A Theory of When and How Learners Construct Tiers: Implications for Opaque and Transparent Vowels

Caleb Belth
University of Utah
`caleb.belth@utah.edu`

Abstract

Some vowel harmony systems have neutral vowels, which need not agree along the harmonizing dimensions of vowel quality. Neutral vowels differ in whether other vowels in turn harmonize with them: those that are harmonized with are *opaque* while those that are not are *transparent*. Prior artificial language learning studies have found opaque vowels to be more readily learned in laboratory settings than transparent vowels. This was initially thought to be because transparent vowels intervene between harmonizing vowels on a vowel tier, making harmony non-local. However, subsequent computational work has demonstrated that vowel harmony is typically tier-strictly-local, even with transparent and opaque vowels, indicating that there may be less difference between them than once believed. I propose an explanation for the different learning results between transparent and opaque vowels by making use of a recent learning model that proposes learners create tier-like representations in response to being unable to sufficiently generalize without them, as measured by the Tolerance Principle. I demonstrate how the representations that this model constructs make sense of different learning results between transparent and opaque vowels, despite their shared formal properties.

1 Introduction and Background

Vowel harmony involves non-local dependencies, as vowels agree along the harmonizing dimensions across intervening consonants. In the following example (1) from Turkish, the underlined suffix vowels harmonize in backness with the vowel to their left (Nevins 2010, p. 28; Kabak 2011, p. 3).

(1) [dɑ̲l-lɑ̲r-ɯ̲n] branch-PL-GEN
[je̲r-le̲r-i̲n] place-PL-GEN
[i̲p-le̲r-i̲n] rope-PL-GEN

In some vowel harmony systems, a subset of vowels are not required to harmonize—they are *neutral*. These neutral vowels are coarsely grouped into two categories: *opaque* and *transparent*. Opaque vowels participate in harmony in that other vowels harmonize with them. For example, in addition to backness harmony, Turkish high vowels [i, y, ɯ, u] also harmonize in roundness (2a). Low vowels [e, ø, ɑ, o] are neutral to the rounding harmony (2b), but high vowels nevertheless harmonize with them *opaquely* (2c).

(2) a. [i̲p-i̲n] rope-GEN
[jy̲z-y̲n] face-GEN
[kɯ̲z-ɯ̲n] girl-GEN
[bu̲z-u̲n] ice-GEN
b. [kɯz-lɑr] gril-PL
[buz-lɑr] ice-PL
c. [e̲l-i̲n] hand-GEN
[sø̲z-y̲n] word-GEN
[sɑ̲p-ɯ̲n] stalk-GEN
[jo̲l-u̲n] road-GEN

Transparent vowels, on the other hand, are inert, neither harmonizing nor being harmonized with. For instance, while Hungarian has backness harmony (3a), the vowels [iː, eː] are transparent, with the DAT vowel skipping them to harmonize with the next vowel to the left of (3b; examples from Benus and Gafos 2007).

(3) a. [ørø̲m-nɛ̲k] joy-DAT
[moːku̲ʃ-nɔ̲k] squirrel-DAT
b. [ɛ̲miːr-nɛ̲k] emir-DAT
[pɔ̲piːr-nɔ̲k] paper-DAT
[my̲ːveːs-nɛ̲k] artist-DAT
[ka̲ːveː-nɔ̲k] coffee-DAT

The development of autosegmental theory (Goldsmith, 1976) allowed for treating vowel harmony as local on a vowel tier (Clements, 1976, 1980). Opaque vowels do not harmonize with the preceding vowel on a vowel tier, but their features take over the harmony, so all remains local. However, vowel harmony must cross transparent vowels, which introduces non-locality even on a vowel tier (Goldsmith, 1985; Bakovic and Wilson, 2000; Hayes and Londe, 2006; Finley, 2009). Finley (2015) hypothesized that this makes transparent vowels harder to learn than opaque vowels and tested this hypothesis with a series of artificial grammar learning (AGL) experiments. Finley found that adults indeed succeeded at learning the behavior of an opaque vowel but failed to learn the behavior of a transparent vowel under equivalent conditions. Only by increasing the amount of evidence of the neutral vowel's transparency, by increasing the amount of exposure to items that unambiguously indicated transparency, did learners eventually succeed at learning transparent vowel behavior. Chen (2024) found compatible results: when adults were trained on an artificial harmony system with a neutral and a transparent vowel, they either failed to learn the harmony system altogether or appeared to treat both the opaque and transparent vowels as opaque (depending on the presentation of the training stimuli).

However, work in computational phonology has found that from a formal-language-theoretic perspective, neither opaque nor transparent vowels meaningfully change the computational character of vowel harmony: vowel harmony is typically tier-strictly-local ($k = 2$) (Heinz et al., 2011), with or without opaque and/or transparent vowels (Burness et al., 2021). Learners could project a tier that excludes transparent vowels along with the consonants, and this renders all relevant dependencies local on the tier. Moreover, as Finley (2015) observed, transparent vowels must be learnable, since they appear in numerous natural language harmony systems. Indeed tier-strictly-local constraints and processes are provably efficiently learnable (Jardine and Heinz, 2016; Jardine and McMullin, 2017; Burness and McMullin, 2019) and Finley (2015) did find that under the right conditions, transparent vowel harmony can be learned in the lab. Similarly, Ozburn et al. (2016) found that adult Canadian French speakers succeeded at learning the behavior of a transparent vowel in an artificial vowel harmony system built around the French vowel inventory.

Given that vowel harmony with opaque and transparent vowels shares a fundamental underlying computational structure and both must be learnable in natural languages, it is worth revisiting what might underlie the picture from experimental results that vowel harmony is harder to learn with transparent vowels than opaque vowels.

To do so, I build on my prior work (Belth, 2024), where I proposed that humans learn phonological alternations by tracking dependencies between alternating segments and the segments adjacent to them—using the well-attested ability to track adjacent dependencies over many kinds of representations (Saffran et al., 1996, 1997; Aslin et al., 1998; Saffran et al., 1999; Fiser and Aslin, 2002). In that proposal, if adjacent dependencies are not sufficiently predictive of the alternation, where *sufficiency* is measured by the Tolerance/Sufficiency Principle (Yang, 2016), learners use the same sensitivity to adjacent dependencies to form a new representation that excludes any adjacent segments that led to incorrect predictions. The resulting representations can be interpreted as tiers, which are constructed in dynamic response to the input. In Belth (2024), I implemented this proposal as a learning model. The model succeeded at learning natural language harmony processes, including Turkish vowel harmony, in which low vowels are opaque to rounding harmony, and Finnish vowel harmony, in which, similarly to Hungarian, [i, e] are transparent to backness harmony (Ringen and Heinämäki, 1999). In Turkish, the learner constructed a vowel tier and in Finnish it constructed a tier that excluded the transparent vowels. Thus, the proposal already accounts for the learnability of vowel harmony with opaque and transparent vowels in natural languages. In this paper, I will demonstrate that it simultaneously accounts for the difference in experimental settings between artificial vowel harmony systems with opaque vs. transparent vowels.

Consider a transparent vowel harmony system, such as the artificial one from Finley (2015), where a suffix vowel harmonizes in backness with the final vowel of the stem (4a), but where the vowel [ɛ] is neutral (4b)-(4c). Since the neutral vowel is itself front, only when the penultimate stem vowel is back (4b) do we get unambiguous evidence that [ɛ] is transparent.

(4) a. [budok-o]
[degib-e]
b. [dotɛb-o]
c. [tedɛt-e]

It is thus possible that the learner treats the neutral vowel as *opaque* and handles the cases like (4b), which contradict this, as lexicalized exceptions. If, during learning, enough of these exceptions accumulate that the learner's harmony generalization is no longer tenable with them as exceptions (which, as in Belth 2024, will be measured with the Tolerance principle), then the learner will again change representations, excluding the neutral vowel because it is no longer sufficiently predictive, thereby rendering it transparent. Thus, transparent vowels can for a time be tolerated as opaque vowels with lexicalized exceptions. This is the main idea underlying my proposed explanation for the observed experimental differences in learning.

In the next section § 2, I introduce the model from Belth (2024) (D2L) in more detail. In § 3, I survey prior experimental work on learning transparent and opaque vowels. I then demonstrate how D2L accounts for these experimental results, as conceptually described above, and also demonstrate that a number of other models fail to account for them § 4. I conclude with a discussion § 5.

2 Model

The model from Belth (2024), named D2L, was based on the developmental trajectory of children's ability to track adjacent and non-adjacent dependencies. Children show evidence of tracking adjacent dependencies at a younger age—as young as 8 months (Saffran et al., 1996, 1997; Aslin et al., 1998)—than tracking non-adjacent dependencies, which appears to develop around 15-18 months (Santelmann and Jusczyk, 1998; Gómez, 2002; Gómez and Maye, 2005). Tracking of adjacent dependencies has been observed over a range of different kinds of structures, linguistic and non-linguistic, including shapes (Fiser and Aslin, 2002) and non-linguistic tones (Saffran et al., 1999). These results serve as evidence of a language-independent psychological mechanism—the ability to track adjacent dependencies—that could underlie the learning of phonological alternations.

D2L implements the proposal that when learning a phonological alternation, a learner's attention is drawn to the alternating segment, and they begin tracking segments adjacent to it. I will use Finley (2015)'s artificial vowel harmony system as an example (see § 3) to describe the model as it pertains to the present paper. In (5), the underlying /-V/ suffix alternates between [-e] ~ [-o].[1]

[1] See Belth (2023a,b) for a proposal on how learners might come to attend to learning an alternation in the first place, and where the underlying forms might come from.

(5) /budok-V/ → [budoko]
/degib-V/ → [degibe]
/gemit-V/ → [gemite]
/kukop-V/ → [kukopo]
/tedɛt-V/ → [tedɛte]
/dotɛb-V/ → [dotɛbe]

D2L's attention is centered around /-V/ and the segments adjacent to it—here, the stem-final segments. D2L attempts to enforce harmony using the final segments, but since they are all consonants, the harmony fails. The learner then creates a new representation, excluding any adjacent segments that harmonizing with fails to yield the observed surface form for /V/—here /k, b, t, p/. D2L attempts to form a natural class for these segments, in this case [−syl]. The new representation is the complement of this *deletion set*, namely [+syl]. Clearly, this has the interpretation of a vowel tier (6).

(6) /uo-V/ → [uoo]
/ei-V/ → [eie]
/ei-V/ → [eie]
/uo-V/ → [uoo]
/eɛ-V/ → [eɛe]
/oɛ-V/ → [oɛe]

D2L then tracks segments adjacent to /-V/ on this new representation. The vowel [ɛ] here is opaque, so harmonizing with the adjacent vowel on this representation yields the expected surface realizations of /-V/ and D2L has succeeded in forming a representation and generalization that sufficiently accounts for the alternation. Following the notation from Belth (2024), (7) shows the generalization, where the vowel /V/ agrees in the value for feature [back] with an adjacent [+syl] segment after projecting vowels.

(7) Agree(V, [back]) / [+syl] ___ ∘ proj([+syl])

If, on the other hand, the vowel [ɛ] were transparent, the surface form of /dotɛb-V/ would be [dotɛbo], in which case enforcing harmony on the new representation would yield the wrong surface form for /-V/: *[e] instead of [o] (8).

(8) /oɛ-V/ → *[oɛe]

In this way, stems where a back vowel precedes a transparent front vowel will be exceptions to the generalization D2L forms on the new representation. D2L changes representations whenever the

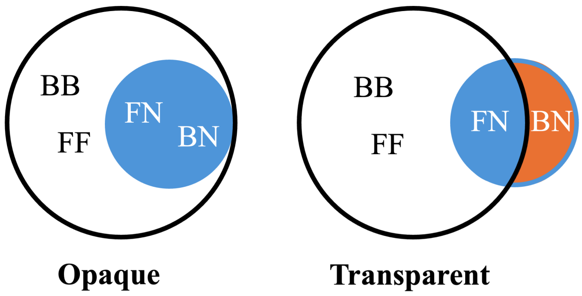

Figure 1: B = Back, F = Front, N = Neutral (opaque or transparent, depending on condition). The black circle represents the stems that are predictable from an adjacent vowel once D2L has constructed a vowel tier. The blue circle represents stems where the adjacent vowel is neutral. The orange sub-circle represents the only stems for which the suffix is not predictable from the tier-adjacent vowel.

generalization it forms over its current representation fails to sufficiently account for the alternation. D2L uses the Tolerance Principle (TP; Yang 2016), which has been evaluated in experimental settings (Schuler et al., 2016; Shi and Emond, 2023), to decide whether the generalization can sustain a particular number of exceptions (9).

(9) **Tolerance Principle**: a rule applying to n items with e exceptions is productive *iff* $e \leq \frac{n}{\ln n}$

Thus, D2L will only change representations again if the number of exceptions due to harmonizing with the transparent vowel, relative to the total number of alternating items, rises above the TP threshold (9). If the number of exceptions fall below the threshold, then D2L lexicalizes the exceptions and may overextend harmony with the final vowel (7) to new words with a final transparent vowel. On the other hand, if the number of exceptions grows too large, D2L will recursively construct a new representation, this time excluding the vowel [ɛ]—the culprit behind the exceptions—in addition to the consonants, as (10) shows.

(10) AGREE(V, [back]) / [+syl] __
$\circ$ proj([+syl] \ {ɛ})

This core idea is visualized in Figure 1. Once D2L has constructed a new representation that excludes consonants (i.e., a vowel tier), the suffix vowel is entirely predictable from the newly-adjacent vowel if the neutral vowel (N) is opaque. This set of stems, for which the suffix is adjacently predictable, is represented by the large black circle.

Table 1: The four basic kinds of training items in Finley (2015)'s study. B = Back, F = Front, N = Neutral (opaque or transparent, depending on condition). The right two columns give the suffix corresponding to the condition (only the BN items differ between conditions)

Kind	Types	Example	Opaque	Transparent
BB	8	[budok]	[-o]	[-o]
FF	8	[degib]	[-e]	[-e]
FN	4	[tedɛt]	[-e]	[-e]
BN	4	[dotɛb]	[-e]	[-o]

On the other hand, if the neutral vowel is transparent, only BN words are not adjacently predictable (the orange sub-circle). But if the orange part of the diagram is small enough, then it may be relegated to lexicalization, at least for a time.

3 Prior Experimental Studies

Finley (2015) carried out a series of artificial grammar learning studies with adults, involving opaque and transparent vowels. Finley first compared each of two experimental groups—one OPAQUE and one TRANSPARENT—to relevant control groups. The experimental groups were trained on CVCVC nonce words, each of which could be suffixed with either front [-e] or back [-o]. The artificial language also had the vowels [i, u] and the neutral vowel [ɛ], which only occurred as the final vowel. The choice of suffix was based on harmony with the final stem vowel, except for the words in the TRANSPARENT condition that had the transparent [ɛ] as the final vowel; for these the choice was based instead on harmony with the penultimate vowel. This is summarized in Table 1. There were 8 stems each with two harmonizing vowels (8 BB and 8 FF), 4 stems with a front vowel before the neutral [ɛ] (FN), and 4 with a back vowel before it (BN).

In the OPAQUE condition, if learners choose the suffix based on the final stem vowel, they would accurately generalize to test words of all four kinds. In the TRANSPARENT condition, however, accurate generalization to test BN words would require learning the transparency of [ɛ]. In other words, because [ɛ] is front, only BN words show unambiguous evidence that [ɛ] is transparent rather than opaque. Finley's first experiment, which presented each stem-suffixed pair 5 times, suggested that the participants in the OPAQUE condition learned vowel harmony, including the behavior of the opaque vowel. However, participants in the TRANSPARENT condition learned the basic vowel harmony pattern, but showed no evidence of learning the behavior of the transparent vowel.

Finley then attempted to find conditions in which participants would succeed at learning the transparent vowel's behavior. In a second experiment, the 4 FN words, for which it is ambiguous whether the [-e] vowel is harmonizing with the final or penultimate vowel (which are both front), were replaced with 4 additional BN words (all taking [-o]). This decreased the learners' test performance across the board. One interpretation is that because the suffix [-o] became more dominant—now occurring with 2/3 of training items—learners failed to attend to learning the alternation at all.

Finley then returned to the original setup (balanced items between FN and BN), and tried replacing the neutral vowel [ɛ] with [ɪ]. The participants again learned the overall harmony pattern, but not the transparent vowel. In another experiment, each word was presented 10 times instead of 5. This led to an increase in performance on the transparent vowel, but the increase over the control group was not statistically significant. The next experiment added 6 additional unambiguously transparent (BN) stems, with all words being presented 10 times. This also led to an increase, though not statistically significant, in performance on the transparent vowel. Finally, increasing the number of presentations of the BN stems to 20, while keeping the others at 10, led to an increase in performance on the transparent vowel that was significantly higher than the control group's.

The overall picture is that under some conditions where adults will learn a vowel harmony system with an opaque vowel, they will fail to learn a transparent vowel. But, if sufficient exposure to words that demonstrate the transparency of a vowel is available, adults will succeed at learning its transparency. While this overall picture is clear, the precise conditions in which learning a transparent vowel will or will not succeed are less so. In multiple of Finley (2015)'s experiments, the results showed a numerical increase in performance that was not statistically significant. The number of participants in some experiments was small (often < 20 per condition), thus warranting a level of caution in drawing strong conclusions from any particular significance test. The study involved adults, but we also know that children acquire vowel harmony systems with transparent vowels (MacWhinney, 1978; Gósy, 1989; Leiwo et al., 2006; Gonzalez-Gomez et al., 2019). Moreover, the stimuli were presented in auditory form only, with no accompanying image. It is difficult to know in such a scenario whether participants treated multiple tokens of the same type as in fact being part of a single word type or of multiple. Consequently, the relative role of type and token frequencies is not entirely clear.

Furthermore, Ozburn et al. (2016) note that Finley's artificial language used the English vowel inventory, which leads to both roundness and backness alternating ([-e] is front, unround; [-o] back, round), which is not typical in natural language backness harmony with transparent vowels. Ozburn et al. trained adult Canadian French speakers in a similar setting as Finley's, but using harmony centered around the French vowel inventory, which includes front rounded vowels, allowing for the rounding dimension to stay fixed. Ozburn et al.'s participants did show evidence of learning vowel harmony transparency in this setting. However, whether this difference in results from Finley's was due to the difference in stimuli and participant populations or to difference in type frequency is not clear: Ozburn et al. do not report how many items of each kind they used in their experiment, but they do say that 1/4 of the items were unambiguously transparent (BN), which is a higher proportion than in Finley's experiments (1/6 to 1/5).

In a related study, Chen (2024) trained adult speakers of Taiwan Mandarin on an artificial vowel harmony pattern with both an opaque and a transparent vowel. The study was primarily interested in a possible "starting small" effect—whether presenting bisyllabic stems before trisyllabic stems, and a disproportionate number of bisyllabic stems, would yield better learning than presenting a balanced number all at once. In the results, only in the "starting small" condition did participants show evidence of learning the vowel harmony pattern. However—more relevant to the current discussion—even in this condition, participants only showed learning of the non-transparent vowels. They appeared to treat the transparent vowel as also opaque. Thus, while this study deviates substantially from the prior two in goals and design, the results largely corroborate the big picture of Finley (2015)'s study: opaque vowels are learned more readily by adults than transparent vowels.

4 Evaluation

To evaluate whether D2L makes sense of the experimental results on opaque and transparent vowels, I tested whether D2L learns an opaque vowel in conditions where it does not learn a transparent vowel (§ 4.2), and whether increasing the amount of training on items showing transparency eventually leads it to learn a transparent vowel (§ 4.3). First, I will introduce the setup (§ 4.1).

4.1 Data and Setup

I used data from Finley (2015)'s study for training and evaluation. As the base training set, I used the same 24 stem-suffixed pairs that Finley, p. 22 reports; these are summarized in Table 1.

In experimental settings (as in natural language learning), participants likely do not learn every word they are trained on. Yet it is over the words that are learned that generalizations can be formed.[2] To simulate this variability in attained vocabulary, I carried out 30 simulations with different samples of training words. For each, I sampled an integer n from a Gaussian distribution with mean 20 and standard deviation of 4 to represent the vocabulary size. I then sampled n unique words from the 24 training words, weighted by frequency. In the first experiment (§ 4.2) all words were given equal frequency, so the sampling was uniform. In the second experiment (§ 4.3), where the amount of exposure to unambiguously transparent (BN) words is increased, this sampling procedure allows for manipulating the saliency of BN words, as Finley (2015) did, by increasing their relative token frequency.

For testing, I used the novel stems from Finley, p. 23. These include 8 stems with two harmonizing vowels (BB or FF) and 11 ending in the neutral [ɛ]. Of the latter, 9 are BN.

4.1.1 Comparison Models

In a study of vowel harmony in Hungarian, Hayes and Londe (2006) proposed two harmony constraints, applying over a vowel tier. The first, local, constraint incurred a violation whenever a front vowel immediately followed a back vowel on the vowel tier, and the second, distal, constraint incurred a violation whenever a front vowel followed a back vowel anywhere on the tier. The distal constraint was necessary because of Hungarian's transparent vowels. Finley (2015) reasoned that the distal constraint is more complex than the local constraint, and thus could make harmony more difficult to learn when transparent vowels are present. This forms the first comparison model: I trained a Maximum Entropy Harmonic Grammar model using distal and local constraints like Hayes and Londe's. The model learns to map underlying forms (e.g., /dotɛb-V/) to surface forms, using a Maximum Entropy model, as described by Goldwater and Johnson (2003). For each underlying form, two candidates are generated—one with [-o] and one with [-e]—and the number of violations of local and distal harmony constraints are used as the features of each candidate. I will call this model H&L, as an homage to Hayes and Londe (2006).

While H&L learns a Maximum Entropy grammar with provided constraints, it is also possible for constraints to be learned. Indeed, building on Hayes and Wilson (2008)'s model, Gouskova and Gallagher (2020) proposed a Maximum Entropy model that automatically learns to project tiers and form phonotactic constraints over the resulting tier projections. I used the model publicly available from the authors.[3] I will call this model G&G.

Lastly, vowel harmony can typically be characterized as 2-Tier-Strictly-Local (2TSL), whether described as phonotactic constraints (Heinz et al., 2011) or processes (Burness et al., 2021).[4] This is usually true even when opaque or transparent vowels are present. Formal learning algorithms have been proposed that allow for proving the efficient learnability of 2TSL languages and functions (Jardine and Heinz, 2016; Burness and McMullin, 2019). However, while these learnability results apply to vowel harmony with opaque or transparent vowels, it does not necessarily imply that languages with either of these kinds of neutral vowels will be learned at equal rates. Like D2L, the Jardine and Heinz (2016) and Burness and McMullin (2019) models start with a representation where all segments are present, and iteratively remove segments to create new tiers. Unlike D2L, they use the formal properties of TSL to deduce conditions where removing segments is provably correct. Thus, I use TSLIA (Jardine and Heinz, 2016; Jardine and McMullin, 2017), which is publicly available (Aksënova, 2020), as an additional comparison model. Formal models of this family

[2] See, for instance, Schuler (2017, ch. 4) for discussion of this point for artificial language learning with children.

[3] github.com/gouskova/inductive_projection_learner

[4] See Mayer and Major (2018) for an example of a harmony pattern than cannot be characterized as TSL.

often benefit from collapsing pattern-irrelevant differences among segments (Aksënova, 2020; Johnson and De Santo, 2023), which simplifies the learning problem and makes it more likely that the characteristic sample (the information needed in the training data for convergence onto an appropriate grammar) will be present. Following this line of work, I collapsed all consonants into the symbol C, back vowels to B, non-neutral front vowels to F, and neutral vowels to N. This collapsing was only applied to TSLIA's input, not the other models'.

For D2L, I used the implementation publicly available in the Python package *algophon*.[5]

In the experiments, each test stem has two possible suffixed forms: [-e] or [-o]. I compute a model's accuracy based on the fraction of stems for which it produces/chooses the form consistent with the relevant vowel harmony pattern. Specifically, the correct choice for BB and FF is the vowel that agrees in backness with the final stem vowel. In OPAQUE conditions, the correct choice for neutral-vowel-final stems is [-e], while in TRANSPARENT conditions, it is the vowel agreeing with the penultimate stem vowel. I report overall accuracy and neutral-vowel accuracy, which is computed over only the neutral-vowel-final test stems.

This scheme can be interpreted as either learning an alternation (mapping a stem with underlying /-V/ to the surface form) or a phonotactic pattern (learning where [-e] and [-o] can/cannot occur). D2L and H&L learn alternations, while G&G and TSLIA learn phonotactics. At test time, the former are probed to produce a surface form for a stem with the underlying suffix /-V/ and the produced form is taken as the choice. Meanwhile, the phonotactic models are asked to score the two choices and the one with the better well-formedness score is chosen. This setup is identical to Belth (2024)'s.

4.2 Opaque vs. Transparent

The first experiment evaluates whether D2L and the comparison models show a difference in generalization between an OPAQUE vowel harmony condition and a TRANSPARENT condition (learning the former better). The experiment uses the training data described above (§ 4.1), training 30 models in each of the two conditions, where the number of words for each simulation is $n \sim Normal(20, 4)$.

Figure 2 shows the accuracy on all test words (All) and accuracy on test words where the final vowel is neutral (Neutral). D2L's accuracy, in both cases, is higher for the OPAQUE condition than the TRANSPARENT condition, consistent with the overall picture that humans are better at learning harmony with an opaque vowel (§ 3). D2L shows this asymmetry because, in most TRANSPARENT samples, the number of exceptions introduced by BN stems does not rise above the TP threshold (9), so D2L does not create a new representation.

No other model shows this pattern. H&L and G&G learn both kinds of harmony equally well. Thus, while Finley (2015) conjectured that the added complexity of Hayes and Londe (2006)'s distal harmony constraint might translate into difficulty learning transparent harmony, when tested on even this quite small amount of data, there is enough input to assign a weight to the distal constraint large enough for the transparent vowel to be learned. Perhaps surprisingly, even G&G, which learns to project tiers and learns its constraints, also fails to show any difference between conditions. In the OPAQUE condition, G&G consistently finds a trigram constraint that marks vowels differing in backness across another segment. This is sufficient to learn the harmony pattern. In the TRANSPARENT condition, G&G learns a similar constraint, but only specific to the harmonizing (non-neutral) vowels. G&G then projects a tier that includes only the vowels in that constraint—the non-neutral vowels. Then, on this projection, G&G learns a new constraint that marks disharmony between vowels on the tier—which excludes the transparent vowel. Thus, G&G learns transparency in conditions where humans do not.

TSLIA does not learn either harmony pattern. This indicates that there is no characteristic sample present in the data. This is true even though I collapsed irrelevant differences among segments (e.g. all consonants were mapped to the symbol C, as described in § 4.1.1), which simplifies the learning problem and in some cases leads learners of this sort to succeed at learning (Aksënova, 2020; Johnson and De Santo, 2023). Running the model without collapsing segments yields the same results.

4.3 Eventual Learning of Transparent

In the second experiment, I evaluated whether D2L and the comparison models get better at learning a transparent vowel as the amount of training exposure to words that unambiguously show the transparency of the vowel increases. This follows the same setup as the TRANSPARENT condition above,

[5] github.com/cbelth/algophon/tree/main/algophon/models

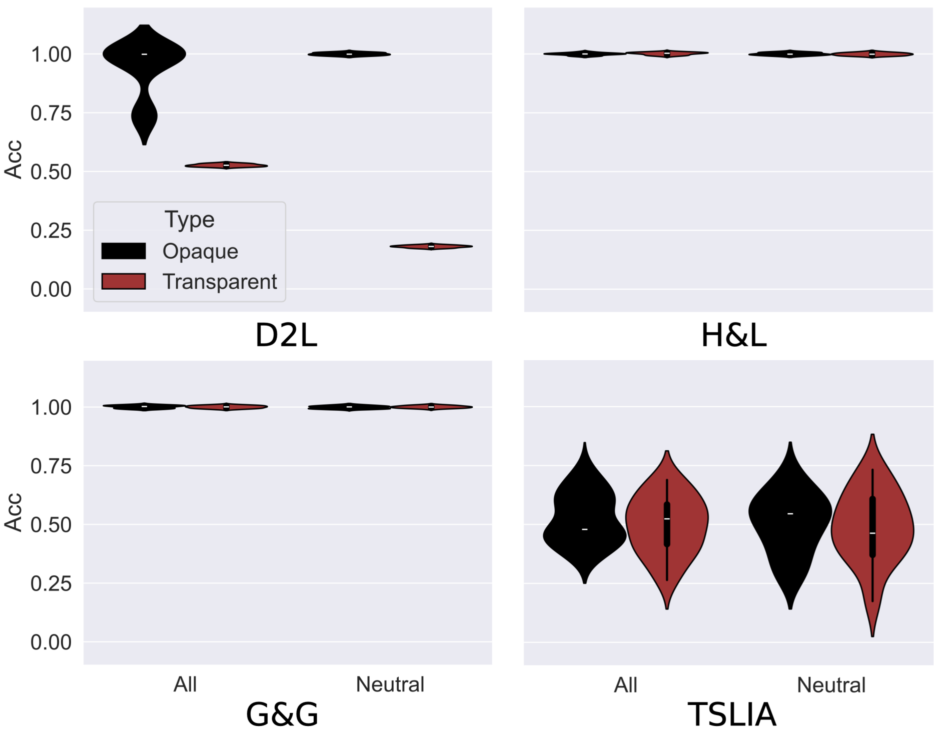

Figure 2: The distribution of accuracies (over All test words and over Neutral test words) of each model in Opaque and Transparent conditions. Only D2L shows a difference in accuracy between conditions, as humans do.

but varies two parameters: the number of BN (unambiguously transparent) types (4, 6 or 8), and the relative token frequency of those types (1x, 2x, or 5x the token frequency of non-BN types). Since the number of words for each simulation is $n \sim Normal(20, 4)$ and the choice of those n words is based on a sample weighted by token frequency, varying the relative token frequency of the BN words increases the probability that they enter into a particular learner's vocabulary. Thus, the token frequency also influences the type frequency of BN words, but in a different way. Increasing the type frequency was accomplished by replacing FN words with BN words (so the total number of words available was always 24). Combining these variations means there are 9 conditions per model. I ran 30 simulations (different seeds) for each model in each condition.

Figure 3 gives the results, where the top row of heatmaps is accuracy over all words and the bottom row is accuracy over words with neutral vowels. If a model mirrors the basic pattern of humans, who get better with transparency as exposure to BN increases, then accuracy should increase (darker colors) as the type frequency increases (rightward movement) and/or relative token frequency increases (downward movement)—in other words if more rightward and lower cells are darker. This is the case for D2L, but no other model. Increasing the prevalence of BN exceptions eventually leads D2L to form a new representation that excludes [ɛ]. H&L and G&G are dark in all cells, mirroring the above results where they learn transparent vowels when humans do not. TSLIA is again at chance across the board. D2L's performance is tied to increases in type frequency, which is consistent with arguments and evidence that type frequency, rather than token frequency, plays the primary role in the formation of linguistic generalizations (Aronoff, 1976; MacWhinney, 1978; Baayen, 1993; Elman, 1998; Pierrehumbert, 2001; Albright and Hayes, 2003; Endress and Hauser, 2011; Yang, 2016).

5 Conclusion and Discussion

Do opaque and transparent vowels do different things to a vowel harmony system? From one perspective, transparent vowels introduce non-locality that opaque vowels do not (Goldsmith, 1985; Bakovic and Wilson, 2000; Hayes and Londe, 2006; Finley, 2009). From another perspective, neither opaque nor transparent vowels change the kind of information needed to capture the harmony generalization: in both cases there is a set of seg-

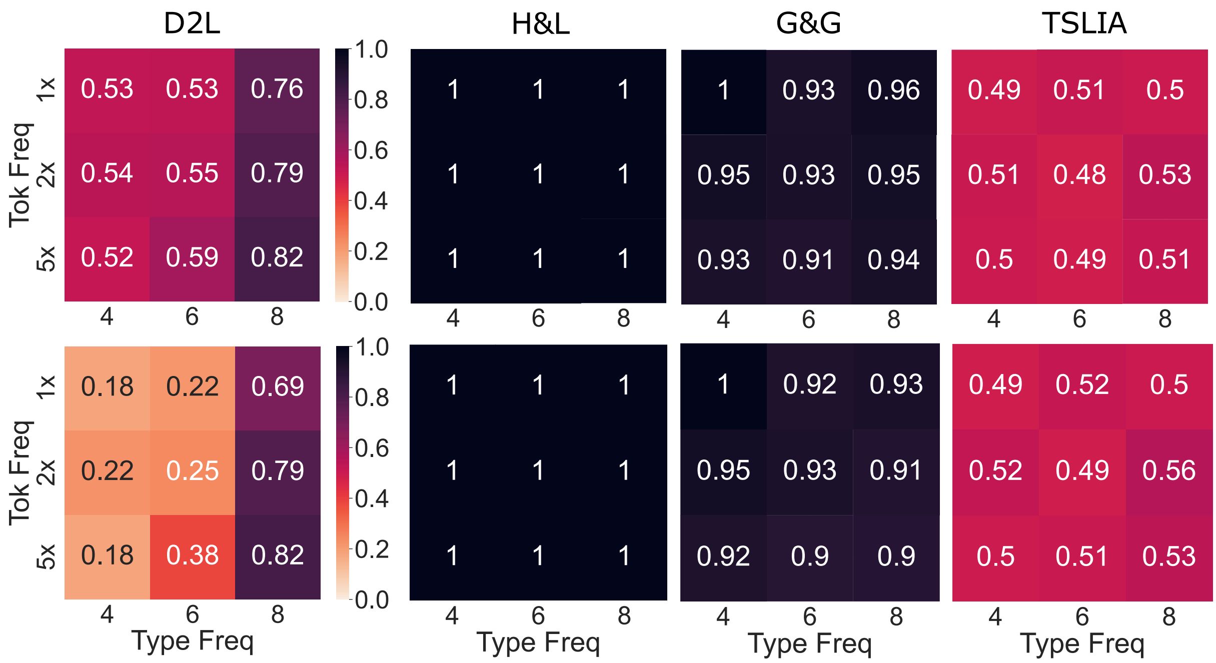

Figure 3: Heatmaps showing the accuracy of each model when trained on a vowel harmony pattern with a transparent vowel. The top row shows accuracy across all test words; the bottom shows accuracy across words where the final vowel is transparent. Matching the trend from human learners in laboratory settings would yield an accuracy gradient that increases as the type and/or token frequency of words exhibiting unambiguous transparent vowel harmony increases. D2L matches this general trend. The same cannot be said of any other evaluated models.

ments that can be projected (a tier) that renders all the dependencies local (Heinz et al., 2011; Burness and McMullin, 2019). One way to approach this question is to take the perspective of the learner. In Belth (2024), I proposed that learners construct new representations only when the ones they are currently generalizing over let them down. The results in that article demonstrated that in natural language harmony systems, this approach leads to accurate generalization to test words. Trained on a few hundred words from Turkish, where low vowels are opaque to rounding harmony, or Finnish, where [i, e] are transparent to backness harmony, D2L constructed representations that allowed for forming a successful harmony generalization. In this paper, I have demonstrated that in Finley (2015)'s setting, the same model constructs a vowel tier and only when a transparent vowel introduces enough exceptions does the model again construct a new representation, then generalizing to transparent vowels. Thus, in this proposal, there is a difference between opaque and transparent vowels—but only for a time.

Further research into the factors influencing human leaning of vowel harmony in the presence of opaque and transparent vowels—in particular children's learning and acquisition—would be of great value. For instance, D2L predicts that, if the conditions are right, there could be a stage of acquisition where learners incorrectly harmonize alternating vowels with preceding transparent vowels. In the limited number of developmental studies on the acquisition of vowel harmony systems with transparent vowels (MacWhinney, 1978; Gósy, 1989; Leiwo et al., 2006), I am not aware of reports of such errors (see Goad and Ozburn 2024 for a recent survey). However, if such a stage exists, D2L predicts it to be transient, since accumulating exceptions would lead to recursive creation of a new representation. Moreover, it is only a subset of words (BN stems in the languages discussed here) that have the potential of showing such overgeneralization. And over-application of generalizations to a particular word is influenced by the strength of the word's lexical representation, which in turn is influenced by its token frequency (Hooper, 1976; Bybee, 1985; Marcus et al., 1992; Bybee, 1995). Errors are thus more likely on low-token-frequency words, which are less represented in child speech. Consequently, identifying whether this is indeed a developmental stage would likely require studies aimed precisely at this question.

References

Alëna Aksënova. 2020. *Tool-assisted induction of subregular languages and mappings*. Ph.D. thesis, State University of New York at Stony Brook.

Adam Albright and Bruce Hayes. 2003. Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition*, 90(2):119–161.

Mark Aronoff. 1976. Word formation in generative grammar. *Linguistic Inquiry Monograph*, 1.

Richard N Aslin, Jenny R Saffran, and Elissa L Newport. 1998. Computation of conditional probability statistics by 8-month-old infants. *Psychological science*, 9(4):321–324.

Harald Baayen. 1993. On frequency, transparency and productivity. In *Yearbook of morphology 1992*, pages 181–208. Springer.

Eric Bakovic and Colin Wilson. 2000. Transparency, strict locality, and targeted constraints. In *West Coast Conference on Formal Linguistics.*, pages 43–56.

Caleb Belth. 2023a. Towards a learning-based account of underlying forms: A case study in Turkish. In *Proceedings of the Society for Computation in Linguistics 2023*, pages 332–342, Amherst, MA. Association for Computational Linguistics.

Caleb Belth. 2023b. *Towards an Algorithmic Account of Phonological Rules and Representations*. Ph.D. thesis, University of Michigan.

Caleb Belth. 2024. A learning-based account of phonological tiers. *Linguistic Inquiry*, pages 1–37.

Stefan Benus and Adamantios I Gafos. 2007. Articulatory characteristics of Hungarian 'transparent'vowels. *Journal of Phonetics*, 35(3):271–300.

Phillip Burness and Kevin McMullin. 2019. Efficient learning of Output Tier-based Strictly 2-Local functions. In *Proceedings of the 16th Meeting on the Mathematics of Language*, pages 78–90, Toronto, Canada. Association for Computational Linguistics.

Phillip Burness, Kevin McMullin, and Jane Chandlee. 2021. Long-distance phonological processes as tier-based strictly local functions. *Glossa: a journal of general linguistics*, 6(1).

Joan Bybee. 1985. *Morphology: A study of the relation between meaning and form*. John Benjamins, Philadelphia.

Joan Bybee. 1995. Regular morphology and the lexicon. *Language and cognitive processes*, 10(5):425–455.

Tsung-Ying Chen. 2024. The "starting-small" effect in phonology: Evidence from biased learning of opaque and transparent vowel harmony. *Language and Speech*, page 00238309241230625.

George N Clements. 1976. *The autosegmental treatment of vowel harmony*. Indiana University Linguistics Club.

George N Clements. 1980. *Vowel harmony in nonlinear generative phonology*. Indiana University Linguistics Club Bloomington.

Jeffrey Elman. 1998. Generalization, simple recurrent networks, and the emergence of structure. In *Proceedings of the twentieth annual conference of the Cognitive Science Society*, page 6. Mahwah, NJ: Lawrence Erlbaum Associates.

Ansgar D Endress and Marc D Hauser. 2011. The influence of type and token frequency on the acquisition of affixation patterns: Implications for language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(1):77.

Sara Finley. 2009. *Formal and cognitive restrictions on vowel harmony*. The Johns Hopkins University.

Sara Finley. 2015. Learning nonadjacent dependencies in phonology: Transparent vowels in vowel harmony. *Language*, 91(1):48.

József Fiser and Richard N Aslin. 2002. Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(3):458.

Heather Goad and Avery Ozburn. 2024. Vowel harmony in language acquisition. In *The Oxford Handbook of Vowel Harmony*. Oxford University Press.

John Goldsmith. 1976. *Autosegmental phonology*. Ph.D. thesis, Massachusetts Institute of Technology.

John Goldsmith. 1985. Vowel harmony in Khalkha Mongolian, Yaka, Finnish and Hungarian. *Phonology Yearbook*, 2(1):253–275.

Sharon Goldwater and Mark Johnson. 2003. Learning OT constraint rankings using a maximum entropy model. In *Proceedings of the workshop on variation within Optimality Theory*, pages 111–120.

Rebecca Gómez and Jessica Maye. 2005. The developmental trajectory of nonadjacent dependency learning. *Infancy*, 7(2):183–206.

Rebecca L Gómez. 2002. Variability and detection of invariant structure. *Psychological Science*, 13(5):431–436.

Nayeli Gonzalez-Gomez, Silvana Schmandt, Judit Fazekas, Thierry Nazzi, and Judit Gervain. 2019. Infants' sensitivity to nonadjacent vowel dependencies: The case of vowel harmony in hungarian. *Journal of Experimental Child Psychology*, 178:170–183.

Mária Gósy. 1989. Vowel harmony: interrelations of speech production, speech perception, and the phonological rules. *Acta Linguistica Hungarica*, 39(1/4):93–118.

Maria Gouskova and Gillian Gallagher. 2020. Inducing nonlocal constraints from baseline phonotactics. *Natural Language & Linguistic Theory*, 38(1):77–116.

Bruce Hayes and Zsuzsa Cziráky Londe. 2006. Stochastic phonological knowledge: The case of Hungarian vowel harmony. *Phonology*, 23(1):59–104.

Bruce Hayes and Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic inquiry*, 39(3):379–440.

Jeffrey Heinz, Chetan Rawal, and Herbert G Tanner. 2011. Tier-based strictly local constraints for phonology. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human language technologies*, pages 58–64.

Joan B. Hooper. 1976. Word frequency in lexical diffusion and the source of morphophonological change. In William M. Christie, editor, *Current progress in historical linguistics*, pages 96–105. North Holland, Amsterdam.

Adam Jardine and Jeffrey Heinz. 2016. Learning tier-based strictly 2-local languages. *Transactions of the Association for Computational Linguistics*, 4:87–98.

Adam Jardine and Kevin McMullin. 2017. Efficient learning of tier-based strictly k-local languages. In *Language and Automata Theory and Applications*, pages 64–76. Springer.

Jacob K Johnson and Aniello De Santo. 2023. Evaluating a phonotactic learner for MITSL-(2, 2) languages. *Society for Computation in Linguistics*, 6(1):379–382.

Bariş Kabak. 2011. Turkish vowel harmony. *The Blackwell companion to phonology*, pages 1–24.

Matti Leiwo, Pirjo Kulju, and Katsura Aoyama. 2006. The acquisition of Finnish vowel harmony. *Finnish Journal of Linguistics*, (19):149–161.

Brian MacWhinney. 1978. The acquisition of morphophonology. *Monographs of the society for research in child development*, pages 1–123.

Gary F Marcus, Steven Pinker, Michael Ullman, Michelle Hollander, T John Rosen, Fei Xu, and Harald Clahsen. 1992. Overregularization in language acquisition. *Monographs of the society for research in child development*, pages i–178.

Connor Mayer and Travis Major. 2018. A challenge for tier-based strict locality from uyghur backness harmony. In *Formal Grammar 2018: 23rd International Conference*, pages 62–83. Springer-Verlag.

Andrew Nevins. 2010. *Locality in vowel harmony*. Linguistic Inquiry Monographs. Mit Press.

Avery Ozburn, G Hansson, and Kevin McMullin. 2016. Learning vowel harmony with transparency in an artificial language. In *Talk Presented at the 2016 NOWCAM Meeting: Eugene Oregon*.

Janet Pierrehumbert. 2001. Stochastic phonology. *Glot international*, 5(6):195–207.

Catherine O Ringen and Orvokki Heinämäki. 1999. Variation in Finnish vowel harmony: An OT account. *Natural Language & Linguistic Theory*, 17(2):303–337.

Jenny R Saffran, Richard N Aslin, and Elissa L Newport. 1996. Statistical learning by 8-month-old infants. *Science*, 274(5294):1926–1928.

Jenny R Saffran, Elizabeth K Johnson, Richard N Aslin, and Elissa L Newport. 1999. Statistical learning of tone sequences by human infants and adults. *Cognition*, 70(1):27–52.

Jenny R Saffran, Elissa L Newport, Richard N Aslin, Rachel A Tunick, and Sandra Barrueco. 1997. Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological science*, 8(2):101–105.

Lynn M Santelmann and Peter W Jusczyk. 1998. Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition*, 69(2):105–134.

Kathryn D Schuler, Charles Yang, and Elissa L Newport. 2016. Testing the tolerance principle: Children form productive rules when it is more computationally efficient to do so. In *CogSci*, volume 38, pages 2321–2326.

Kathryn Dolores Schuler. 2017. *The acquisition of productive rules in child and adult language learners*. Ph.D. thesis, Georgetown University.

Rushen Shi and Emeryse Emond. 2023. The threshold of rule productivity in infants. *Frontiers in Psychology*, 14:1251124.

Charles Yang. 2016. *The price of linguistic productivity: How children learn to break the rules of language*. MIT press.