus. We treat these cases as accidental gaps, resulting from (a) [o] being somewhat less common than other vowels in Tommo So; and (b) Backness harmony being almost obligatory for the Factitive and Reversive suffixes. The probability that a total of zero tokens for these allomorphs could arise by chance can be calculated at .51 for -ndé and .2 for -ilé; thus there is very little support for the view that these absences are meaningful.
the rate at which each suffix undergoes harmony. Application rates are shown first as a table and then in graph form (error bars represent 95% Clopper–Pearson binomial confidence intervals.)

(31) Application rates by suffix and harmony process

a. Table

<table>
<thead>
<tr>
<th>Suffix</th>
<th>Low</th>
<th>%</th>
<th>Backness</th>
<th>%</th>
<th>ATR</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Root</td>
<td>151</td>
<td>97.4</td>
<td>470</td>
<td>98.3</td>
<td>264</td>
<td>100</td>
</tr>
<tr>
<td>Factitive</td>
<td>57</td>
<td>85.1</td>
<td>95</td>
<td>99</td>
<td>80</td>
<td>100</td>
</tr>
<tr>
<td>Reversive</td>
<td>12</td>
<td>19.7</td>
<td>40</td>
<td>90.9</td>
<td>43</td>
<td>100</td>
</tr>
<tr>
<td>Transitive</td>
<td>0</td>
<td>19</td>
<td>40</td>
<td>69</td>
<td>31</td>
<td>100</td>
</tr>
<tr>
<td>Mediopassive</td>
<td>0/167</td>
<td>0</td>
<td>107/243</td>
<td>44</td>
<td>231</td>
<td>100</td>
</tr>
<tr>
<td>Causative</td>
<td>0/42</td>
<td>19</td>
<td>13/72</td>
<td>18.1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Perfective</td>
<td>0/119</td>
<td>0</td>
<td>17/125</td>
<td>13.6</td>
<td>N/A</td>
<td>—</td>
</tr>
</tbody>
</table>

b. Graph

As can be seen, for two of the harmony processes, application peters out gradually going outward from the root. ATR Harmony does not peter out gradually but plummets from 100 to 0% at the Mediopassive-Causative break.

---

19 Root rates calculated as follows: denominator = number of forms that match the structural description of the rules in §4, numerator is the number of harmony-compliant forms; we make no claims about the UR’s of harmonized roots. Our method of calculating root rates is somewhat arbitrary but we assume that other choices would yield very similar results; by any criterion all three harmony processes are very close to being obligatory in roots.
One further point will be important below: the two gradient processes peter out in a particular way, forming sigmoid (S-shaped) curves. When we turn to our analysis below (§7), our primary goal will be to derive this sigmoid shape in terms of principles already well-motivated in linguistic theory.

6. An earlier proposal: the inverted-exponential model

Before doing this, however, we first review what might fairly be taken to be the existing “standard model” for relating morphological distance to phonological process application, namely the widely-cited proposal put forth by Guy (1991). Guy developed his model in the course of studying the system of cluster simplification found in English final C + {t,d} clusters, which likewise applies with lower frequency in the outer layers of the morphology.

In Guy’s model, the diminishing frequency of application emerges directly from the architecture of the grammar. He adopts the general approach to morphological layering assumed in the theory of Lexical Phonology and Morphology (Kiparsky 1982 et seq.), in which the layers are the result of a sequence of derivationally ordered levels. For English, the levels assumed are as follows. The innermost level, to be called Level 0, is the root; thus for instance when the /t/ of a simplex form like act is dropped in fluent speech, this is a case of a Level 0 /t/. The intermediate Level 1 is the domain of the /t/ that occurs as a non-productive past tense suffix, triggering stem allomorphy in words like kept [kɛpt]. The outermost Level 2 is the domain of the regular past tense allomorph seen in words like tripped [tɹɪpt].

Studies examining the dropping of alveolar stops in consonant clusters have consistently found a petering out effect (Labov et al. 1968, Labov 1969, Wolfram 1969, Fasold 1972, Guy 1980, Neu 1980, Nesbitt 1984). For example, in the study of Chicano English by Santa Ana (1991), the root-level alveolar stops are deleted 74.3% of the time; the intermediate-level stops are deleted 59.3% of the time, and the stops of the outermost level are deleted 42.1% of the time.

Guy’s model relates this quantitative pattern to the grammatical architecture in a simple and elegant way, in which the frequencies emerge as a direct consequence of the system of level organization. Guy proposes that the -t,d Deletion rule has a constant application rate, which can vary from dialect to dialect. Moreover, -t,d Deletion applies at all three morphological levels, and failure to apply on one level does not preclude application on a later level. Thus, for roots like act there are three chances for -t,d Deletion to apply, for kept there are two, and for tripped there is just one, as shown in (32).
(32) **Morphological structure and rule application in the model of Guy (1991)**

\[ \begin{align*}
\text{a. Level 0} & \quad \text{Monomorphemic} & \quad [\text{ækt}] \\
\text{3 chances} & \\
\text{b. Level 1} & \quad \text{“Tightly-bound” suffix} & \quad [\text{kɛp} \text{t}] \\
\text{2 chances} & \\
\text{c. Level 2} & \quad \text{“Loosely-bound” suffix} & \quad [\text{trɪp} \text{t}] \\
\text{1 chance} & 
\end{align*} \]

With these assumptions in place, the model makes clear quantitative predictions. Assuming an application rate \( r \), then after \( n \) chances to apply, the “survival rate” is \((1 − r)^n\) and the application rate is therefore \( 1 − (1 − r)^n \). If we plot application rate against \( n \) for various values of \( r \), we get a family of inverted exponential curves, as in (33).

(33) **The inverted-exponential model: schematic predictions of application rate**

More specifically, when a phonological process applies at more than one level, we can predict the application rate for all later levels using the inverted exponential formula (i.e. in the English case, given just one parameter, we predict three observed values). For the data from Santa Ana mentioned above, this is a fairly good fit, as shown in (34).\(^{20}\)

---

\(^{20}\) For a recent critique of the model as applied to English -\( t,d \) Deletion, see Fruehwald (2012: 79-80), who suggests that the best-fit values of \( n \) in the model formula actually do not come out cleanly as integers when we include a random effect for word identity. For a case where the effect of morphological level failed to reach significance for deletion data, see Tagliamonte and Temple (2005).
(34) The inverted-exponential model applied to cluster simplification

Following up on Guy, Kiparsky (1994) proposes a related model in the framework of Optimality Theory. The effect of Guy’s sequential feeding of forms through the three levels is achieved instead by a special nested pattern of constraint violations and probabilistic rankings; see §9.2 for full summary. The empirical predictions of Kiparsky’s model are identical to those of Guy’s.

The data from Tommo So vowel harmony form a challenge to the inverted exponential model, since there are seven instead of three levels. Using the default settings of the Solver package in Excel, we found the inverted exponential curves that minimize mean absolute error, with three base rates (one for each harmony process).21 We set the application rate so as to minimize model error rather than by selecting any particular level (such as the stem) as criterial; our procedure gives the benefit of the doubt to the model. The resulting fit was poor, as (35) indicates.

---

21 The details of curve-fitting apparently matter rather little; we rechecked our work with other software and with hand search, and also using a mean squared error rather than absolute error criterion; model fit is similarly good in all cases. The spreadsheets for model fitting are available in the online Supplemental Materials.
Visually, the model matches poorly to the data; the mean absolute error is 0.181.²² Qualitatively, the inverted exponential curves are a poor match to the empirical curves, which, as graph (31b) above indicates, are not inverted U’s but tend rather to be S-shaped. The model does a particularly poor job with the categorical case of ATR Harmony, which we will show in the next section to be derivable under the same model used for the gradient cases.

7. Deriving the data pattern with Harmonic Grammar

We attempt to find a better fitting analysis, using existing ideas from current phonological theory. Ideally, under this analysis, the sigmoid shapes observed will not be stipulated, but will follow naturally from the principles of the theory.

We adopt two tools from current theorizing. First, we employ the constraint-based framework of Harmonic Grammar (Legendre, Miyata, and Smolensky 1990, Smolensky and Legendre 2006, Pater 2009, Jesney 2010, Potts et al. 2010, Jesney and Tessier 2011). This resembles standard Optimality Theory (Prince and Smolensky 1993/2004) in most respects but uses weighted instead of ranked constraints. In brief, every constraint bears a weight (a real

²²Error by harmony type: Low 0.249, Back 0.134, ATR 0.157. One might argue that for the ATR data, Guy’s model would not be an appropriate choice in the first place because the data are not gradient; even so, the error for the remaining cases is still high.
number, representing its strength); every candidate is assigned a value (its *harmony*), a sort of penalty score which is computed by multiplying weights by violation counts and summing. In the simplest version of the theory, the winning candidate is the one with the best harmony value.

Since the data involve variation, we need to employ a probabilistic implementation of Harmonic Grammar. There are two such implementations, which turn out to work about equally well. We give first an analysis with maxent grammar (Smolensky 1986, Goldwater and Johnson 2003; Wilson 2006; Hayes and Wilson 2008), examining later on Noisy Harmonic Grammar (Boersma and Pater, in press).

The second theoretical tool we will use is the idea that harmony for some constraints is calculated by multiplying the constraint weight by a value along a *scale*. For us, the scale will be an abstract morphological notion, “closeness to the root,” based on the evidence given in §3.5. The values we adopt for our scale are as in (36).

(36) **The “root-closeness” scale**

<table>
<thead>
<tr>
<th>Root</th>
<th>Factitive</th>
<th>Reversive</th>
<th>Transitive</th>
<th>Mediopassive</th>
<th>Causative</th>
<th>Perfective</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

The use of scales in constraint-based grammar has a long history. The original work in Optimality Theory, Prince and Smolensky (1993:§5.2), suggested a constraint HNUC that is gradiently violable, based on the sonority of the segment occupying nuclear position in a syllable.²³ Frisch, Broe, and Pierrehumbert (1997, 2004) took the next step, adopting constraint weights of which one was used to multiply the value on the relevant scale (a similarity metric for pairs of consonants in Arabic verbal roots); indeed, the math they employed is a special case of the maxent framework we will be employing. For Flemming (2001), the crucial scale was phonetic, namely values of second formant frequencies and their (squared) deviations from a specified target. Using a framework essentially identical to Harmonic Grammar with scalar constraints, Flemming found that he could derive the well-known “locus effect” in phonetics (Sussman et al. 1993) from first principles. We pursue the ideas in these works, using the morphological-distance scale.

Our morphological distance scale reduces morphological structure to a single number, as given in (36). In doing this, we do not intend to reject the insights of stratal theories of phonology (Kiparsky 1982, 2000, 2008; Bermudez-Otero 1999); rather, we assume that root-closeness is an abstract property that not only serves to define the system of strata, but is also directly accessible to constraints; that is, the constraints “know” what layer a suffix belongs to and can access this information as a scalar value.

To obtain the scale of (36), we arrange the affixes in descending order of assumed root-closeness. Each suffix is separated from its neighboring(s) by the arbitrarily-chosen value of 1. As we will later show (§7.6), selecting a separation value different from 1 yields exactly the same

---

²³ Prince and Smolensky avoid the use of actual integer values in their constraint, but the analytic effect is quite similar.
predictions, once the weights have been fitted to the data. It also makes no difference if we let
the scale run between different termini, such 6 to 0 or 0 to −6. Spacing the suffixes along the
scale unevenly would change the quantitative predictions of the analysis to a degree depending
on the differences in spacing.24

We seek a grammar that, for each combination of morphological layer and harmony process,
yields a number expressing the probability of vowel harmony as applied to the relevant suffix.
To do this, we adopt a set of ordinary phonological constraints (Markedness and Faithfulness, as
in OT), letting the Markedness constraints refer to the root-closeness scale of (36). The
constraints will be weighted in a way that generates a close match to the empirical values listed
in (31a).

7.1 Maxent calculations

In this section we give a brief review of the calculations of a maxent grammar. For fuller
discussion see e.g. Goldwater and Johnson (2003) or Hayes and Wilson (2008).

We first consider the way in which the grammar assigns probabilities to candidates. For each
candidate for a given input, the first step is to calculate its harmony. This is done by multiplying
each constraint weight by the number of violations of that constraint, then adding up the total
across constraints, as in (37):

\[
H(x) = \sum_{i=1}^{N} w_i C_i(x)
\]

where

- \(x\) is some candidate.
- \(H(x)\) is the harmony value being computed for that candidate
- \(w_i\) is the weight of the \(i\)th constraint
- \(C_i(x)\) is the number of times that \(x\) violates the \(i\)th constraint
- \(\sum_{i=1}^{N}\) denotes summation over all constraints (\(C_1, C_2, \ldots C_N\)).25

For the present application, the calculation can be made particularly simple if we adopt the
idealizing assumption that there are only two plausible candidates to consider for each input, one
for each value of the relevant feature. Thus, it is implicit in our account that other, non-stated
constraints rule out any candidates not mentioned. For example, *[−ndi] for Factitive /-ndɛ/ can
be assigned an arbitrarily low probability by giving IDENT(high) an indefinitely high weight.

24 For nonlinear scales in harmonic grammar see Pater (to appear).
25 Hayes and Wilson (2008) adopt slightly different terminology, calling (37) the formula for “scores”.
Once the harmony of the two relevant candidates (call them Cand1 and Cand2) has been calculated, the probability of Cand1 can be determined as in (38):

(38) Probability of a candidate in maxent (two-candidate system)

\[
p(Cand1) = \frac{\exp(-H(Cand1))}{\exp(-H(Cand1)) + \exp(-H(Cand2))}
\]

where

\[p(x) = \text{the probability of candidate } x\]
\[\exp(y) = e^y, \text{ where } e \text{ is the base of natural logarithms, about } 2.718\]
\[H(x) = \text{the harmony of } x, \text{ as given in (37)}\]

The probability of Cand2 is, of course, 1 minus the probability of Cand1. The overall pattern is large harmony values (deriving from violations of strong constraints) produce low predicted probability values. The probability predicted for a candidate will depend on the aggregate strength of its constraint violations, as well as those of the candidate with which it competes.

### 7.2 Constraint set

In our analysis for Tommo So, harmony is favored by Markedness constraints of the Agree family (Lombardi 1999, Baković 2000), these are opposed by corresponding Faithfulness constraints of the Ident family (McCarthy and Prince 1995).

In (39) are given the three necessary Agree constraints; these are approximations, whose content is to be modified.

(39) Agree( ) constraints used in the analysis (preliminary)

a. Agree (low): Assign a violation for every non-high vowel that disagrees in [low] with the initial vowel.

b. Agree (back): Assign a violation for every vowel that disagrees in [back] with the initial vowel.

c. Agree(ATTR): Assign a violation for every mid vowel that disagrees in [ATTR] with a preceding mid vowel.

The corresponding Ident constraints are given in (40).

---

26 The current theoretical climate for the analysis of vowel harmony is unsettled; in particular a wide variety of constraints for enforcing harmony have been put forth. In addition to Agree, the literature includes two types of Align (of featural autosegments, Kirchner 1993, Ringen and Vago 1998; and of feature spans, McCarthy 2004), as well as Spread (Kaun 1995, Ni Chiosáin and Padgett 1997, Walker 1998) and systems based on agreement by correspondence (Rose and Walker 2004, Hansson 2006, Rhodes 2012). Our concerns are, we think, largely orthogonal to the current debates; what we need is some sort of constraint that can enforce featural agreement and we adopt Agree as one option from among many.
(40) **Ident( ) constraints used in the analysis**

a. **Ident (low):** Assign a violation for every segment that disagrees in [low] with its underlying value.

b. **Ident (back):** Assign a violation for every segment that disagrees in [back] with its underlying value.

c. **Ident (ATR):** Assign a violation for every segment that disagrees in [ATR] with its underlying value.

7.3 **Scalar constraints**

The next step is to modify our Agree constraints in the manner promised above, i.e. so that they are sensitive to the morphologically-defined scale of (36). The idea is that when the target disagrees with the trigger vowel in the relevant way, the degree of the violation is determined by the morphological closeness of the target. In particular, following the references cited above we suggest that the number of violations for any particular Agree constraint is simply the value of the scale (36) for the target vowel. In essence, this turns the calculations of maxent from a multiplication of two values (weights times violations) to a multiplication of three (weights times violations times scalar value), as shown in (41):

(41) **Harmony calculation with scalar constraints**

\[
H(x) = \sum_{i=1}^{N} w_i C_i(x) S_i(x)
\]

where

\(S_i(x)\) is the value of candidate \(x\) along the scale invoked by constraint \(C_i\).

The constraints that invoke a scale in our analysis are the Agree( ) constraints of (39); we restate these constraints with a scale in (42) below.

(42) **Agree( ) constraints used in the analysis**

a. **Agree (low):** Assign a penalty for every non-high vowel that disagrees in [low] with the initial vowel. *Scale the violations according to* (36).

b. **Agree (back):** Assign a penalty for every vowel that disagrees in [back] with the initial vowel. *Scale the violations according to* (36).

c. **Agree (ATR):** Assign a penalty for every mid vowel that disagrees in [ATR] with a preceding mid vowel. *Scale the violations according to* (36).
Only a subset of constraints invoke a scale; for the others — in our analysis, the \textsc{Ident()} constraints — we may assume that if a scale if used for the sake of consistency, its value is always 1.

7.4 Establishing the weights

Most stochastic grammar theories are augmented with a learning model. It is assumed that the human language learner brings the constraints to the task of acquisition and that the rankings or weights are determined algorithmically during the learner’s encounter with language data (see e.g. Tesar and Smolensky 2000, Boersma and Hayes 2001, and Boersma and Pater, in press). We adopt this approach here, using our data corpus as an approximation to real-life learning data.

For maxent, it matters little in practical terms what weighting algorithm is adopted for purposes of mimicking human learning; the search space is free of local maxima (Della Pietra et al. 1997) and a great number of procedures yield essentially the same result. Here, we chose to set the weights so as to minimize mean absolute error, since this is how we report model fit. For convenience and easy replicability we used the Solver application in Microsoft Excel in its default settings. As a check we recomputed the results with the Maxent Grammar Tool; (Wilson and George 2009, www.linguistics.ucla.edu/people/hayes/MaxentGrammarTool) and obtained very similar results. All weight-setting calculations are posted in the Supplemental Materials.

We fitted six numbers, i.e. the constraint weights for the three \textsc{Ident()} constraints and the three \textsc{Agree()} constraints. The procedure is noncircular; we are using 6 values to predict 20 observations, i.e. the frequency of application for all possible combinations of harmony processes in 7 morphological contexts. The weights obtained were as follows.

(43) **Grammar: constraint weights found**

\[
\begin{array}{ll}
\text{\textsc{Ident} (low)} & 15.2 \\
\text{\textsc{Ident} (back)} & 4.0 \\
\text{\textsc{Ident} (ATR)} & 85.6 \\
\text{\textsc{Agree} (low)} & 2.8 \\
\text{\textsc{Agree} (back)} & 1.2 \\
\text{\textsc{Agree} (ATR)} & 34.8 \\
\end{array}
\]

7.5 Results

From the grammar thus constructed we computed the predicted percentage of vowel harmony, using the root-closeness values of (36) and the maxent formulae in (38) and (41). These percentages, given in (44a), closely fit the original data; in fact, the mean absolute error of the maxent model is 0.012 (broken down by harmony type, this is: Low 0.012, Back 0.022, ATR 0.000. This compares very favorably with the 0.181 (Low 0.249, Back 0.134, ATR 0.157) obtained with the inverted exponential model.
(44) Results: grammar predictions vs. observed

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th></th>
<th></th>
<th>Back</th>
<th></th>
<th></th>
<th>ATR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Root</td>
<td>97.4</td>
<td>98.4</td>
<td>98.3</td>
<td>99.1</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>2. Factitive</td>
<td>85.1</td>
<td>79.5</td>
<td>99</td>
<td>97.0</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>3. Reversive</td>
<td>19.7</td>
<td>19.7</td>
<td>90.9</td>
<td>90.4</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>4. Transitive</td>
<td>0</td>
<td>1.5</td>
<td>69.0</td>
<td>73.1</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>5. Mediopassive</td>
<td>0</td>
<td>0.1</td>
<td>44</td>
<td>44.0</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>6. Causative</td>
<td>0</td>
<td>0</td>
<td>18.1</td>
<td>18.5</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>7. Perfective</td>
<td>0</td>
<td>0</td>
<td>13.6</td>
<td>6.2</td>
<td>N/A</td>
<td>N/A27</td>
<td></td>
</tr>
</tbody>
</table>

While the superior fit of the model is encouraging, it is also important to understand the model from the viewpoint of restrictiveness — what sorts of frequency patterns could it model in principle? For this purpose it helps to consider in general terms how the model makes its predictions, a topic to which we now turn.

7.6 Sigmoid curves and their derivation

Harmonic grammars — in the general case — can be quite complex. Notably, Prince (1997) showed that under suitable idealizing assumptions, a harmonic grammar can mimic the behavior of any grammar expressed in Optimality Theory,28 hence all the intricate data patterns that can be derived with ranking in OT can also be derived in maxent. However, in the present case we have made an additional assumption that greatly simplifies analysis, namely that all candidates other than simple vowel harmony or the faithful realization are ruled out by other constraints. Hence only two candidates have any chance of winning, and it is solely the weights of AGREE() and IDENT() that determine the outcome.

Simple algebra tells us29 that under these circumstances, the only frequency patterns that can be derived by the grammar will take the form of sigmoid curves, of the particular type known as the logistic function, widely used in statistical analysis. The form of the logistic function relating application frequency to morphological distance is given in (45) (it is obtained by substituting (37) and (41) into (38)). In the formula, we generalize from the specific Markedness and Faithfulness constraints we used here to the case of any phonological process governed by a conflicting scalar Markedness constraints and nonscalar Faithfulness constraint.

27 We write Not Applicable here because as a high voweled suffix the Perfective is ineligible for ATR Harmony. Were a mid-voweled suffix to occur in this slot the application rate predicted by the grammar would be zero.

28 The crucial assumption is that there is some maximum number of possible violations for every constraint; for maxent we further need to assume that constraint weights can be unboundedly high.

29 We suppress the math here (it is textbook material, not original research) but invite readers to peruse the online Supplemental Materials for a full account.
(45) **Sigmoid generation under our analysis: the general case**

\[
P(\text{candidate undergoing phonology}) = \frac{1}{1 + e^{W_{\text{faith}} - (W_{\text{markedness}} \cdot s)}}
\]

where \(s\) is the value of the candidate along the relevant scale.

Making use of this formula, and plugging in our particular weights for Markedness and Faithfulness ((43)), we can restate the table of (44) as a graph, this time plotting the predictions of the analysis as the sigmoid curves generated by the formula.

(46) **Results:** grammar predictions plotted as sigmoid curves

![Graph showing sigmoid curves for different grammatical categories.]  

Further math, given in the Supplemental Materials, demonstrates the restrictiveness of the approach, as well as illuminating the role of the constraints weights in the analysis. The following mathematical results concerning the sigmoid curves are relevant.

**Asymptotes at zero and one.** Assuming sufficient room on the horizontal scale, the predicted application rate will asymptote at zero on the low end of the scale and at one on the high end. Of course, it is possible that the actual values available in the language will not permit us to see these asymptotes; the curves can be “cut off” to some degree.
Location of 50% point. The application rate of the phonological process reaches 50% at a point that is determined by the ratio of the constraint weights, specifically at \( s = \frac{w_{\text{Faith}}}{w_{\text{Markedness}}} \). Thus, for instance, the application rate for Low Harmony crosses the 50% point on our morphological closeness scale at 5.4 (= 15.2/2.8; for weights see (43)).

Symmetry. The sigmoid curve is symmetrical about the 50% point.

Steepest slope. The steepest slope of the curve occurs at the symmetry point and is determined solely by the weight of the Markedness constraint; specifically, it is equal to \( \frac{w_{\text{Markedness}}}{4} \). Thus, for instance, in our analysis the essentially categorical cut-off for ATR Harmony is captured in our analysis with a very high weight (34.8) for Agree (ATR).

We emphasize that these predictions are quite specific to the Harmonic Grammar model; other theories do not make these predictions and for them the Tommo So data are correspondingly problematic; either the model fit is bad, as with the inverted exponentials discussed in §6, or the theory is not restrictive enough and can fit essentially any data (see discussion below in §9).

We can now justify an assertion made earlier in connection with our root-closeness scale (36): it does not matter what we pick as the baseline value or interval size. As formula (45) suggests, different choices made for the root-closeness scale get canceled out when the best-fit weights are found. Thus, if we made the interval size of our scale be 2 instead of 1, the best-fit model would be one that used a value for \( w_{\text{Agree}} \) that was half as big. More generally, the expression \( w_{\text{Faith}} - w_{\text{Markedness}} \) in (45) represents a linear rescaling (i.e. of the form \( y = mx + b \)) of \( s \), which means that appropriate choices of weights exist that can compensate for any linear rescaling we might make of (36).\(^{30}\)

7.7 A variant with Noisy Harmonic Grammar

We expressed our account in maxent, but this is not the only available framework for stochastic analysis in Harmonic Grammar. In the alternative approach of Noisy Harmonic Grammar (Boersma and Pater, in press), a “noise” factor drawn from the Gaussian distribution is added to the weight of each constraint every time the grammar is applied, and the output for any given input for that particular application is simply the most harmonic candidate. The probabilities of the candidates in the general case are defined as the probability distribution over multiple applications. The best-fit constraint weights can be calculated in various ways.\(^{31}\) Mathematically, NHG generates symmetrical zero-to-one sigmoid curves just like maxent, but the sigmoids are from a different curve family (the cumulative normal distribution) that is strikingly similar visually to the logistic function employed by maxent.

---

\(^{30}\) If the scale were actually inverted in its direction we would have to use negative constraint weights (i.e., rewarding violations); so it makes sense to use a scale where closer affixes get higher values, as in (36).

\(^{31}\) In Praat (Boersma and Weenink 2014) and OTSoft (Hayes et al. 2013), weights are found using a version of Boersma’s Gradual Learning Algorithm, which employs random sampling. In the present case, since there are just two conflicting constraints, we found we could obtain a more accurate grammar by evaluating the Gaussian distributions directly using the Solver application in Excel; see Supplemental Materials for this calculation.
NHG turned out to work slightly better than maxent, with a mean absolute error of 0.009 (cf. maxent .012). For separate harmony processes the values are Low 0.005, Back 0.021, ATR 0.000.

A technical issue in NHG arises of just how the noise should be added in the case scalar constraints: in principle it could be added in before the harmony contribution is multiplied by the distance on the scale, so that noise itself gets multiplied. This produces asymmetrical sigmoids and indeed a slightly worse model fit (mean absolute error .031). The value .009 just mentioned is obtained by adding in the noise after the harmony contribution.  

7.8 Returning to the English -t,d Deletion data

It is appropriate to ask whether the English -t,d Deletion data that motivated the inverted exponential theory of Guy can be modeled under our theory. The answer is that it can. In brief, we set up a -t,d Deletion model employing the weighted constraints *CT (cluster penalty, w = .69) and Max(T) (Faithfulness, w = 1.01), with the root-closeness scale 3=root, 2=Level 1, 1=Level 2. For the Santa Ana data this gets highly accurate results.

8. Sigmoid curves in linguistic theory

Sigmoids have proven to provide good data fit in various places in linguistics, whenever similar math is used and the circumstances present here arise: specifically, when a constraint with scalar violations stacks up against a constraint with constant violations. Albright (2012), in a study of phonotactics, finds that variable phonotactic disharmony scores get a closer fit to lexical frequency when pitted against the non-scalar constraint MPARSE(). In McClelland and Vander Wyck’s (2006) study of interacting phonotactic constraints in English, a term $\beta$, analogous in effect to Albright’s MPARSE(), also improves model fit. Earlier, Frisch, Broe, and Pierrehumbert (1997) used a constant term K pitted against a scalar constraint in order to derive logistic curves that predict the frequency of consonant pairs in Arabic roots from their similarity.

The Noisy Harmonic Grammar analysis can be replicated in the framework of Stochastic OT (Boersma 1998, Boersma and Hayes 2001), provided we adopt a stratagem originated by Boersma (1998: §6, §8.4): instead of a single Markedness constraint, we adopt a family of Markedness constraints, one for each morphological level, and spaced at equal intervals along the ranking scale. This, too, will generate sigmoid curves from the cumulative normal distribution, which can be used to match the data just as well as can be done with NHG itself. The slope of the sigmoid will depend on the spacing of the members of the Markedness family, and the 50% crossing point will be determined by the relative ranking of the Markedness family with Faithfulness; see Supplemental Materials. The consequences of fragmented constraints of this kind have been little explored in the literature, and the issue of the learnability of ranking values for Stochastic OT is also problematic (Keller and Asudeh 2002, Pater 2008).

It is true that our own model has a greater number of parameters and thus would be expected to have a better fit a priori. However, the fit of the inverted-exponential model to the Tommo So data does not improve substantially when we increase its number of parameters in the most obvious way, specifically by adding an intercept term (which can raise or lower the overall height of the curve). We tried this approach and found that the mean absolute errors were still quite high: overall 0.172, Low 0.223, Back 0.134, ATR 0.156 (earlier 0.181/0.249/0.134/0.157). A reviewer also asked whether our own model could do as well if we reduced the number of parameters. We tried this first by setting Markedness to a UG-determined constant (fitting it to data from all three harmony processes at once) then by doing the same for UG-determined constant Faithfulness. The error values for the former model are 0.076/0.068/0.042/0.107 and for the latter 0.090/0.123/0.055/0.080. These are much higher than the 0.012/0.012/0.022/0.000 obtained under our maxent model.
Zuraw (2012), studying morphophonemic Nasal Mutation in Tagalog, finds that the Tagalog prefixes differ in propensity to trigger Nasal Mutation, and stem-initial consonants differ in their propensity to undergo it; each prefix gives rise to its own sigmoid curve. Much earlier, Kroch (1989) presented a model in which a syntactic constraint rises in weight over time, overtaking a set of opposed statically-weighted constraints, thus creating a family of diachronic sigmoids; the model matches well with Kroch’s diachronic data; for further discussion and examples of work inspired by Kroch see Zuraw (2003).

9. **Critiquing alternative approaches**

We have thus far considered two models that work well in fitting the Tommo So data (maxent, NHG), and also some models that work poorly (inverted exponentials and the stripped-down constraint sets of fn. 33). Here we critique two models that fit the Tommo So data perfectly, but at great cost in restrictiveness.

9.1 **Morpheme indexation**

It is likely that there are phonological processes that are triggered by specific morphemes or undergone by specific morphemes. Thus, Pater (2000, 2010) suggests that particular Markedness or Faithfulness constraints can be indexed to particular morphemes. Work supporting this view includes Flack (2007), Gouskova (2007), Steriade (2008), and Jurgec (2010).

We have found that if we affiliate a separate Markedness constraint with each of the affixes of Tommo So (along with one for roots), we can achieve a very good — indeed, perfect — fit to the Tommo So data. For instance, we factored our system of AGREE constraints to include 21 constraints total — one for each morphological type, multiplied by three since there are three harmony processes. An example is AGREE (back)\textsubscript{Facitive}. Coupled with our three Faithfulness constraints (40) and implemented in maxent, this achieves an essentially perfect fit to the data (mean error 0.0000057).

As a reviewer points out, this “would be a less restrictive model … since there would be no predicted connection between linear order of morphemes and frequency of alternation.” Indeed, the model is the least restrictive conceivable, because it can describe any frequency pattern whatsoever. For instance, in (47) we match hypothetical data in which Low Harmony follows a sawtooth frequency pattern, ATR Harmony follows a reverse sawtooth, and Backness Harmony is pyramidal. Only one line is given for each data series; the “predicted” and “observed” lines superpose because the model fit is perfect.
Fitting arbitrary data patterns with morpheme-specific constraints

Morpheme indexation does nothing to explicate the Kiparskian generalization that is our empirical focus, namely that application frequency diminishes as we approach outer morphological layers. We think our approach is a better theory because it actually predicts the Kiparskian pattern. It remains a problem to explain why morpheme-specific Markedness or Faithfulness constraints do not commonly subvert the Kiparskian pattern in ways like (47); perhaps these constraints represent something of a “last resort” option of language learners, invoked only when more natural constraints cannot deal with the data.

9.2 Domain-indexed constraints following a stringency hierarchy

A similar approach indexes constraints not to individual morphemes but rather to morphological domains. Kiparsky (1994), in his reanalysis of the English /t/-deletion pattern (see §6), uses a model of this sort. The key idea is that “a violation at any given level is necessarily a violation at all superordinate levels (but not conversely)” (p. 4). Applied to /t/-deletion, the pattern would look like (48). Such a pattern is what later came to be characterized as a stringency hierarchy (Prince 1997, de Lacy 2004) and may be related to the Strong Domain Hypothesis of Lexical Phonology (Kiparsky 1984, Myers 1991; rules may “turn off” at later levels but not “turn on”).
(48) Stringency hierarchies of Markedness constraints as defined on levels (Kiparsky 1994)

<table>
<thead>
<tr>
<th>Candidates</th>
<th>*C + t_{root}</th>
<th>*C + t_{stem}</th>
<th>*C + t_{word}</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a. [cost]</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>1b. [cos]t</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2a. [los+t]</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>2b. [los]t</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3a. [toss#t]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3b. [toss]#t</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Kiparsky stipulated that the ranking probability between each indexed Markedness constraint and the opposing Faithfulness constraints must be exactly the same. With this condition, the theory makes exactly the same predictions as the rule-based approach of Guy that inspired it (see Kiparsky 1994:4). In particular, it works well for the English t-deletion data — but by the same token, it would fail badly when applied to Tommo So (§6).

On the other hand, as a reviewer suggested, the approach might be adapted to account for Tommo So by abandoning the assumption that the ranking probability with respect to Faithfulness must remain the same for each constraint in the stringency hierarchy. Pursuing this suggestion, we set up the following hierarchy for the Tommo So markedness constraints (i.e. AGREE), as illustrated for Backness Harmony in (49); example forms are taken from §4 above.

(49) A stringency hierarchy of Markedness constraints for Tommo So Backness Harmony

<table>
<thead>
<tr>
<th>Candidates</th>
<th>AGREE-Level 1 (root)</th>
<th>AGREE-Level 2 (Fact.)</th>
<th>AGREE-Level 3 (Rev.)</th>
<th>AGREE-Level 4 (Trans.)</th>
<th>AGREE-Level 5 (Med.)</th>
<th>AGREE-Level 6 (Caus.)</th>
<th>AGREE-Level 7 (Perf.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[kõjě] (violation in root)</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>[dzɔ́bɔ́-ndé] (violation in Factitive)</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>[úm-ilé] (violation in Reversive)</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[ɔ́g-írε] (violation in Transitive)</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>[tɔ́m-ijɛ́] (violation in Mediopassive)</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>[kɛ́rɛ́-mɔ́] (violation in Causative)</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>[bɔ́d-i] (violation in Perfective)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

The essential result is that this model, coupled with any stochastic grammar framework that expresses ranking probability, can provide an exact match to the Tommo So data pattern.
Like the morpheme indexation analysis, the domain-indexed constraint model is insufficiently restrictive; it can derive any non-ascending quantitative pattern.\textsuperscript{34} To illustrate this, we give three hypothetical frequency curves below that are perfectly matched under the proposed approach, but only poorly fitted by our scalar model.

\textit{(50) Three hypothetical curves fittable using Domain-indexed constraints, but ill-fitted by the scalar model}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{hypothetical_curves.png}
\caption{Three hypothetical curves fittable using Domain-indexed constraints, but ill-fitted by the scalar model.}
\end{figure}

\textsuperscript{34} Why? Here is a simple demonstration, taking the form of a way of setting weights or ranking values that will always match the data perfectly. The rate of harmony in Level 1 (Perfective) is determined solely by the ranking or weighting of IDENT vs AGREE-Level 1; this can be matched exactly by choosing the appropriate ranking probability or weight. The rate of harmony in Level 2 (Causative) will be (by hypothesis) the same as or higher than that in Level 1; this can be matched exactly by selecting a suitable weight or ranking value for AGREE-Level 2, creating an additional penalty on the non-harmonized candidate. Proceeding inductively through the hierarchy, we can obtain an exact match for every level in succession.
The hypothetical data (diamonds) in (50a) fall on a curve that starts at 1 (at the Root level) and asymptotes to 0.5. They cannot be accurately fitted in our scalar model (best maxent fit shown with a line), because the sigmoid curves it derives can only asymptote at 0 or 1 (§7.6). The hypothetical data in (50b) involve a level sequence at 0.5 going out from the root, then descending in later levels; our scalar model cannot fit them accurately for the same reason.
Lastly, (50c) is a perfectly straight line descending from one to zero. Our scalar model cannot accurately match the zero and one at the termini because the sigmoid curves it derives must level out gradually.

9.3 Summary and evaluation of all models

In this article we have considered multiple models to account for the Tommo So vowel harmony data. By far the most successful models are those that were based on the combination of scalar constraints and Harmonic Grammar. Both the maxent version and the Noisy Harmony Grammar version achieve a close fit to the data.

Poor fits to the data, on the other hand, were obtained under a number of different approaches. The inverted-exponential model (§6) was one such case, whether one includes an intercept term or not (fn. 33). Our own Harmonic Grammar models also performed poorly if one attempts to simplify them by removing the possibility of language-specific weights for either Faithfulness or Markedness; both weights are necessary.

Lastly, both the morpheme-indexation approach (§9.1) and domain indexation with stringency (§9.2) can fit our data perfectly, but in a sense they are not interesting because they can fit either any data pattern whatsoever (morpheme indexation) or any descending pattern (domain indexation).

The chart summarizes the comparative modeling picture, giving the error values for each model.

(51) Summary of successful and unsuccessful models

<table>
<thead>
<tr>
<th>Model Description</th>
<th>Total error</th>
<th>Low error</th>
<th>Back error</th>
<th>ATR error</th>
</tr>
</thead>
<tbody>
<tr>
<td>core model in maxent</td>
<td>§7.5</td>
<td>0.012</td>
<td>0.012</td>
<td>0.022</td>
</tr>
<tr>
<td>core model in NHG</td>
<td>§7.7</td>
<td>0.009</td>
<td>0.005</td>
<td>0.021</td>
</tr>
<tr>
<td>Stochastic OT with exploded constraints</td>
<td>fn. 32</td>
<td>0.009</td>
<td>0.005</td>
<td>0.021</td>
</tr>
<tr>
<td>classic inverted exponentials</td>
<td>§6</td>
<td>0.181</td>
<td>0.249</td>
<td>0.134</td>
</tr>
<tr>
<td>inverted exponentials plus intercept</td>
<td>fn. 33</td>
<td>0.172</td>
<td>0.223</td>
<td>0.134</td>
</tr>
<tr>
<td>core model, fixed Markedness</td>
<td>fn. 33</td>
<td>0.076</td>
<td>0.068</td>
<td>0.042</td>
</tr>
<tr>
<td>core model, fixed Faithfulness</td>
<td>fn. 33</td>
<td>0.090</td>
<td>0.123</td>
<td>0.055</td>
</tr>
<tr>
<td>morpheme indexation</td>
<td>§9.1</td>
<td>zero error</td>
<td>zero error</td>
<td>zero error</td>
</tr>
<tr>
<td>domain indexation</td>
<td>§9.2</td>
<td>zero error</td>
<td>zero error</td>
<td>zero error</td>
</tr>
</tbody>
</table>

10. Further issues

10.1 What kind of variation?

Variation in phonology takes several forms. For instance, it can be either interspeaker or intraspeaker variation. In the former case, every speaker always produces the same outputs, and
the appearance of variation in the data as a whole is the merely the result of mixing in data from different speakers. We think the variation we describe here is intraspeaker variation; there is no independent evidence of any dialect differences among the consultants from whom the data were gathered, and more importantly, the collected data include numerous instances in which the very same speaker said the same word on different occasions with different harmony outcomes; see e.g. (9), (10), and (19).

Intraspeaker variation itself can involve either types or tokens. In type variation, different words or stems of similar phonological makeup behave differently on an idiosyncratic, lexically-determined basis. For example, the words of Hungarian that end in a back plus a neutral vowel differ in whether they take back or front suffixes, but normally a particular stem behaves consistently; for discussion see Hayes and Londe (2006), Hayes et al. (2009). In token variation, all words or stems that are phonologically eligible for variation do in fact vary (the same words are said differently on different occasions).

The question of which of these two kinds of variation occurs in Tommo So is difficult to assess; we report tentative conclusions based on the following calculations on our corpus. The corpus contains 828 distinct suffixed underlying representations (root plus one or more suffixes). Of these, 455 occur just once and thus cannot be used to assess free variation. Of the 373 underlying representations that occur more than once, 216 involve vowel sequences that, according to our analysis, either categorically disallow harmony or categorically require it; such UR’s likewise cannot be used to assess the type of variation. Of the remaining 157 UR’s, 39 show free variation due to optional vowel harmony. We have calculated that if harmony applied completely at random, following the probabilities of (31), there would be about 78 cases of free variation.

Our best guess at present is that Tommo So is, fundamentally, a case of token variation, with the norm being more than one possibility for any one underlying form. However, some forms — particularly those with noncompositional meaning — may be lexicalized with a particular suffix allomorph. For instance, the root /dùló/ ‘turn’ appears suffixed with Mediopassive /-íjɛ́/ a total of eight times in the data corpus, and in all eight, the suffix allomorph that appears is [-íjó], with Backness Harmony; i.e. we only get [dùl-íjó], which has the idiosyncratic meaning ‘return’. This suggests there may be a separate lexical entry for this form, something like /dùl-íjó/; for data and theory on the lexical listing of multimorphemic entries see Baayen et al. (2002), Zuraw (2000). The use of such lexically listed entries artificially inflates the observed rate of non-variation; thus the total amount of free variation ends up being below the level that would be expected under a system of pure token variation.

Lastly, we suggest that whatever sort of variation Tommo So has, our model should be responsible for it. Repeated experimental investigation has now demonstrated that even type variation gets internalized by speakers, since they tend to replicate the quantitative pattern when tested psycholinguistically in nonce probe studies; for discussion and citations see Hayes et al.

---

35 If \( n \) tokens surface derived from a UR that permits variation, with a probability of harmony of \( p \), then the expected number of forms with free variation is: \( E = 1 - p^n - (1 - p)^n \). Adding up the calculated values of \( E \) for all 157 inputs yields the expected value of 78.
(2009:826). While it has not yet proven possible to perform such testing on Tommo So speakers, there is little reason to think that the outcome of such a study would differ from what has been found for other languages.

10.2 Is harmony truly nonlocal?

We have formulated Low Harmony and Backness Harmony (§3.3) as fundamentally nonlocal processes: the trigger is the initial root vowel, but the target is sometimes a segment more than one syllable away; cf. [póó-nd-íjó] 'become bigger', where the backness of the root vowel induces a back [o] in the suffix, despite the intervening reduced vowel. Given this formulation, and the optionality of harmony, it is theoretically possible to find harmony applying in a counterintuitive way to forms with two suffixes: the outer suffix harmonizes and the inner one does not.\(^{36}\) Here is a scenario of this kind, involving Backness Harmony.

(52) “Outer but not inner” Backness Harmony (hypothetical)

\[
\begin{array}{c}
/\text{ṣẹ́n}s-ijé-mó-ì/ \\
\text{tire-MEDIOPASSIVE-CAUSATIVE-PERFECTIVE}
\end{array}
\rightarrow
\begin{array}{c}
[\text{ṣẹ́n}-ijè-m-ù]
\end{array}
\]

We have combed through the data corpus seeking cases in which this scenario could arise. Due to the “petering out” of harmony processes, the disappearance of non-final suffix vowels in hiatus, and the limitations on affix order, it turns out that there are very few. These all involve Backness Harmony, with the three-suffix combination Mediopassive-Causative-Perfective (the Causative loses its vowel in hiatus), and a back vowel in the root. There are only three such cases; two of them show no Backness Harmony and the third has inner harmony but not outer harmony (this is the real-world counterpart of (52), which in fact was pronounced [ṣẹ́n-ijè-m-ì]).

We are reluctant to draw any conclusions from such limited data. Estimating application probabilities from the data (\((31)\)), we calculate that the probability of scenario (52) arising independently of issues of locality is only .076 \((= (1-.44) \times .136)\), meaning there is about a 79% chance that such cases would be absent purely by accident \((= 1 - (1-.076)^3)\). Thus the data do not suffice to determine whether the absence of real cases like (52) is meaningful.

10.3 Opacity with Hiatus Resolution

ATR Harmony in Tommo So is opaque: it is in a counterbleeding relationship with Hiatus Resolution. In roots with an initial high vowel and a final mid vowel, the only possible ATR trigger is the final vowel, since only mid vowels may trigger ATR Harmony (§4.4). When a vowel-initial suffix is added to such a root, this final vowel deletes, yet the vowel of the suffix agrees in [ATR] as though it were present. Examples are given in (53).

\(^{36}\) Thanks to Anonymous for pointing this out to us.

\(^{37}\) The morphemes /ṣẹ́nS/ and /-mS/ have final vowels that are removed by Hiatus Resolution (11). Further, surface tones are actually the result of a low tone overlay (McPherson 2013, §12.4.1).
(53) a. Deleted [+ATR] root vowel: suffix is [+ATR]

/\texttt{didé-ilé}/ \rightarrow /\texttt{did-ilé}/ \quad \text{‘prop.up-REVERSIVE’ = ‘remove prop’}
/\texttt{kúmbó-iré}/ \rightarrow /\texttt{kúmb-iró}/ \quad \text{‘fist-TRANSITIVE’ = ‘put in (sb’s) fist’}
/\texttt{wigílé-ijé}/ \rightarrow /\texttt{wigíl-ijé}/ \quad \text{‘swing-MEDIOPASSIVE’ = ‘swing’}

b. Deleted [−ATR] root vowel: suffix is [−ATR]

/\texttt{diŋé-ilé}/ \rightarrow /\texttt{diŋ-ilé}/ \quad \text{‘tie-REVERSIVE’ = ‘untie’}
/\texttt{túŋ-iré}/ \rightarrow /\texttt{túŋ-iró}/ \quad \text{‘kneel-TRANSITIVE’ = ‘make (sb) kneel’}
/\texttt{timbé-ijé}/ \rightarrow /\texttt{timb-ijé}/ \quad \text{‘stack-MEDIOPASSIVE’ = ‘become stacked’}

As these examples show, the [ATR] value on the suffix is determined by an underlying vowel no longer present on the surface. In a rule-based analysis, this result would be obtainable by ordering ATR Harmony before Hiatus Resolution.

In constraint-based phonology, there are several theories that would permit the derivation of these opaque cases. Employing autosegmental theory, we could let vowel deletion strand a floating Root node (Clements 1985). Provided that this node is visible to the AGREE constraints, opacity would result. Opacity could also be obtained by framing our analysis within various standard approaches within OT: Sympathy Theory (McCarthy 1999), Turbidity Theory (Goldrick 1999), Stratal OT (Kiparsky 2000, 2008; Bermudez-Otero 1999), Candidate Chain Theory (McCarthy 2007). Since the analysis of opacity is tangential to our main concerns, we will not explore the matter further.

10.4 Are vowel-zero alternations the result of epenthesis?

Whenever we see vowel-zero alternations, as in Tommo So, it is appropriate to ask whether they are due to deletion (as we propose here) or to epenthesis. Consider for instance an alternation from (53), [\texttt{díd}-\texttt{é}] \sim \texttt{\textit{didd}-ilé} \ ‘prop.up’ \sim \texttt{\textit{prop.up-REVERSIVE}}. While we have assumed that the underlying form of the root is /\texttt{didé}/, with Hiatus Resolution occurring before the initial vowel of [\texttt{ilé}], another possibility is that the root ‘prop up’ is actually underlying /\texttt{did}/, and that the appearance of [\texttt{é}] is due to some kind of epenthesis.

Standard considerations militate against this idea. Specifically, vowel quality in the putatively epenthetic position is contrastive. Thus, for instance, there is a near-minimal pair [\texttt{tínd}-\texttt{é}] \sim \texttt{\textit{tind}-ilé} \ ‘block.passage(-MEDIOPASSIVE)’ \vs. [\texttt{timbé}-\texttt{é}] \sim \texttt{\textit{timb}-ijé} \ ‘double.up(-MEDIOPASSIVE)’. Such cases are not unusual in Tommo So.

The epenthesis analysis could be rescued if we are willing to set up abstract vowels in underlying forms. So, for instance, if [\texttt{timbé}] has an underlying [−ATR] vowel in its initial syllable, this could trigger the appearance of epenthetic [\texttt{ê}] in the surface form, with some other process assumed that would neutralize the [ATR] distinction among high vowels on the surface.
In contrast, [tìndé] would have an underlying [+ATR] stem vowel, causing its second vowel to surface as [+ATR]. In sum, under this analysis /tìmb/ gives rise to [tìmbɛ́] while /tìnd/ gives rise to [tìndɛ́].

The abstract analysis strikes us as a rather costly move. Abstract segments are hardly needed to derive the observed surface pattern; we know that [dìd-ilé] should have the suffix allomorph [-ilé] because the stem in isolation is pronounced [dìdɛ́]; and that likewise [dìŋ-ilɛ́] should have the suffix allomorph [-ilɛ́] because its stem in isolation is pronounced [dìŋɛ́]. This is opaque phonology to be sure, but it is not the sort of pattern that has traditionally been used to justify abstract underlying vowels.38

Finally, we note that even if Tommo So were to be analyzed with pervasive epenthesis, the problem of relating vowel harmony frequency to morphological structure would remain unchanged.

11. Conclusion

We list what we take to be the main results of this work.

First, Tommo So vowel harmony is a fairly clear case of the Kiparskian generalization that phonological processes “turn off” as they extend into the outer reaches of the morphology. With a rich set of ordered affixes, Tommo So forms a test case for formal theories that seek to account for this phenomenon in general.

Second, Guy’s inverted-exponential theory, though simple and principled, does not provide a good fit to the Tommo So data. Other models of phonology-morphology interactions, namely morpheme-indexed constraints and domain-indexed constraints with a stringency hierarchy, are able to provide a essentially perfect match to the data, but this result is uninformative because these models are unrestrictive.

Third, Harmonic Grammars (both maxent and NHG), augmented to include scalar constraints, provide a good fit to our data. Moreover, these are highly restrictive theories; they inherently generate sigmoid probability functions and can only be used to model curves that fit this family.

This said, we judge that our work remains speculative until a greater number of “petering out” systems have been located and documented with enough data to do quantitative analysis. We hope the method described here will facilitate the task of taking on new cases.

38 Thanks to an anonymous Phonology reviewer for pointing out the possibility of an epenthesis analysis for Tommo So. The epenthesis-cum-abstract-vowels account described above follows in outline the proposals of Hantgan and Davis (2012) for the related language Bondu-so. The Bondu-so data pattern appears to be rather different from that of Tommo So; we make no claims about the applicability of our analysis to any other Dogon language.
References

Albright, Adam (2012). Additive markedness interactions in phonology. Ms., Department of Linguistics and Philosophy, MIT.


Fruehwald, Josef T. (2012). Redevelopment of a morphological class. University of Pennsylvania Working Papers in Linguistics Vol. 18: Iss. 1, Article 10. Available at: http://repository.upenn.edu/pwpl/vol18/iss1/10


Tommo So Vowel Harmony


Kirchner, Robert (1993). Turkish vowel harmony and disharmony: an optimality theoretic account. Paper presented at the Rutgers Optimality Workshop 1, Rutgers University, New Brunswick, NJ.


Zuraw, Kie (2012). Quantitative patterns of constraint interaction. Talk given at WCCFL 30, University of California, Santa Cruz.