Learnability of derivationally opaque processes in the Gestural Harmony Model

Caitlin Smith
Department of Cognitive Science
Johns Hopkins University
csmit372@jhu.edu

Charlie O'Hara
Department of Linguistics
University of Southern California
charleso@usc.edu

Abstract

In this paper, we examine the learnability of two apparently derivationally opaque vowel harmony patterns: attested stepwise height harmony and unattested saltatory height harmony. We analyze these patterns within the Gestural Harmony Model (Smith, 2018) and introduce a learning algorithm for setting the gestural parameters that generate these harmony patterns. Results of the learning model indicate a learning bias in favor of the attested stepwise pattern and against the unattested saltation pattern, providing a potential explanation for the differences in attestation between these two derivationally opaque patterns.

1 Introduction

In stepwise partial height harmony, nonhigh vowels raise one step along a height scale in the presence of a high vowel trigger. Such harmony processes resemble chain shifts, a type of underapplication opacity. This is illustrated by the vowel raising harmony of Nzebi, a Bantu language spoken in Gabon (Guthrie, 1968; Kirchner, 1996; Parkinson, 1996; Smith, 2020). In Nzebi, the suffix /-i/ occurs immediately after verb roots in some tenses and triggers one-step raising of preceding root vowels. Before this harmony triggering suffix, high-mid vowels /e/ and /o/ surface as [i] and [u], respectively; low-mid vowels /ɛ/ and /ɔ/ surface as [e] and [o], respectively; and low /a/ surfaces as [ɛ]. This is illustrated by the data in (1).

(1) Root vowels in non-raising vs. raising contexts

a.	[be̲t]	[bi̲t-i]	'carry'
b.	[βo̲ːm]	[βu̲ːm-i]	'breathe'
c.	[sɛ̲b]	[se̲b-i]	'laugh'
d.	[mɔ̲n]	[mo̲n-i]	'see'
e.	[sa̲l]	[sɛ̲l-i]	'work'

The chain-shifting nature of this pattern of stepwise vowel raising is illustrated in Figure 1. This represents a case of underapplication opacity as the process that raises underlying high-mid vowels to high vowels appears to have underapplied to high-mid vowels that were derived from low-mid vowels. Likewise, the process that raises underlying low-mid vowels to high-mid vowels has underapplied to low-mid vowels that were derived from low vowels.

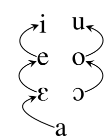

Figure 1: Pattern of vowel raising in the stepwise height harmony of Nzebi

Another example of underapplication opacity is saltation, a type of phonologically derived environment effect. A hypothetical case of saltatory height harmony is provided in Figure 2. Low-mid vowels raise to reach the trigger's height (/ɛ-i/ → [i-i], apparently 'skipping over' more faithful [e-i]), while a high-mid vowel does not raise (/e-i/ → [e-i], not [i-i]). the opacity of this pattern is apparent when we conceptualize two-step raising as the application of two separate raising processes: underlying low-mid vowels raise first to high-mid and then to high, while underlying high-mid vowels are not raised by the latter of these two processes. The high-mid to high raising process has therefore underapplied to underlying high-mid vowels.

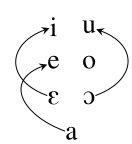

Figure 2: Pattern of unattested saltatory two-step vowel raising

While stepwise height harmonies are well attested, saltatory processes are rare in general and (to our knowledge) unattested in height harmony.

Proceedings of the Society for Computation in Linguistics (SCiL) 2021, pages 396-400.
Held on-line February 14-19, 2021

Because they are derivationally opaque, neither stepwise nor saltatory height harmony patterns can be generated in output-oriented Harmonic Grammar assuming standard faithfulness constraints and standard features (Albright et al., 2008; Farris-Trimble, 2008; Hayes and White, 2015).

The Gestural Harmony Model (Smith, 2016, 2018), on the other hand, is sufficiently powerful to generate both of these apparently derivationally opaque patterns. In this paper, we illustrate that while this model is able to generate unattested saltatory height harmony, this overgeneration can be resolved by appealing to factors beyond the grammatical model itself, such as the learnability of the pattern. We show that when analyzed within the Gestural Harmony Model, stepwise and saltatory harmony exhibit significant differences in their learnability, providing an explanation for the difference in their typological attestation.

2 A Gestural Model of Height Harmony

Gestures are the dynamically-defined, task-based units of sub-segmental representation assumed within the framework of Articulatory Phonology (Browman and Goldstein, 1986, 1989). Each gesture is specified for a target articulatory state to be achieved during its period of activation; this target is specified in terms of a primary articulator, constriction location, and constriction degree. The Gestural Harmony Model (Smith, 2016, 2018) adopts many of the representational assumptions of Articulatory Phonology. In this model, harmony is the result of a gesture extending its period of activation to overlap the gestures of surrounding segments. In the case of regressive (leftward) harmony, an anticipatory, or early-activating, trigger gesture activates before its scheduled starting point, extending to overlap the gestures of preceding undergoer segments.

Gestural overlap may result in the concurrent activation of two gestures with conflicting target articulatory states (e.g., narrow vs. wide constriction degree between the tongue body and the upper surface of the vocal tract). According to the Task Dynamic Model of speech production (Saltzman and Munhall, 1989; Fowler and Saltzman, 1993), intergestural conflict is resolved by blending the conflicting target articulatory states of two gestures to create an intermediate target state that holds during the period of their concurrent activation. This blended target state is the weighted average of the gestures' individual target articulatory states, with the weighting in this averaging function contributed by the gestures' strength parameters, denoted α. This blending function is provided in Equation 1.

$$\frac{Target_1 \times \alpha_1 + Target_2 \times \alpha_2}{\alpha_1 + \alpha_2} \tag{1}$$

Smith (2020) proposes that the partial, stepwise vowel raising harmony of Nzebi is the result of gestural blending resulting from overlap of root vowels by the anticipatory, harmony-triggering tongue body gesture of the high suffix vowel /-i/. In this analysis, the four vowel heights observed in Nzebi are represented by vowel gestures with one of four possible constriction degrees between the tongue body and the upper surface of the vocal tract: narrow (4mm), narrow-mid (8mm), wide-mid (12mm), and wide (16mm).

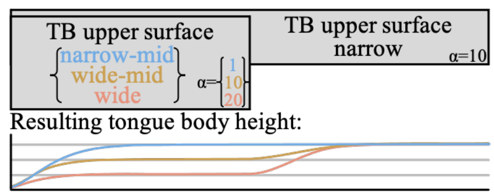

Figure 3: Stepwise raising via gestural blending

With appropriate blending strengths for these vowels, the stepwise raising pattern of Nzebi is generated from overlap between vowel gestures, as in Figure 3. When overlapped by suffix /-i/, high-mid root vowels surface as high and do not resist raising at all, indicating that they have a blending strength much lower than that of trigger /-i/. Wide-mid vowels, however, raise to only an intermediate degree when overlapped by harmony-triggering /-i/, indicating that /ɛ/ and /ɔ/ have blending strengths equal to that of /-i/. Finally, low /a/ also raises only partially when overlapped by /-i/. With /a/ specified for twice the strength of the trigger gesture that overlaps it, the result of blending is wide-mid [ɛ], closer to the intrinsic target constriction degree of wide /a/ than to that of narrow /-i/.

With a different set of strength value settings, the Gestural Harmony Model is also able to generate unattested saltatory height harmony. Figure 4 illustrates and provides a set of gestural strengths that will generate a two-step raising process consistent with the one depicted in Figure 2.

The Gestural Harmony Model, then, is able to generate two types of apparently derivationally opaque height harmony patterns. However, on its

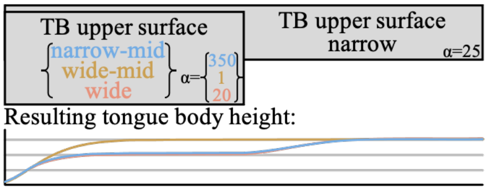

Figure 4: Stepwise raising via gestural blending

own the model does not provide an explanation for the lack of attestation of saltatory height harmony. In the next section, we address the question of whether these two patterns are equally learnable when analyzed in this framework.

3 The Gestural Gradual Learning Algorithm

With the correct gestural strength settings, the Gestural Harmony Model can generate both attested stepwise and unattested saltatory height harmonies. However, there are additional factors beyond what patterns a grammar can generate that impact how likely a given pattern is to be crosslinguistically attested. Recent work has explored the hypothesis that learning biases can affect whether a pattern is common or rare (Pater and Moreton, 2012; Staubs, 2014; Stanton, 2016; Hughto, 2019; O'Hara, 2021). If a pattern is difficult to learn, it is more likely to be mislearned and to change across generations, eventually becoming typologically underrepresented.

To address the question of whether a learning bias is responsible for the lack of attestation of saltatory height harmony, we designed a gesture-based computational learning model. The learner was tasked with setting constriction degree targets and blending strengths for vowels and dorsal consonant gestures such that the learner reproduced its teacher's height harmony pattern. We tested the learner on two types of harmony in a four-height vowel inventory: (1) a Nzebi-like pattern of one-step raising before high vowel triggers, and (2) an unattested pattern of two-step saltatory raising before high vowel triggers.

The learner utilizes a learning algorithm that we introduce here: the Gestural Gradual Learning Algorithm (GGLA), which is defined as in (2).

(2) The Gestural Gradual Learning Algorithm

1. For each segment, initialize segment target constriction degree of 16 mm (i.e., all vowels start as [a]) and random strength (between 1 and 20)
2. On each training iteration, randomly generate (V)CV sequence
3. Check for gestural blending:
   (i) If V_2 is a trigger of harmony, it overlaps V_1 and blending occurs
   (ii) If C is dorsal /g/, blending with following V occurs
4. If learner produces error (segment with target farther than 0.2 mm from teacher's production):
   (i) Update constriction degree target of learner's tongue body gesture by 0.1 to produce a constriction degree that better matches teacher's output
   (ii) In cases of blending: update strength of learner's tongue body gesture by 0.1 to produce a constriction degree that better matches teacher's output

We trained 100 GGLA models each on the target stepwise and saltatory harmony patterns. All models were trained until convergence, which occurred when all (V)CV sequences were produced without errors (i.e. with every segment's constriction degree produced within 0.2 mm of the teacher's production).

4 Results and Discussion

We compared the average number of training iterations necessary for the learner to converge upon each target pattern; the result is shown in Figure 5. We observe that stepwise height harmony was learned more than five times faster than saltatory harmony. We interpret this result as an explanation

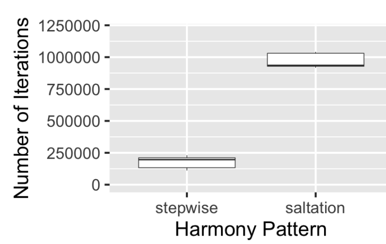

Figure 5: Average number of iterations necessary to converge upon stepwise and saltatory height harmony

for the attestation of stepwise height harmony and the lack of attestation of saltatory height harmony. The difference in learning rate between these two types of harmony makes the saltatory pattern more likely to be mislearned across generations, causing it to become less typologically frequent.

The reason why saltatory harmony is more time consuming (i.e. harder) to learn lies in the more extreme strength values necessary to generate it. Due to the GGLA's linear updating of gestural parameter settings, a learner will take longer to set more extreme strength values. To better understand why saltatory harmony crucially relies on more extreme gestural strengths, we introduce the idea of overpowering relationships between blended gestures. In order for one gesture to fully assimilate to another, or to fully resist assimilation, one of the gestures must overpower the other by being specified for an exponentially greater strength value. Figure 6 presents the average gestural strengths learned for each segment for each target harmony pattern rounded to the nearest integer. Overpowering relationships between segments are represented with arrows. For one segment to overpower another, it must have a gestural strength approximately ten times greater. When one segment partially assimilates to another, their strengths are more similar, differing by a factor of one or two rather than ten.

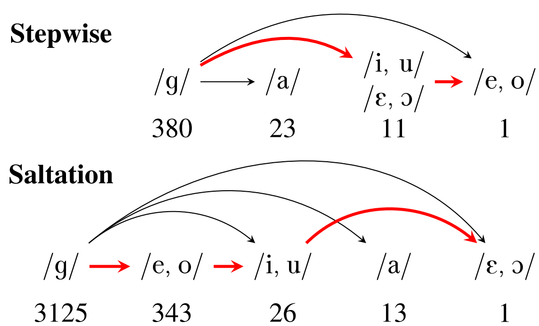

Figure 6: Segments' average learned strengths and overpowering relationships necessary to generate stepwise and saltatory harmony. Red arrows indicate a pattern's longest chain of overpowering relationships.

In order to determine how extreme the greatest segmental strength value must be in order to produce a certain pattern, we can examine chains of overpowering relationships that exist between gestures of different strengths. These chains are indicated by red arrows in Figure 6. Each additional overpowering relationship in a chain indicates an order of magnitude of strength that must be reached by the strongest gesture in the chain.

In both harmony patterns, the dorsal consonant /g/ must overpower all vowels. In every (V)CV sequence in which /g/ appears, it is blended with V_2 but always completely resists assimilation. Overpowering relationships between vowels always involve the narrow trigger vowels /i/ and /u/, which overlap and blend with preceding vowels. In saltatory height harmony, narrow vowels must be overpowered by narrow-mid /e/ and /o/, as these vowels completely resist assimilation to the narrow vowels. In addition, the narrow vowels must themselves overpower wide-mid /ɛ/ and /ɔ/, as those vowels completely assimilate to the narrow vowels. Therefore, the longest chain of overpowering relationships in the saltatory harmony pattern is three links long, indicating that the strongest gesture in the inventory must be about three orders of magnitude stronger than the weakest gesture.

In stepwise height harmony, no vowels completely resist assimilation, and only the narrow-mid vowels /e/ and /o/ assimilate completely to the high vowel triggers of harmony. Therefore, the high vowels must only overpower the narrow-mid vowels, and the longest chain of overpowering relations in the stepwise harmony pattern is two links long. This indicates that the strongest segment in the inventory must be about two orders of magnitude stronger than the weakest segment.

These differences in the extremeness of gestural strengths necessary to generate saltatory and stepwise height harmony have consequences for the rates of learning for these patterns. Saltatory harmony requires a longer chain of overpowering relationships than stepwise harmony, and therefore more extreme strength values. As a result, learning of saltatory height harmony takes much longer for a learner utilizing the GGLA.

5 Conclusion and Future Work

The Gestural Harmony Model is sufficiently powerful to generate not only common stepwise height harmony, but also unattested saltatory height harmony. We have shown that for a learner relying on the GGLA, unattested saltatory height harmony is significantly slower and harder to learn than attested stepwise height harmony. We argue that this difference in learnability can explain the difference in attestation between these two opaque patterns without restricting the power of the Gestural Harmony Model. Future work will compare the learnability of these patterns in the Gestural Harmony Model to their learnability in feature-based models of grammar that are capable of generating opaque patterns, such as those incorporating scalar features (Gnanadesikan, 1997) and *MAP constraints (Zuraw, 2007; Hayes and White, 2015).

Acknowledgements

For helpful discussion of this work, we thank Canaan Breiss, Brian Hsu, Larry Hyman, Karen Jesney, Jack Merrill, Elliott Moreton, Eric Rosen, Jennifer L. Smith, Paul Smolensky, and Rachel Walker; audiences at AMP 2020 and SCiL 2021; the members of Phonlunch at USC; and the members of Zoom Phonology, organized by Karee Garvin and Hannah Sande. This work is supported by Microsoft Research.

References

Adam Albright, Giorgio Magri, and Jennifer Michaels. 2008. Modeling doubly marked lags with a split additive model. In *Proceedings of the 32nd Annual Boston University Conference on Language Development*, pages 36–47. Cascadilla Press.

Catherine P. Browman and Louis Goldstein. 1986. Towards an articulatory phonology. *Phonology Yearbook*, 3:219–252.

Catherine P. Browman and Louis Goldstein. 1989. Articulatory gestures as phonological units. *Phonology*, 6(2):201–251.

Ashley Farris-Trimble. 2008. *Cumulative Faithfulness Effects in Phonology*. Ph.D. thesis, Indiana University.

Carol A. Fowler and Elliot Saltzman. 1993. Coordination and coarticulation in speech production. *Language and Speech*, 36:171–195.

Amalia Gnanadesikan. 1997. *Phonology with Ternary Scales*. Ph.D. thesis, University of Massachusetts Amherst.

Malcolm Guthrie. 1968. Notes on Nzɛbi (Gabon). *Journal of African Languages*, 7(2):101–129.

Bruce Hayes and James White. 2015. Saltation and the P-map. *Phonology*, 32:267–302.

Coral Hughto. 2019. *Emergent Typological Effects of Agent-Based Learning Models in Maximum Entropy Grammar*. Ph.D. thesis, University of Massachusetts Amherst.

Robert Kirchner. 1996. Synchronic chain shifts in Optimality Theory. *Linguistic Inquiry*, 27(2):341–350.

Charlie O'Hara. 2021. *Soft Biases in Phonology: Learnability meets grammar*. Ph.D. thesis, University of Southern California.

Frederick Parkinson. 1996. *The Representation of Vowel Height in Phonology*. Ph.D. thesis, The Ohio State University.

Joe Pater and Elliott Moreton. 2012. Structurally biased phonology: complexity in language learning and typology. *The EFL Journal*, 3(2):1–44.

Elliot Saltzman and Kevin G. Munhall. 1989. A dynamical approach to gestural patterning in speech production. *Ecological Psychology*, 1(4):333–382.

Caitlin Smith. 2016. A gestural account of neutral segment asymmetries in harmony. In *Proceedings of the 2015 Annual Meeting on Phonology*.

Caitlin Smith. 2018. *Harmony in Gestural Phonology*. Ph.D. thesis, University of Southern California.

Caitlin Smith. 2020. Stepwise height harmony as partial transparency. In *Proceedings of the 50th Annual Meeting of the North East Linguistic Society*, pages 141–154. GLSA.

Juliet Stanton. 2016. Learnability shapes typology: the case of the midpoint pathology. *Language*, 92(4):753–791.

Robert Staubs. 2014. *Computational modeling of learning biases in stress typology*. Ph.D. thesis, University of Massachusetts Amherst, Amherst.

Kie Zuraw. 2007. The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog infixation. *Language*, 83(2):277–316.