MaxEnt fails at reasoning by transitivity

Giorgio Magri & Arto Anttila
CNRS & Stanford University

1 Introduction

Generative phonology has taken a probabilistic turn (see for instance Coetzee & Pater 2011 and Pierrehumbert 2022 for overviews). At the empirical level, categorical data collected through introspection and field work are now complemented with quantitative data extracted from corpora and experiments. At the theoretical level, analysts are moving from **categorical** OT grammars (Prince & Smolensky 2004) and harmonic grammars (HG; Potts et al. 2010) to **probabilistic** grammars that make quantitative predictions, such as maximum entropy grammars (ME; Hayes & Wilson 2008; Wilson 2025) and noisy harmonic grammars (NHG; Boersma & Pater 2016). These developments bring along new analytical challenges. The analysis of the typological structure predicted by categorical OT and HG is relatively easy because these theories predict finite factorial typologies. Probabilistic ME and NHG typologies instead always consist of infinitely many grammars. As a result, new tools are needed to analyze these probabilistic typologies.

In series of recent papers, Anttila, Magri, and their collaborators (AM; Anttila & Magri 2018; Anttila et al. 2019; Magri & Anttila 2023; Ikwut-Ukwa et al. 2025; Magri & Anttila 2025, 2026) develop one such tool. They extend the Greenbergian **implicational universals** from the categorical to the probabilistic setting through probability inequalities that must be satisfied by every probabilistic grammar in the typology considered, as reviewed in section 2. To illustrate, it has been reported (Guy 1980; Coetzee & Kawahara 2013) that the frequency of t-deletion in English coda clusters is always smaller before vowels (/cost us/ rarely surfaces as [cos us] than before consonants (/cost me/ often surfaces as [cos me]). The implication (/cost us/, [cos us]) → (/cost me/, [cos me]) captures this observation because it qualifies a ME or NHG universal provided that the ME or NHG probability of t-deletion is always smaller in the pre-vocalic environment than in the pre-consonantal environment, no matter how we choose the constraint weights.

These implicational universals reveal a surprising difference between NHG and ME: while any ME implicational universal is also an NHG universal, the reverse fails. To establish this difference, AM develop a number of necessary constraint conditions on ME implicational universals and use these conditions to systematically detect NHG implicational universals that fail in ME. Qualitatively, these failures are often counterintuitive. In particular, the distinction between the universals that survive in ME and those that fail seems to follow no sensible phonological rationale. Quantitatively, these failures can be so massive as to wipe away all implicational universals that deal with markedness asymmetries (Ikwut-Ukwa et al. 2025).

Our paper contributes to this research program. Section 3 spells out a new formal constraint condition on ME implicational universals that strengthens a condition previously obtained by AM. The proof of this new result is relegated to an appendix. Section 4 illustrates the phonological implications of our strengthened condition. We focus on implicational universals that *hold by transitivity* in categorical OT and HG as well as in probabilistic NHG. We use our new strengthened condition to explain why this class of universals instead fails in probabilistic ME. Finally, we observe that AM's weaker condition is unable to diagnose this ME failure. Section 5 provides additional examples of implicational universals that hold by transitivity and thus fail in ME, suggesting that these examples can be easily multiplied. Section 6 concludes that ME is not a satisfactory mode of constraint interaction.

* Financial support from the France-Stanford Center for Interdisciplinary Studies is gratefully acknowledged.

Proceedings of *AMP 2025*

Before developing our argument, we sketch informally what we mean by *reasoning by transitivity*. We consider some implicational universal antecedent → consequent such that the antecedent mapping comes in particular with two losers that can be described in ERC notation (Prince 2002) as in (1a) while the consequent mapping comes in particular with the loser in (1b). In order for the antecedent mapping to beat the first antecedent loser in OT, constraint C_1 must outrank constraint C_2. In order for the antecedent mapping to also beat the second antecedent loser, C_2 must in turn outrank C_3. By transitivity, C_1 must outrank C_3. This ranking $C_1 \gg C_3$ ensures that the consequent mapping beats its consequent loser. As a result, every OT grammar that contains the antecedent mapping, also contains the consequent mapping, as required by the Greenbergian implicational universal antecedent → consequent. We conclude that this implication is an OT universal that holds *by transitivity*.

(1) a.

antecedent mapping	C_1	C_2	C_3
first antecedent loser	W	L	
second antecedent loser		W	L

b.

consequent mapping	C_1	C_2	C_3
consequent loser	W		L

Analogous considerations hold in HG (when the ERCs correspond to reasonable actual numbers of constraint violations). Finally, this reasoning by transitivity extends to NHG because AM show that categorical HG and probabilistic NHG always share the same universals: every NHG grammar assigns at least as much probability to the consequent mapping as to the antecedent mapping. The situation is very different in ME: sections 4 and 5 use the new formal result of section 3 to show that all such examples of universals that hold in OT, HG, and NHG by transitivity actually fail in ME: the ME probability of the antecedent mapping corresponding to some non-negative constraint weights is larger than the ME probability of the consequent mapping, flouting the implicational universal antecedent → consequent.

2 Implicational universals of categorical and probabilistic phonology

This section briefly reviews the HG and NHG/ME implementations of categorical and probabilistic constraint-based phonology together with the corresponding notions of implicational universals.

2.1 *HG, NHG, and ME grammars* A phonological mapping is a pair (x, y) consisting of some underlying form x (such as /cost me/) and some surface form y (such as [cos' me]). As usual, we list in Gen all the phonological mappings that are relevant for the description of the phonological system of interest. For every underlying form x, we collect into the set $Gen(\mathsf{x})$ all the surface forms y that are candidate surface realizations of that underlying form x because Gen contains the mapping (x, y). A constraint set $\boldsymbol{C}$ lists a finite number n of constraints $C_1, \ldots, C_n$ for the mappings listed by Gen. The number of violations assigned by a constraint C_k to a mapping (x, y) is denoted by $C_k(\mathsf{x}, \mathsf{y})$. These numbers are collected into the constraint violation vector $\boldsymbol{C}(\mathsf{x}, \mathsf{y}) = (C_1(\mathsf{x}, \mathsf{y}), \ldots, C_n(\mathsf{x}, \mathsf{y}))$ of the mapping (x, y).

A weight vector $\boldsymbol{w} = (w_1, \ldots, w_n)$ assigns a non-negative weight w_k to each constraint C_k. The relative size of a weight w_k quantifies the relative importance of the corresponding constraint C_k. The product $\boldsymbol{w} \cdot \boldsymbol{C}(\mathsf{x}, \mathsf{y})$ between the weight vector $\boldsymbol{w}$ and the constraint violation vector $\boldsymbol{C}(\mathsf{x}, \mathsf{y})$ is the sum of the products of the corresponding entries, namely $\boldsymbol{w} \cdot \boldsymbol{C}(\mathsf{x}, \mathsf{y}) = \sum_{k=1}^{n} w_k C_k(\mathsf{x}, \mathsf{y})$. Equivalently, this product $\boldsymbol{w} \cdot \boldsymbol{C}(\mathsf{x}, \mathsf{y})$ is the weighted sum of the constraint violations of the phonological mapping (x, y). Since the weights are non-negative, the size of this product quantifies how badly the mapping of the underlying form x to the surface candidate y violates the constraints, taking their relative importance into account.

The categorical **HG grammar** $G_{\boldsymbol{w}}$ realizes an underlying form x as the unique candidate winner surface form $\mathsf{y} = G_{\boldsymbol{w}}(\mathsf{x})$ from the candidate set $Gen(\mathsf{x})$ that violates the constraints the least because the inequality $\boldsymbol{w} \cdot \boldsymbol{C}(\mathsf{x}, \mathsf{y}) < \boldsymbol{w} \cdot \boldsymbol{C}(\mathsf{x}, \mathsf{z})$ holds for every other surface realization z from the candidate set $Gen(\mathsf{x})$. Next, we consider some probability density function $p_{\boldsymbol{w}}$ among n-dimensional vectors that depends on a weight vector $\boldsymbol{w}$.[1] The probabilistic **NHG grammar** $G_{\boldsymbol{w}}$ predicts that a given underlying form x is realized as a surface form y from the candidate set $Gen(\mathsf{x})$ with a probability $G_{\boldsymbol{w}}(\mathsf{y} \mid \mathsf{x})$ equal to the probability of sampling according

[1] The NHG literature assumes that $p_{\boldsymbol{w}}$ is a gaussian density with mean $\boldsymbol{w}$. AM assume instead that $p_{\boldsymbol{w}}$ is the exponential density that starts at $\boldsymbol{w}$, defined as $p_{\boldsymbol{w}}(v_1 \ldots v_n) = \prod_{k=1}^{n} \exp(v_i - w_i) \mathbb{I}_{[w_i, +\infty)}(v_i)$. The choice of the gaussian density runs into the problem of negative HG weights (Hayes & Kaplan 2023). The choice of the exponential density avoids that problem because it is concentrated "at the right" of the non-negative weight vector $\boldsymbol{w}$.

to the density $p_{\boldsymbol{w}}$ some weight vector $\boldsymbol{v} = (v_1, \ldots, v_n)$ such that the corresponding HG grammar $G_{\boldsymbol{v}}$ realizes that underlying form x as that surface form y. Finally, the probabilistic **ME grammar** $G_{\boldsymbol{w}}$ predicts that a given underlying form x is realized as a surface form y from the candidate set $Gen(\mathsf{x})$ with a probability $G_{\boldsymbol{w}}(\mathsf{y}\,|\,\mathsf{x})$ proportional to the positive score $\exp(-\boldsymbol{w} \cdot \boldsymbol{C}(\mathsf{x},\mathsf{y}))$. The categorical HG typology $\mathfrak{T}_{\mathrm{HG}}(Gen, \boldsymbol{C})$ and the probabilistic NHG and ME typologies $\mathfrak{T}_{\mathrm{NHG}}(Gen, \boldsymbol{C})$ and $\mathfrak{T}_{\mathrm{ME}}(Gen, \boldsymbol{C})$ collect together the HG, NHG, and ME grammars $G_{\boldsymbol{w}}$ corresponding to all vectors $\boldsymbol{w}$ of non-negative constraint weights.

2.2 *HG, NHG, and ME implicational universals* How can one study typologies that can be very large (HG) or infinite (NHG, ME)? Here is a natural answer: instead of inspecting the grammars listed by these typologies, we extract the **universals** they predict, namely the properties that hold true of every single grammar in the typology and thus plausibly explain why those grammars have been bundled together into a typology. To implement this intuition, AM focus on implications $(\mathsf{x},\mathsf{y}) \rightarrow (\widehat{\mathsf{x}},\widehat{\mathsf{y}})$ between two mappings (x,y) and $(\widehat{\mathsf{x}},\widehat{\mathsf{y}})$ from Gen. They say that this implication is a universal of the categorical HG typology $\mathfrak{T}_{\mathrm{HG}}(Gen, \boldsymbol{C})$ provided every HG grammar that realizes the antecedent underlying form x as the antecedent surface form y, also realizes the consequent underlying form $\widehat{\mathsf{x}}$ as the consequent surface form $\widehat{\mathsf{y}}$, whereby the material implication (1) holds universally for any vector $\boldsymbol{w}$ of non-negative constraint weights. To illustrate, (/cost us/, [cos us]) → (/cost me/, [cos me]) is an HG universal if any HG grammar that applies t-deletion categorically before a vowel also applies it before a consonant. Anttila (2018) provides an overview.

(1) $$\text{If } \overbrace{G_{\boldsymbol{w}}(\mathsf{x}) = \mathsf{y}}^{\mathsf{x}\text{ is realized as }\mathsf{y}}, \text{ then also } \underbrace{G_{\boldsymbol{w}}(\widehat{\mathsf{x}}) = \widehat{\mathsf{y}}}_{\widehat{\mathsf{x}}\text{ is realized as }\widehat{\mathsf{y}}}$$

(2) $$\overbrace{G_{\boldsymbol{w}}(\mathsf{y}\,|\,\mathsf{x})}^{\text{probability that }\mathsf{x}\text{ is realized as }\mathsf{y}} \leq \underbrace{G_{\boldsymbol{w}}(\widehat{\mathsf{y}}\,|\,\widehat{\mathsf{x}})}_{\text{probability that }\widehat{\mathsf{x}}\text{ is realized as }\widehat{\mathsf{y}}}$$

How should this definition be extended from a typology of categorical grammars (HG) to a typology of probabilistic grammars (NHG, ME)? AM propose that the implication $(\mathsf{x},\mathsf{y}) \rightarrow (\widehat{\mathsf{x}},\widehat{\mathsf{y}})$ is a universal of the probabilistic ME typology $\mathfrak{T}_{\mathrm{ME}}(Gen, \boldsymbol{C})$ provided that the ME probability $G_{\boldsymbol{w}}(\mathsf{y}\,|\,\mathsf{x})$ of realizing the antecedent underlying form x as the antecedent surface form y is never larger than the ME probability $G_{\boldsymbol{w}}(\widehat{\mathsf{y}}\,|\,\widehat{\mathsf{x}})$ of realizing the consequent underlying form $\widehat{\mathsf{x}}$ as the consequent surface form $\widehat{\mathsf{y}}$, whereby the inequality (2) holds universally for any vector $\boldsymbol{w}$ of non-negative constraint weights. This definition extends unchanged to the NHG typology $\mathfrak{T}_{\mathrm{NHG}}(Gen, \boldsymbol{C})$. To illustrate, (/cost us/, [cos us]) → (/cost me/, [cos me]) is a ME or NHG universal provided the ME or NHG probability of t-deletion before a vowel is never larger than the probability of t-deletion before a consonant.

Both conditions (1) and (2) formalize the Greenbergian intuition that the consequent mapping $(\widehat{\mathsf{x}},\widehat{\mathsf{y}})$ is easier to get and thus better than the antecedent mapping (x,y). As a result, AM observe that probabilistic NHG and categorical HG share the same universals: $(\mathsf{x},\mathsf{y}) \rightarrow (\widehat{\mathsf{x}},\widehat{\mathsf{y}})$ is a universal of the probabilistic NHG typology $\mathfrak{T}_{\mathrm{NHG}}(Gen, \boldsymbol{C})$ in the sense of the universal probability inequality (2) if and only if it is also a universal of the categorical HG typology $\mathfrak{T}_{\mathrm{HG}}(Gen, \boldsymbol{C})$ in the sense of the universal material implication (1). AM show that the situation is different for ME. Indeed, if $(\mathsf{x},\mathsf{y}) \rightarrow (\widehat{\mathsf{x}},\widehat{\mathsf{y}})$ is a universal of the probabilistic ME typology $\mathfrak{T}_{\mathrm{ME}}(Gen, \boldsymbol{C})$, it is also a universal of the categorical HG typology $\mathfrak{T}_{\mathrm{HG}}(Gen, \boldsymbol{C})$—and therefore also of the probabilistic NHG typology $\mathfrak{T}_{\mathrm{NHG}}(Gen, \boldsymbol{C})$, because HG and NHG share the same universals. Yet, the reverse spectacularly fails. Indeed, AM consider a number of simple constraint sets drawn from the literature and observe that the corresponding ME typologies miss a score of universals that are reasonable and indeed predicted by the corresponding HG/NHG typologies.

3 The main formal result

This section presents the main result of the paper, namely a new necessary constraint condition on ME implicational universals. This result is introduced as an extension of some previous results obtained by AM.

3.1 *Previous results* Given an implication $(\mathsf{x},\mathsf{y}) \rightarrow (\widehat{\mathsf{x}},\widehat{\mathsf{y}})$, we call the surface form y the **antecedent winner** while we dub **antecedent loser** any other surface candidate z (if any) of the antecedent underlying form x different from the antecedent winner y. Analogously, we call the surface form $\widehat{\mathsf{y}}$ the **consequent winner** while we dub **consequent loser** any other surface candidate $\widehat{\mathsf{z}}$ (if any) of the consequent underlying form $\widehat{\mathsf{x}}$ different from the consequent winner $\widehat{\mathsf{y}}$. We denote by $Gen(\mathsf{x},\mathsf{y})$ and by $Gen(\widehat{\mathsf{x}},\widehat{\mathsf{y}})$ the sets of the antecedent and of the consequent losers, respectively. We now review some conditions on the constraint violations of winners and losers that are shown by AM to be necessary for the implication $(\mathsf{x},\mathsf{y}) \rightarrow (\widehat{\mathsf{x}},\widehat{\mathsf{y}})$ to qualify as a universal of the ME typology $\mathfrak{T}_{\mathrm{ME}}(Gen,\boldsymbol{C})$.

To start, AM show that the consequent winner $\widehat{\mathsf{y}}$ ought to be at least as good as the antecedent winner y in the sense that $\widehat{\mathsf{y}}$ cannot violate a constraint more than y. Equivalently, the inequality (2) holds for every constraint C in the constraint set $\boldsymbol{C}$.[2] In other words, the consequent winner of a ME implicational universal **harmonically bounds** the antecedent winner (to borrow a term from Prince & Smolensky 2004).

(2) $C(\mathsf{x},\mathsf{y}) \geq C(\widehat{\mathsf{x}},\widehat{\mathsf{y}})$

Next, AM show that the reverse holds for the losers: the antecedent losers ought to be at least as good as the consequent losers in the following sense. To put the winner comparison (2) to rest, AM focus on those constraints C that are **even** between the antecedent and consequent winners, namely assign them the same number of violations $C(\mathsf{x},\mathsf{y}) = C(\widehat{\mathsf{x}},\widehat{\mathsf{y}})$. They then use these even constraints to compare the antecedent and consequent losers. They show that every consequent loser $\widehat{\mathsf{z}}$ from the set $Gen(\widehat{\mathsf{x}},\widehat{\mathsf{y}})$ must come with some antecedent loser z from the set $Gen(\mathsf{x},\mathsf{y})$ such that z satisfies every even constraint satisfied by $\widehat{\mathsf{z}}$. Equivalently, the implication (3) holds for every constraint C in the constraint set $\boldsymbol{C}$. In conclusion, the antecedent losers of a ME implicational universal ought to be at least as good as the consequent losers when compared in terms of satisfaction of the constraints that are even between antecedent and consequent winners.

(3) If $C(\mathsf{x},\mathsf{y}) = C(\widehat{\mathsf{x}},\widehat{\mathsf{y}})$ and $C(\widehat{\mathsf{x}},\widehat{\mathsf{z}}) = 0$, then also $C(\mathsf{x},\mathsf{z}) = 0$.

These constraint conditions (2) and (3) make intuitive sense as follows. A ME implicational universal requires the ME probability of the consequent mapping to be always at least as large as the ME probability of the antecedent mapping. Intuitively, this requires *the consequent winner to be at least as good as the antecedent winner*. Condition (2) says that this winner-based intuition is correct. Yet, in a competitive world where candidates fight off for probability mass, a winner is only as good as its losers are bad. Thus, intuitively, a ME implicational universal also requires *the antecedent losers to be at least as good as the consequent losers*, so that the antecedent losers receive a larger share of the probability mass than the consequent losers, leaving the antecedent winner with a smaller share than the consequent winner, as desired. Condition (3) says that this loser-based intuition is correct as well, at least when we compare antecedent and consequent losers in terms of satisfaction of the constraints that are even between the antecedent and consequent winners.

3.2 *A new stronger result* The main formal result of this paper is boxed below. It strengthens AM's comparison of antecedent versus consequent losers of ME implicational universals by adding condition (4).

Suppose that $(\mathsf{x},\mathsf{y}) \rightarrow (\widehat{\mathsf{x}},\widehat{\mathsf{y}})$ is a universal of the ME typology $\mathfrak{T}_{ME}(Gen,\boldsymbol{C})$ corresponding to some sets Gen and $\boldsymbol{C}$ of candidates and constraints. Consider an arbitrary consequent loser $\widehat{\mathsf{z}}$ from the set $Gen(\widehat{\mathsf{x}},\widehat{\mathsf{y}})$ and an arbitrary constraint C_0 that is even between the antecedent and consequent mappings (x,y) and $(\widehat{\mathsf{x}},\widehat{\mathsf{y}})$ because it assigns them the same number $C_0(\mathsf{x},\mathsf{y}) = C_0(\widehat{\mathsf{x}},\widehat{\mathsf{y}})$ of violations. Some antecedent loser z from the set $Gen(\mathsf{x},\mathsf{y})$ must satisfy the implication (3) for every constraint C in the set $\boldsymbol{C}$ and furthermore z must violate the designated even constraint C_0 at most as much as $\widehat{\mathsf{z}}$, as stated in (4).

(4) $C_0(\mathsf{x},\mathsf{z}) \leq C_0(\widehat{\mathsf{x}},\widehat{\mathsf{z}})$

The additional condition (4) can be understood as follows. The comparison (2) between antecedent and consequent winners looks at their actual numbers of constraint violations. To illustrate, if the consequent winner $\widehat{\mathsf{y}}$ violates a constraint C ten times, this condition (2) requires the antecedent winner y to violate that

[2] AM's result actually says that the inequality (2) holds for every constraint C that is **reasonable** because each underlying form comes with some (winner or loser) candidate surface realization that satisfies that constraint. This reasonableness assumption is a mere technicality. In fact, a faithfulness constraint is usually reasonable because it is satisfied in particular by the faithful candidate. Furthermore, a markedness constraint is usually reasonable because the combinatorics of candidacy usually makes available some surface string that displays none of the structures penalized by that markedness constraint (possibly at the price of severe violations of other conflicting markedness constraints).

constraint at least ten times as well. The comparison (3) between antecedent and consequent losers instead only looks at whether they satisfy the even constraints, not at the actual numbers of violations. To illustrate, if the antecedent loser z violates an even constraint C ten times, this condition (3) only requires the consequent loser $\widehat{z}$ to violate that constraint at least once. Our strengthened result partially remedies this weakness of AM's loser comparisons. In fact, the additional comparison (4) between antecedent and consequent losers looks at their actual number of constraint violations, although it is limited to the designated even constraint C_0. To illustrate, if the antecedent loser z violates the designated even constraint C_0 ten times, this additional condition (4) does require the consequent loser $\widehat{y}$ to violate that constraint C_0 at least ten times as well.

In the rest of this paper, we argue that the improvement afforded by our strengthened result has substantial phonological implications. Indeed, we single out a class of implicational universals that hold in categorical HG and in probabilistic NHG by transitivity but fail in ME. We show that this failure cannot be diagnosed using AM's necessary constraint conditions (2) or (3) as both are satisfied in these cases. The failure can instead be diagnosed using the additional condition (4) that is incompatible with reasoning by transitivity.

4 The phonological implications of the formal result

This section introduces the notion of implicational universals that hold by transitivity with an example and uses our new strengthened result to diagnose their failure in ME.

4.1 *A puzzle* Vowel rounding is marked especially in front vowels (Kaun 1995). This generalization is captured by the markedness constraint *ROFRO against front rounded vowels, counterbalanced by the faithfulness constraint IDENTROUND. We further assume that velar place of articulation is marked relative to coronal place (de Lacy 2004, but see Ikwut-Ukwa 2026). This generalization is captured by the markedness constraint *DORSAL, counterbalanced by the faithfulness constraint IDENTPLACE. Thus, we expect the implicational universal antecedent → consequent to hold whenever the antecedent and consequent mappings only differ because the antecedent mapping has rounded front vowels and dorsal consonants while the consequent mapping has the corresponding unrounded vowels and coronal consonants.

To test this intuition, we focus on the implication (/kyg/, [kʰyg]) → (/tag/, [tʰag]) that compares two mappings that only differ for onset place and vowel rounding, while both mappings aspirate the voiceless onset and preserve the voiced coda. We complement the constraints for rounding and place mentioned above with three laryngeal constraints: ASPONS requires voiceless stops to be aspirated in syllable onsets; *VOICEDCODA penalizes voiced codas; finally, IDENTVOT is violated by corresponding segments with different VOTs (Flemming 2002) and thus penalizes discrepancies in both voicing and aspiration simultaneously.[3] Finally, we assume that *Gen* lists the two underlying forms /kyg/ and /tag/ with the 128 candidate surface realizations obtained by changing underlying stop voicing, aspiration, and place (between velar and coronal) as well as vowel rounding in all possible ways.

The resulting OT and HG typologies satisfy the implicational universal (/kyg/, [kʰyg]) → (/tag/, [tʰag]): every OT or HG grammar that contains the antecedent mapping, also contains the consequent mapping. This universal also holds in NHG, because categorical HG and probabilistic NHG always share the same universals. Finally, it is not hard to see that this universal holds in ME as well. No matter how we choose the non-negative constraint weights, every NHG or ME grammar assigns at least as much probability to the consequent mapping as to the antecedent mapping. This makes good intuitive sense given our constraints. In terms of faithfulness, the antecedent and consequent mappings are on an equal standing: both violate IDENTVOT once to satisfy ASPONS. In terms of markedness, the antecedent winner [kʰyg] is worse than the consequent winner [tʰag] because of the antecedent front rounded vowel penalized by *ROFRO and because of the antecedent velar penalized by *DORSAL. We conclude that categorical OT and HG as well as probabilistic NHG and ME capture the markedness of vowel rounding and velar place given the constraints considered.

Yet, ME's success is ephemeral. To see that, we add the constraint OCP that is violated when a syllable's onset and coda share both place and laryngeal specifications, as in the English word *gag* (Berkley 2000; McCarthy 1994; Frisch et al. 2004; Coetzee & Pater 2008). Since neither winners [kʰyg] nor [tʰag] violates OCP, it is not surprising that the implicational universal (/kyg/, [kʰyg]) → (/tag/, [tʰag]) is not compromised by adding OCP in categorical OT and HG. That is the case also in probabilistic

[3] Although such a broad faithfulness constraint is crucial for our specific example, our general point is independent of any assumptions about the phonetics/phonology interface, as shown by the additional examples presented in section 5.

NHG, because HG and NHG always share the same universals. The situation is dramatically different in ME. As soon as the apparently innocuous OCP is added to the constraint set, the implicational universal (/kyg/, [k^hyg]) → (/tag/, [t^hag]) fails in ME: it becomes possible to find non-negative weights that make the probability of the antecedent mapping larger than the probability of the consequent mapping, whereby ME violates the intuitive generalization about the markedness of rounded vowels and velar place. We now use the result boxed in subsection 3.2 to make sense of this puzzling observation.

4.2 *Universals that hold by transitivity in OT and HG* In order to focus on the formal properties of this example that are crucial, we only consider the constraints and the candidates described in ERC notation (Prince 2002) in (5). Furthermore, we represent each candidate as a sequence of two squares, with the convention that the first (second) square is filled if AspOns (*VoicedCoda, respectively) is violated. To illustrate, the antecedent winner candidate [k^hyg] can be represented as □■ because it satisfies AspOns (the first square is empty) but it violates *VoicedCoda (the second square is filled).

As shown in (5), if the antecedent winner [k^hyg] □■ is to beat the antecedent loser [kyg] ■■ in OT, AspOns must outrank Ident. This is because every loser-favoring constraint (marked as L) must be dominated by some winner-favoring constraint (marked as W). Furthermore, if the antecedent winner [k^hyg] □■ is to beat the antecedent loser [k^hyk] □□, Ident must in turn outrank *VoicedCoda. By transitivity, AspOns must outrank *VoicedCoda. This ranking AspOns $\gg$ *VoicedCoda ensures that the consequent winner [t^hag] □■ beats the consequent loser [tak] ■□. Without OCP, the ranking AspOns $\gg$ *VoicedCoda does not need to be inferred by transitivity: it is required for the antecedent winner [k^hyg] □■ to beat the antecedent loser [kyk] ■□. However, with OCP, the antecedent loser [kyk] provides no evidence for the ranking AspOns $\gg$ *VoicedCoda, which must therefore be inferred by transitivity. We conclude that, when OCP is added, the implication (/kyg/, [k^hyg]) → (/tag/, [t^hag]) is an OT universal that holds *by transitivity*.

(5)

☞[k^hyg] □■	AspOns	Ident	*VoiCo	OCP
[kyg] ■■	W	L		
[k^hyk] □□		W	L	
[kyk] ■□	W		L	W

☞[t^hag] □■	AspOns	Ident	*VoiCo	OCP
[tak] ■□	W		L	

Analogous considerations hold in HG. To see that, we replace the ERCs in (5) with the actual numbers of violations in (6). If the antecedent winner [k^hyg] □■ is to beat the antecedent loser [kyg] ■■, AspOns must have a larger weight than Ident. Furthermore, if the antecedent winner [k^hyg] □■ is to beat the antecedent loser [k^hyk] □□, Ident must in turn have a larger weight than *VoicedCoda. By transitivity, AspOns must have a larger weight than *VoicedCoda. This weight inequality $w_{\text{AspOns}} > w_{*\text{VoicedCoda}}$ ensures that the consequent winner [t^hag] □■ beats the consequent loser [tak] ■□. Without OCP, the weight inequality $w_{\text{AspOns}} > w_{*\text{VoicedCoda}}$ does not need to be inferred by transitivity because it is required for the antecedent winner [k^hyg] □■ to beat the antecedent loser [kyk] ■□. However, with OCP, the weight inequality $w_{\text{AspOns}} > w_{*\text{VoicedCoda}}$ must be inferred by transitivity. We conclude that, when OCP is added, the implication (/kyg/, [k^hyg]) → (/tag/, [t^hag]) is an HG universal that holds *by transitivity*.

(6)

/kyg/ ■■	AspOns	Ident	*VoiCo	OCP
☞[k^hyg] □■		1	1	
[kyg] ■■	1		1	
[k^hyk] □□		2		
[kyk] ■□	1	1		1

/tag/ ■■	AspOns	Ident	*VoiCo	OCP
☞[t^hag] □■		1	1	
[tak] ■□	1	1		

4.3 *Universals that hold by transitivity fail in ME* The four constraints in (5)-(6) are all even between the antecedent and consequent mappings (/kyg/, [k^hyg]) and (/tag/, [t^hag]): both mappings satisfy AspOns and OCP and both violate Ident and *VoicedCoda only once. These even constraints can thus be used to compare antecedent versus consequent losers. We focus on the consequent loser [tak] ■□. Can we find an antecedent loser that is as good as this consequent loser? The answer is positive when we compare antecedent and consequent losers in terms of AM's weaker condition (3). In fact, the consequent loser [tak] ■□ satisfies the two even constraints *VoicedCoda and OCP and the antecedent loser [k^hyk] □□ satisfies these two even constraints as well, as required by condition (3). Analogous considerations hold for all other consequent losers. We conclude that AM's original weaker result is unable to diagnose the puzzling failure of the ME universal (/kyg/, [kʰyg]) → (/tag/, [tʰag]) when OCP is added to the constraint set.

The situation is different when we instead compare antecedent versus consequent losers using our result in subsection 3.2, that is stronger because of the additional condition (4). To illustrate, we focus again on the consequent loser [tak] ■□ and we choose IDENT as the even constraint C_0 in the additional condition (4). The antecedent losers [kyg] ■■ and [kyk] ■□ violate the even constraints *VOICEDCODA and OCP and are therefore not as good as the consequent loser [tak] ■□ which instead satisfies both of them. The remaining antecedent loser [kʰyk] □□ does satisfy both of these even constraints *VOICEDCODA and OCP. Yet, it is not as good as the consequent loser [tak] ■□ because the antecedent loser violates IDENT more (twice) than the consequent loser (only once), thus flouting the additional condition (4) with C_0 = IDENT. We conclude that our new strengthened result manages to explain why the universal (/kyg/, [kʰyg]) → (/tag/, [tʰag]) fails in ME when OCP is added. Without OCP, the antecedent and consequent losers [kyk] ■□ and [tak] ■□ share the same violation profile and thus count as equally good.

5 Additional examples

In this section, we discuss other examples of implicational universals that hold in OT, HG, and NHG by transitivity but fail in ME, to show that these examples are relatively easy to construct.

5.1 *Vowel nasality and obstruent voicing* Nasal vowels are marked relative to corresponding oral vowels (Ferguson 1963; Greenberg 1966). This generalization is captured by the markedness constraint *NASAL, counterbalanced by the faithfulness constraint IDENTNASAL. Thus, we expect the implicational universal antecedent → consequent to hold whenever the antecedent and consequent mappings compared only differ because the antecedent mapping has a nasal vowel while the consequent mapping has the corresponding oral vowel. To test this intuition, we focus on the implication (/æ̃b+z/, [æ̃bs]) → (/æb+z/, [æbs]) that compares two mappings that only differ for vowel nasality, while both mappings preserve obstruent voicing in roots but not in suffixes. We complement the constraints for nasality mentioned above with three constraints for voicing: *VOICE, IDENTROOT, and IDENTSUFFIX.[4] Furthermore, we assume that *Gen* lists the two underlying forms /æ̃b+z/ and /æb+z/ with the eight candidates obtained by changing vowel nasality and obstruent voicing in all possible ways. The resulting categorical OT and HG typologies as well as the probabilistic NHG and ME typologies all validate the implicational universal (/æ̃b+z/, [æ̃bs]) → (/æb+z/, [æbs]) and thus capture the markedness of nasality.

Yet again, ME's success is ephemeral. To see that, we add the constraint SSP-PLACE that prohibits rising sonority in codas if the segments share place, a kind of partial geminate behavior (Rose & Walker 2004; Coetzee & Pater 2008:491). For example, the coda sequence [tz] violates SSP-PLACE but [pz] does not. Since neither winners [æ̃bs] nor [æbs] violate SSP-PLACE (the coda segments do not share place), it is not surprising that the implicational universal (/æ̃b+z/, [æ̃bs]) → (/æb+z/, [æbs]) is not compromised by adding SSP-PLACE in categorical OT and HG. That is the case also in probabilistic NHG, because HG and NHG always share the same universals. Again, the situation is dramatically different in ME. As soon as SSP-PLACE is added, the implicational universal (/æ̃b+z/, [æ̃bs]) → (/æb+z/, [æbs]) fails in ME, which therefore flouts the basic generalization about the markedness of nasals recalled at the beginning. The culprit is again that reasoning by transitivity fails in ME.

To see that, we focus on the constraints and the candidates described in (7) using the ERC notation. If the antecedent winner [æ̃bs] □■ is to beat the antecedent loser [æ̃ps] ■■ in OT, IDENTROOT must outrank *VOICE. Furthermore, if the antecedent winner [æ̃bs] is to beat the antecedent loser [æ̃bz] □□, *VOICE must in turn outrank IDENTSUFFIX. By transitivity, IDENTROOT must outrank IDENTSUFFIX. This ranking IDENTROOT ≫ IDENTSUFFIX ensures that the consequent winner [æbs] □■ beats the consequent loser [æpz] ■□. Without SSP-PLACE, this ranking IDROOT ≫ IDSUFFIX does not need to be inferred by transitivity

4 Aaron Kaplan (p.c.) observes that our example would not work if we were to pair IDENTROOT with the general constraint IDENT instead of the restricted constraint IDENTSUFFIX. While that is true, our example could be rescued by replacing IDENTSUFFIX with any restricted faithfulness constraint that protects voicing on the suffix /-z/. Alternatives include IDENTMINIMAL that protects voicing only on a segment which is the only exponent of a morpheme (in the spirit of MAXMINIMAL from Casali 1998) or IDENTINITIAL that protects voicing on the initial segment of a morpheme (as proposed by Beckman 1999).

because it is required in order for the antecedent winner [æ̃bs] □■ to beat the antecedent loser [æ̃pz] ■□. With SSP-PLACE, the latter antecedent loser provides no evidence for IDROOT ≫ IDSUFFIX which must thus be inferred by transitivity. We conclude that, when SSP-PLACE is added, the implication (/æ̃b+z/, [æ̃bs]) → (/æb+z/, [æbs]) is an OT universal that holds *by transitivity*. This universal also holds in HG by transitivity, as can be seen by reasoning on the actual numbers of violations in (8) exactly as we did in (6) above.

(7)

☞ [æ̃bs] □■	IDRT	*VCE	IDSFX	SSP-PLACE
[æ̃ps] ■■	W	L		
[æ̃bz] □□		W	L	
[æ̃pz] ■□	W		L	W

☞ [æbs] □■	IDRT	*VCE	IDSFX	SSP-PLACE
[æpz] ■□	W		L	

Once SSP-PLACE is added, the universal fails in ME as follows. All four constraints in (7) and (8) are even between the antecedent and consequent mappings (/æ̃b+z/, [æ̃bs]) and (/æb+z/, [æbs]). These even constraints are thus relevant for establishing whether the antecedent losers are as good as the consequent losers. We focus on the consequent loser [æpz] ■□ and we choose *VOICE as the even constraint C_0 in the additional condition (4). The antecedent losers [æ̃ps] ■■ and [æ̃pz] ■□ violate the even constraints IDENTSUFFIX and SSP-PLACE and are therefore not as good as the consequent loser [æpz] ■□ which instead satisfies both of them. The remaining antecedent loser [æ̃bz] □□ does satisfy both of these even constraints IDENTSUFFIX and SSP-PLACE. Yet, it is not as good as the consequent loser [æpz] ■□ because the antecedent loser violates *VOICE more (twice) than the consequent loser (only once), thus flouting the additional condition (4) with C_0 = *VOICE. Our new strengthened result thus explains why the universal (/æ̃b+z/, [æ̃bs]) → (/æb+z/, [æbs]) fails in ME once SSP-PLACE is added. Without SSP-PLACE, the antecedent and consequent losers [æ̃pz] and [æpz] share the same constraint violation profiles and thus count as equally good. As highlighted by the square notation, this example with SSP-PLACE has exactly the same formal structure as the preceding example with OCP.

(8)

/æ̃b+z/	IDRT	*VCE	IDSFX	SSP-PLACE
☞ [æ̃bs] □■		1	1	
[æ̃ps] ■■	1		1	
[æ̃bz] □□		2		
[æ̃pz] ■□	1	1		1

/æb+z/	IDRT	*VCE	IDSFX	SSP-PLACE
☞ [æbs] □■		1	1	
[æpz] ■□	1	1		

5.2 *Nasal fricatives and vowel rounding* Nasal fricatives are particularly marked: they are rare, only result from nasal spreading, and are never contrastive (Ladefoged & Maddieson 1996:§4.4; Shosted 2006). This generalization is captured by the markedness constraint *NASFRIC, counterbalanced by the faithfulness constraints IDENTNASAL. Thus, we expect the implicational universal antecedent → consequent to hold whenever the antecedent and consequent mappings compared only differ because the antecedent mapping has a nasal fricative while the consequent mapping has the corresponding oral fricative. To test this intuition, we focus on the implication (/ys̃+ɒ/, [is̃+ɒ]) → (/ys#ɒ/, [is#ɒ]) that compares two mappings that only differ for nasality of the fricative (and for internal boundaries), while both mappings pattern alike for vowel rounding, namely preserve it for the low back vowel /ɒ/ but not for the front high vowel /y/. We complement the constraints mentioned above with three constraints for vowel rounding: *ROFRO against front rounded vowels; *ROLO against rounded low vowels; and IDENTROUND. Furthermore, we assume that *Gen* lists the two underlying forms /ys̃+ɒ/ and /ys#ɒ/ with the eight candidates obtained by changing vowel rounding and obstruent nasality in all possible ways. The resulting categorical OT and HG typologies as well as the probabilistic NHG and ME typologies all validate the implicational universal (/ys̃+ɒ/, [is̃+ɒ]) → (/ys#ɒ/, [is#ɒ]) and thus capture the markedness of nasal fricatives.

Yet, once again, ME's success is ephemeral. To see that, we add the constraint AGREE against rounding disharmony subject to two qualifications. First, AGREE only enforces progressive rounding harmony, not regressive harmony. In other words, it penalizes a rounded vowel on the left followed by an unrounded vowel on the right, but not a rounded vowel on the right preceded by an unrounded vowel on the left, capturing a well known directional asymmetry in harmony systems (Rose & Walker 2014:section 4.1). Second, AGREE only penalizes rounding disharmony between two vowels that belong to (possibly different morphemes of) the same word but not disharmony between two vowels separated by a word boundary #, capturing the well known sensitivity of harmony processes to the hierarchy of morphological boundaries (McPherson & Hayes 2016). Since neither winners [is̃+ɒ] nor [is#ɒ] violate AGREE (because it does not penalize regressive disharmony), it is not surprising that the implicational universal (/ys̃+ɒ/, [is̃+ɒ]) → (/ys#ɒ/, [is#ɒ]) is not compromised by adding AGREE in categorical OT and HG. That is the case also in probabilistic NHG, because HG and NHG share the same universals. Once again, the situation is dramatically different in ME. As soon as AGREE is added, the implicational universal (/ys̃+ɒ/, [is̃+ɒ]) → (/ys#ɒ/, [is#ɒ]) fails in ME. ME thus flouts the basic generalization about the markedness of nasal fricatives recalled at the beginning. Once again, the culprit is that reasoning by transitivity fails in ME.

To see that, we focus on the constraints and the candidates described in (9) using the ERC notation. We reason in the by now familiar way. If the antecedent winner [is̃+ɒ] □■ is to beat the antecedent loser [ys̃+ɒ] ■■ in OT, *ROFRO must outrank IDENT. Furthermore, if the antecedent winner [is̃+ɒ] □■ is to beat the antecedent loser [is̃+ɑ] □□, IDENT must in turn outrank *ROLO. By transitivity, *ROFRO must outrank *ROLO. This ranking *ROFRO ≫ *ROLO ensures that the consequent winner [is#ɒ] □■ beats the consequent loser [ys#ɑ] ■□. Without AGREE, this ranking *ROFRO ≫ *ROLO does not need to be inferred by transitivity because it is required in order for the antecedent winner [is̃+ɒ] □■ to beat the antecedent loser [ys#ɑ] ■□. With AGREE, the latter antecedent loser provides no evidence for the ranking *ROFRO ≫ *ROLO, which must thus be inferred by transitivity. We conclude that, when AGREE is added, the implication (/ys̃+ɒ/, [is̃+ɒ]) → (/ys#ɒ/, [is#ɒ]) is an OT universal that holds *by transitivity*.

(9)

☞ [is̃+ɒ] □■	*ROFRO	IDENT	*ROLO	AGREE
[ys̃+ɒ] ■■	W	L		
[is̃+ɑ] □□		W	L	
[ys̃+ɑ] ■□	W		L	W

☞ [is#ɒ] □■	*ROFRO	IDENT	*ROLO	AGREE
[ys#ɑ] ■□	W		L	

This universal also holds in HG by transitivity, as can be seen by reasoning on the actual numbers of constraint violations in (10) exactly as we did in (6) and (8) above. Once AGREE is added, the universal fails in ME because we can reason as follows. All four constraints in (9) and (10) are even between the antecedent and consequent mappings (/ys̃+ɒ/, [is̃+ɒ]) and (/ys#ɒ/, [is#ɒ]). These even constraints are thus relevant to establish whether the antecedent losers are as good as the consequent losers. We focus on the consequent loser [ys#ɑ] ■□ and we choose IDENT as the even constraint C_0 in the additional condition (4). The antecedent losers [ys̃+ɒ] ■■ and [ys̃+ɑ] ■□ violate the even constraints *ROLO and AGREE and are therefore not as good as the consequent loser [ys#ɑ] ■□ which instead satisfies both of them. The remaining antecedent loser [is̃+ɑ] □□ does satisfy both of these even constraints *ROLO and AGREE. Yet, it is not as good as the consequent loser [ys#ɑ] ■□ because the antecedent loser violates IDENT more (twice) than the consequent loser (only once), thus flouting the additional condition (4) with C_0 = IDENT. Our new strengthened result thus explains why the universal (/ys̃+ɒ/, [is̃+ɒ]) → (/ys#ɒ/, [is#ɒ]) fails in ME once AGREE is added. Without AGREE, the antecedent and consequent losers [ys̃+ɑ] ■□ and [ys#ɑ] ■□ share the same constraint violation profiles and thus count as equally good.

(10)

/ys̃+ɒ/ ■■	*ROFRO	IDENT	*ROLO	AGREE
☞ [is̃+ɒ] □■		1	1	
[ys̃+ɒ] ■■	1		1	
[is̃+ɑ] □□		2		
[ys̃+ɑ] ■□	1	1		1

/ys#ɒ/ ■■	*ROFRO	IDENT	*ROLO	AGREE
☞ [is#ɒ] □■		1	1	
[ys#ɑ] ■□	1	1		

6 Conclusions

The logic of the argument developed in this paper has been to look at basic phonological generalizations and combine them with unrelated phonological processes. Section 4 looked at the basic generalization that rounded vowels are marked compared to corresponding unrounded vowels and combined this generalization with unrelated laryngeal processes (voiceless onsets are aspirated and voiced codas faithfully realized). Subsection 5.1 looked at the basic generalization that nasal vowels are marked compared to corresponding oral vowels and combined this generalization with unrelated voicing processes (obstruent voicing is preserved in roots but not in suffixes). Finally, subsection 5.2 looked at the basic generalization that nasal fricatives are marked compared to corresponding oral fricatives and combined this generalization with unrelated rounding processes (rounding is preserved for low vowels but not for front vowels).

The basic generalizations considered hold as implicational universals both in categorical OT and HG and in probabilistic NHG. These implicational universals are resilient in the sense that they are not affected by adding constraints that model unrelated phonological processes. This is not so in ME where these implicational universals are instead brittle: an additional constraint entirely irrelevant to the generalization breaks the universal. Indeed, section 4 has shown that the basic generalization that rounded vowels are marked is broken in ME by an unrelated constraint OCP against two consonants that share place and voicing. Subsection 5.1 has shown that the basic generalization that nasal vowels are marked is broken in ME by an unrelated constraint SSS-PLACE against codas with rising sonority. Finally, subsection 5.2 has shown that the basic generalization that nasal fricatives are marked is broken in ME by an unrelated constraint AGREE against two vowels that differ in rounding.

These failures of ME can all be traced back to a common cause. The unrelated constraints have the surprising effect that the relevant implicational universals hold in OT, HG and NHG only by reasoning by transitivity. Reasoning by transitivity fails in ME because of the idiosyncratic way in which ME compares the antecedent versus consequent losers of implicational universals, as distilled in the result boxed in subsection 3.2. This failure at reasoning by transitivity suggests that ME is not a satisfactory model of constraint interaction for probabilistic phonology.

Appendix

Let us suppose that an implication $(\mathsf{x},\mathsf{y}) \to (\widehat{\mathsf{x}},\widehat{\mathsf{y}})$ is a universal of the ME typology $\mathfrak{T}_{\mathrm{ME}}(Gen,\boldsymbol{C})$ corresponding to some sets Gen and $\boldsymbol{C}$ of candidates and constraints. AM then show that the following necessary condition holds for every consequent loser $\widehat{\mathsf{z}}$ from the set $Gen(\widehat{\mathsf{x}},\widehat{\mathsf{y}})$: each antecedent loser z from the set $Gen(\mathsf{x},\mathsf{y})$ can be paired with some non-negative coefficient λ_{z} such that these coefficients add up to one and furthermore satisfy the constraint inequality (11).[5] The lefthand side of this inequality features the difference between the constraint violation vector of the consequent winner $\widehat{\mathsf{y}}$ minus the constraint violation vector of the antecedent winner y. The righthand side features the difference between the constraint violation vector of the consequent loser $\widehat{\mathsf{z}}$ considered minus the sum of the constraint violation vectors of the antecedent losers z, each rescaled by the corresponding coefficient λ_{z}.

(11) $$\boldsymbol{C}(\widehat{\mathsf{x}},\widehat{\mathsf{y}}) - \boldsymbol{C}(\mathsf{x},\mathsf{y}) \le \boldsymbol{C}(\widehat{\mathsf{x}},\widehat{\mathsf{z}}) - \sum_{\mathsf{z}\in Gen(\mathsf{x},\mathsf{y})} \lambda_{\mathsf{z}}\boldsymbol{C}(\mathsf{x},\mathsf{z}) \qquad \lambda_{\mathsf{z}} \ge 0 \text{ and } \sum_{\mathsf{z}\in Gen(\mathsf{x},\mathsf{y})} \lambda_{\mathsf{z}} = 1$$

Let S be the subset of $Gen(\mathsf{x},\mathsf{y})$ consisting of those antecedent losers z whose corresponding coefficient λ_{z} is not just non-negative but actually strictly positive. This set S is non-empty because the coefficients λ_{z} add up to one and therefore cannot be all equal to zero. The sum on the righthand side of the inequality (11) therefore effectively runs over this subset S, as the antecedent losers z from $Gen(\mathsf{x},\mathsf{y})$ that do not belong to S have a corresponding coefficient λ_{z} equal to zero and therefore contribute nothing to the sum. Furthermore, let $\widehat{\boldsymbol{C}}$ be the subset of the constraint set $\boldsymbol{C}$ consisting of those constraints that are even between the antecedent and consequent mappings (x,y) and $(\widehat{\mathsf{x}},\widehat{\mathsf{y}})$ and furthermore are satisfied by the consequent loser $\widehat{\mathsf{z}}$, namely $\widehat{\boldsymbol{C}}(\mathsf{x},\mathsf{y}) = \widehat{\boldsymbol{C}}(\widehat{\mathsf{x}},\widehat{\mathsf{y}})$ and $\widehat{\boldsymbol{C}}(\widehat{\mathsf{x}},\widehat{\mathsf{z}}) = \mathbf{0}$.

The inequality (11) restricted from the whole constraint set $\boldsymbol{C}$ to this constraint subset $\widehat{\boldsymbol{C}}$ simplifies to $\sum_{\mathsf{z}\in S} \lambda_{\mathsf{z}}\widehat{\boldsymbol{C}}(\mathsf{x},\mathsf{z}) \le \mathbf{0}$. The latter inequality says in turn that every antecedent loser z that has a strictly positive coefficient λ_{z} and therefore belongs to the set S must satisfy the constraints in $\widehat{\boldsymbol{C}}$, as stated by the set inclusion in (12). In other words, every antecedent loser z in the non-empty set S satisfies every even constraint that is satisfied by the consequent loser $\widehat{\mathsf{z}}$, thus establishing that AM's implication (3) holds for every constraint C in the constraint set $\boldsymbol{C}$.

(12) $$S \subseteq \left\{\mathsf{z} \in Gen(\mathsf{x},\mathsf{y}) \,\middle|\, \widehat{\boldsymbol{C}}(\mathsf{x},\mathsf{z}) = \mathbf{0}\right\}$$

Next, we consider some constraint C_0 that is even between the antecedent and consequent mappings (x,y) and $(\widehat{\mathsf{x}},\widehat{\mathsf{y}})$, namely $C_0(\mathsf{x},\mathsf{y}) = C_0(\widehat{\mathsf{x}},\widehat{\mathsf{y}})$. Yet, contrary to the even constraints collected into the subset

[5] As usual, all vector operations are defined component-wisely and all inequalities apply component-wisely.

$\widehat{\boldsymbol{C}}$, we allow for this constraint C_0 to be possibly violated by the consequent loser $\widehat{\mathsf{z}}$. The inequality (11) for this designated even constraint C_0 becomes (13a). The inequality (13b) holds by replacing the generic term $C_0(\mathsf{x},\mathsf{z})$ in the sum with the smallest term in the sum. The identity (13c) holds because the coefficients λ_{z} add up to one. Finally, the inequality (13d) holds because of the set inclusion (12) together with the fact that the minimum over a set can only decrease when the set is enlarged.

(13)
$$
\begin{aligned}
C_0(\widehat{\mathsf{x}},\widehat{\mathsf{z}}) &\overset{(a)}{\geq} \sum_{\mathsf{z}\in S} \lambda_{\mathsf{z}} C_0(\mathsf{x},\mathsf{z}) \\
&\overset{(b)}{\geq} \min\left\{C_0(\mathsf{x},\mathsf{z}) \,\middle|\, \mathsf{z}\in S\right\} \sum_{\mathsf{z}\in S} \lambda_{\mathsf{z}} \\
&\overset{(c)}{=} \min\left\{C_0(\mathsf{x},\mathsf{z}) \,\middle|\, \mathsf{z}\in S\right\} \\
&\overset{(d)}{\geq} \min\left\{C_0(\mathsf{x},\mathsf{z}) \,\middle|\, \mathsf{z}\in Gen(\mathsf{x},\mathsf{y}), \widehat{\boldsymbol{C}}(\mathsf{x},\mathsf{z}) = \mathbf{0}\right\}
\end{aligned}
$$

The antecedent loser z that realizes the minimum in (13d) satisfies every even constraint in $\widehat{\boldsymbol{C}}$ satisfied by the consequent loser $\widehat{\mathsf{z}}$. Furthermore, it violates the designated even constraint C_0 at most as much as the consequent loser $\widehat{\mathsf{z}}$, as required by the additional condition (4).

References

Anttila, Arto (2018). Variation in phonology. *Oxford Research Encyclopedia of Linguistics.*, Oxford University Press, Oxford, 76–91.

Anttila, Arto & Giorgio Magri (2018). Does MaxEnt overgenerate? Implicational universals in Maximum Entropy grammar. Gallagher, Gillian, Maria Gouskova & Yin Sora (eds.), *AMP 2017: Proceedings of the 2017 Annual Meeting on Phonology*, Linguistic Society of America, Washington, DC.

Anttila, Arto, Scott Borgeson & Giorgio Magri (2019). Equiprobable mappings in weighted constraint grammars. Nicolai, Garrett & Ryan Cotterell (eds.), *Proceedings of the 16th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology*, Association for Computational Linguistics, 125–134.

Beckman, Mary E. (1999). *Positional faithfulness: An Optimality Theoretic treatment of phonological asymmetries*. Garland Publishing (Outstanding dissertations in Linguistics), New York.

Berkley, Deborah (2000). *Gradient OCP Effects*. Ph.D. thesis, Northwestern University.

Boersma, Paul & Joe Pater (2016). Convergence properties of a gradual learning algorithm for Harmonic Grammar. McCarthy, John & Joe Pater (eds.), *Harmonic Grammar and Harmonic Serialism*, Advances in Optimality Theory, Equinox Publishing, Sheffield, UK and Bristol, CT.

Casali, Roderic F. (1998). *Resolving Hiatus*. Outstanding dissertations in Linguistics, Garland, New York.

Coetzee, Andries W. & Shigeto Kawahara (2013). Frequency biases in phonological variation. *Natural Language and Linguistic Theory* 31:1, 47–89.

Coetzee, Andries W. & Joe Pater (2008). Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic. *Natural Language and Linguistic Theory* 26:2, 289–337.

Coetzee, Andries W. & Joe Pater (2011). The place of variation in phonological theory. Goldsmith, John, Jason Riggle & Alan Yu (eds.), *Handbook of phonological theory*, Blackwell, Cambridge, 401–434.

Ferguson, Charles A. (1963). Assumptions about nasals: a sample study in phonological universals. Greenberg, J.H. (ed.), *Universals of language*, MIT Press, Cambridge, 53–60.

Flemming, Edward S (2002). *Auditory representations in phonology*. Routledge.

Frisch, Stefan A, Janet B Pierrehumbert & Michael B Broe (2004). Similarity avoidance and the OCP. *Natural Language & Linguistic Theory* 22:1, 179–228.

Greenberg, Joseph H. (1966). *Language Universals: With Special Reference to Feature Hierarchies*. Mouton de Gruyter, Berlin and New York.

Guy, Gregory R. (1980). Variation in the group and the individual: The case of final stop deletion. Labov, William (ed.), *Locating language in time and space*, Academic Press, New York, 1–36.

Hayes, Bruce & Aaron Kaplan (2023). Zero-weighted constraints in Noisy Harmonic Grammar. *Linguistic Inquiry* 1–14.

Hayes, Bruce & Colin Wilson (2008). A Maximum Entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39, 379–440.

Ikwut-Ukwa, Emiyare (2026). Dorsals are the least-marked codas in Obolo. Talk given at the 2026 LSA Annual Meeting, January 9, 2026, New Orleans.

Ikwut-Ukwa, Emiyare, Kushal Thaman, Annalisa Welinder, Arto Anttila & Giorgio Magri (2025). Probabilistic syllable structure. *WCCFL 42: Proceedings of the 42nd meeting of the West Coast Conference in Linguistics*.

Kaun, Abigail Rhoades (1995). *The typology of Rounding Harmony: An Optimality Theoretic Approach*. Ph.D. thesis, UCLA.

de Lacy, Paul (2004). Markedness conflation in Optimality Theory. *Phonology* 21, 145–199.

Ladefoged, Peter & Ian Maddieson (1996). *The Sounds of the World's Languages*. Wiley-Blackwell.

Magri, Giorgio & Arto Anttila (2023). Paradoxes of MaxEnt markedness. Elkins, Noah, Bruce Hayes, Jinyoung Jo & Jian-Leat Siah (eds.), *AMP 2022: Supplemental Proceedings of the 2022 Annual Meeting on Phonology*, Linguistic Society of America.

Magri, Giorgio & Arto Anttila (2025). Probabilistic phonology is intrinsically categorical. Submitted.

Magri, Giorgio & Arto Anttila (2026). The stochastic generalization and its implications for probabilistic constraint-based phonology. Monograph in preparation.

McCarthy, John J. (1994). The phonetics and phonology of Semitic pharyngeals. Keating, P. (ed.), *Papers in laboratory phonology III: Phonological structure an phonetic form*, Cambridge University Press, London.

McPherson, Laura & Bruce Hayes (2016). Relating application frequency to morphological structure: the case of Tommo So vowel harmony. *Phonology* 33, 125–167.

Pierrehumbert, Janet B. (2022). More than seventy years of probabilistic phonology. Dresher, B. Elan & Harry van der Hulst (eds.), *The Oxford history of phonology*, Oxford University Press, Oxford, 639–655.

Potts, Christopher, Joe Pater, Karen Jesney, Rajesh Bhatt & Michael Becker (2010). Harmonic Grammar with linear programming: from linear systems to linguistic typology. *Phonology* 27(1), 1–41.

Prince, Alan (2002). Entailed ranking arguments, URL `http://roa.rutgers.edu/files/500-0202/500-0202-PRINCE-0-1.PDF`. Manuscript (Rutgers University). Available from the Rutgers Optimality Archive as ROA 500.

Prince, Alan & Paul Smolensky (2004). *Optimality Theory: constraint interaction in generative grammar*. Blackwell, Oxford.

Rose, Sharon & Rachel Walker (2004). A typology of consonant agreement as correspondence. *Language* 80:3, 475–531.

Rose, Sharon & Rachel Walker (2014). Harmony systems. Goldsmith, John A., Jason Riggle & Alan C. L. Yu (eds.), *The handbook of phonological theory (second edition)*, Wiley Blackwell, 240–290.

Shosted, Ryan Keith (2006). *The aeroacoustics of nasalized fricatives*. Ph.D. thesis, University of California, Berkeley, URL `https://escholarship.org/uc/item/00h9g9gg`.

Wilson, Colin (2025). Maximum entropy grammars. Jardine, Adam & Paul de Lacy (eds.), *The Cambridge Handbook of Phonology (2nd edition)*, Cambridge University Press.