Modeling the Learning of Tier Projection with Blocking in Non-Local Phonological Processes

Yue Yin
The Ohio State University

1 Introduction

Phonological theories often address long-distance dependencies by treating them as adjacent at a certain level of representation, which can be achieved through tier projection (Vergnaud, 1977; Goldsmith, 1979; McCarthy, 1979). Previous studies have shown that tier projection is necessary for models to learn long-distance phonological processes (Hayes & Wilson, 2008; Goldsmith & Riggle, 2012) [1]. However, they provided a relevant tier projection a priori, assuming it to be an innate component of Universal Grammar (UG). Then, a critical question is whether a learner can discover the tiers automatically during the learning process.

To answer this question, some recent work, for example, the Tier-based Strictly k-Local Inference Algorithm (kTSLIA) (Jardine & McMullin, 2017), the Inductive Projection Learner (IPL) (Gouskova & Gallagher, 2020) and the more restrictive version (Kim, 2022), the Restrictive Tier Learner (RTL), as well as the Distant to Local model (D2L) (Belth, 2024), explored automatic tier induction mechanisms. Belth (2024) has shown that the D2L model can not only learn the correct generalizations from large-scale realistic language data, covering both assimilation and dissimilation, but also acquire these patterns from small-scale toy data used in artificial language learning experiments, and that its behavior closely resembles that of human participants across various conditions. In contrast, other models failed to reach a comparable level of accuracy. kTSLIA performed poorly on both kinds of datasets. IPL appeared unable to generalize from these small artificial language learning datasets; although its performance on the realist dataset was reasonably good, it still fell short of D2L, and it learned the dissimilation patterns very poorly. The better performance of the D2L model is because it is derived from more cognitively realistic mechanisms, i.e., the human preference for adjacency, which has been proved by previous empirical studies (e.g., Finley, 2011).

However, the D2L model fails to learn the correct tier and rule when the phonological processes are blocked by inert segments [2], which do not initiate or undergo assimilation (Clements, 1987). Unlike transparent segments, they do not allow the assimilating feature to pass through them. Instead, they prevent the target from alternating and cause it to take the default form. For example, in Khalkha Mongolian (Nevins, 2010: 137), the backness (rounding) harmony is blocked by high back (round) vowels, and the perfect suffix

[*] I would like to thank Micha Elsner, Rebecca Morley, Caleb Belth, the audience of OSU phonies discussion group, the anonymous reviewers, organizers, and audience of AMP 2025 for their helpful feedback and support. I assume responsibility for any and all errors.

[1] It is controversial whether all non-local phonological phenomena can be represented as feature spreading on a tier. For example, Heinz (2010) argued that consonantal harmony is not feature spreading since it doesn't allow blocking patterns. However, blocking patterns are widely attested in vowel harmony (Rose & Walker, 2011). Therefore, given that the current paper focuses on the blocking patterns in vowel harmony, it is reasonable to assume a tier-based phonological representation.

[2] Belth (2024) has tested the D2L model's capacity of capturing blocking on Finley (2011)'s experimental data and the Latin liquid dissimilation data. However, the "blockers" in those two cases are not inert blockers. For more details, see 4.1.

Proceedings of *AMP 2025*

would take the default [-back] ([-round]) value. These patterns are illustrated in the examples below [3].

(1) a. tor-oːd ‘be.born-PERF’
b. ɔr-ɔːd ‘enter-PERF’
c. tor-uːl-eːd ‘be.born-CAUS-PERF’
d. ɔr-ʊːl-aːd ‘enter-CAUS-PERF’

The D2L model learns the tier by removing all the projected adjacent segments that cannot account for the alternation from the tier incrementally, including such inert blockers, thus leading to wrong predictions.

The current research improves the D2L model by distinguishing transparent segments from inert blockers during tier updating, ensuring blockers remain on the tier but are excluded from the rule condition, thereby correctly modeling blocking effects in long-distance phonological processes.

2 Model

The learning procedures of the original D2L model and the current adjusted model are shown in the two flow charts in Figure 1 and Figure 2.

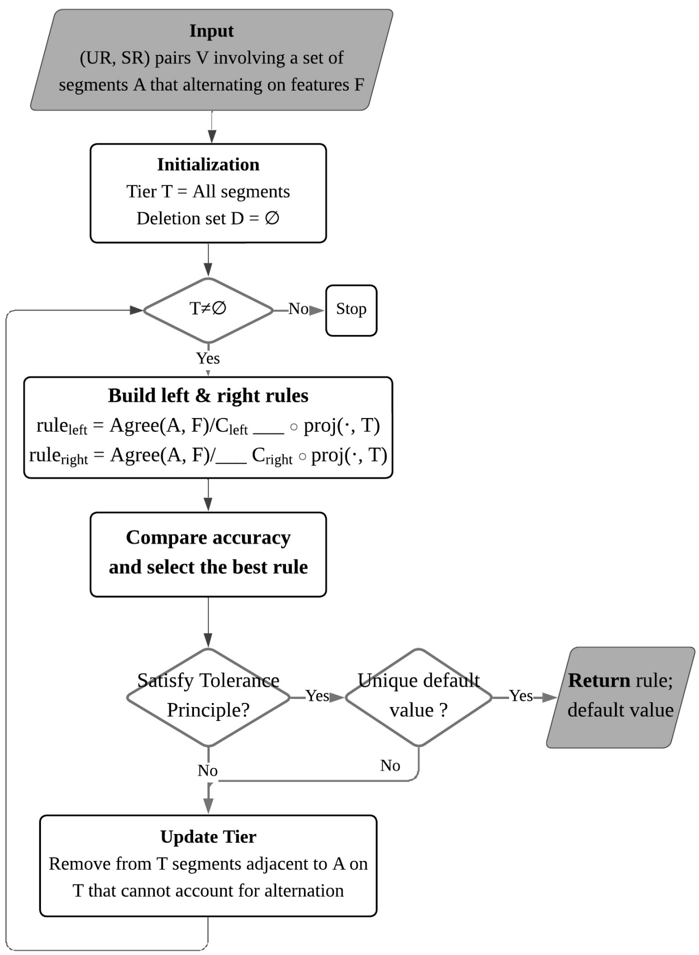

Figure 1: The learning procedures of the original D2L model.

[3] In addition to the backness (rounding) harmony, there is also [ATR] harmony in Khalkha Mongolian, but only the backness (rounding) harmony is blocked by high back vowels. /a, ɔ/ are [-ATR] and /e, o/ are [+ATR].

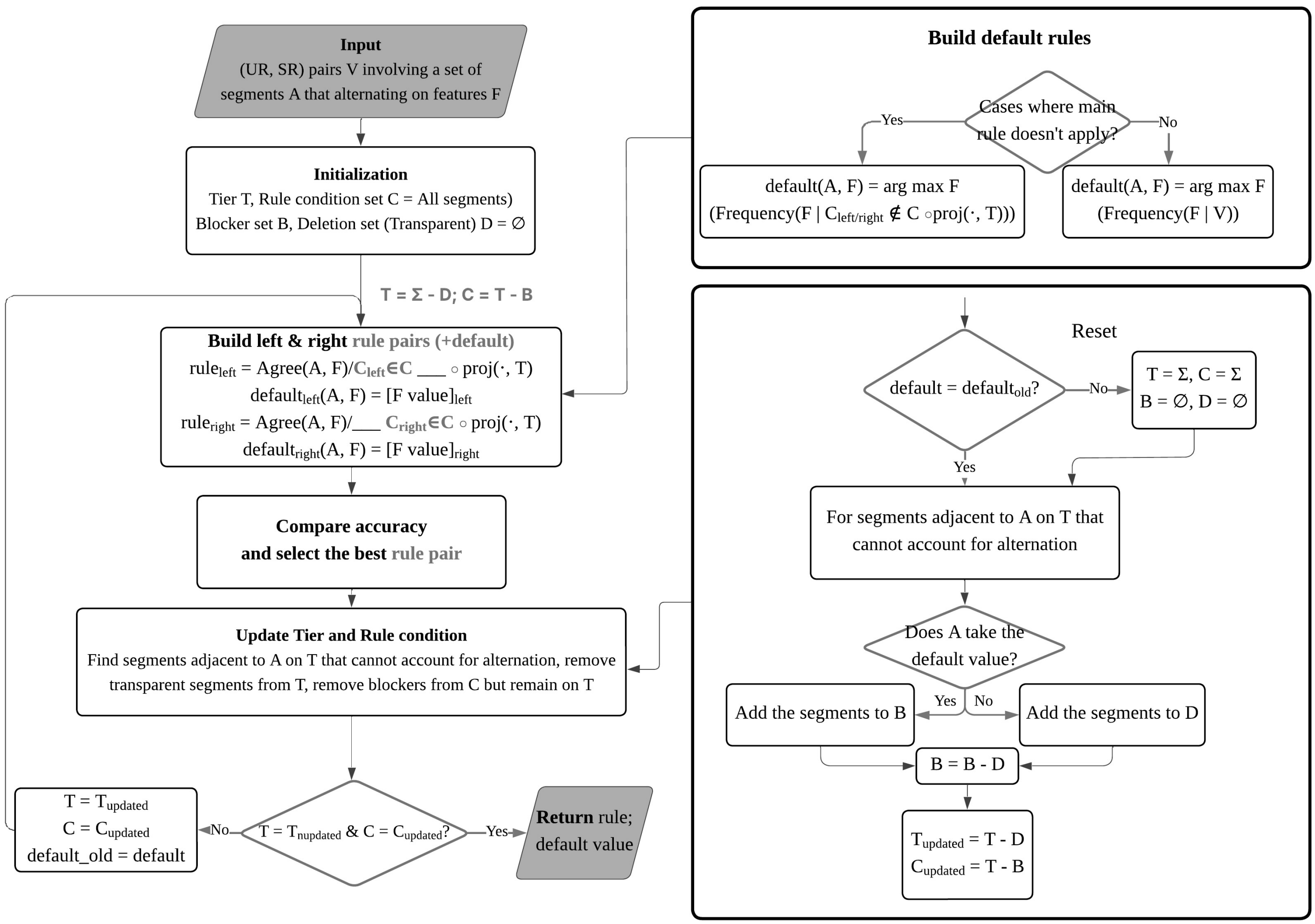

Figure 2: The learning procedures of the current adjusted model.

2.1 *The D2L Model* The model learns from input consisting of pairs of underlying representations (UR) and their corresponding surface representations (SR). In the URs, the alternating segments, i.e., the targets, are treated as underspecified with the alternating feature(s). In other words, there is no "default" value for the alternating feature in the underlying level. Rather, the assignment of the "default" value is treated as a surface-level operation. This input setting indicates that the learning of UR is before the learning of alternation patterns, which could be problematic for some more complicated situations that include rule interaction, e.g., counterfeeding and counterbleeding, but is fine for the current problem. Based on the working assumptions above, the learning of UR can also be automatically implemented (Belth, 2023). It should also be noted that the feature values of each phoneme are provided to the learner, which can be considered as part of the Universal Grammar (UG).

The model output is a local rule with tier projection, which can be either assimilation or dissimilation, and can apply leftward or rightward. The rule is applied to projected URs, unless no context segments are remaining in the projected tier. The target then takes the same (assimilation) or opposite (dissimilation) value as the left or right adjacent segment, depending on the direction of the rule. The default rule, if it exists, applies to all cases that the main rule doesn't apply to and assigns the target segment a default form.

As shown in Figure 1, the tier T is initialized as the full inventory, with the deletion set D, i.e., the set containing segments that should be deleted from the tier, as an empty set. Then the tier is iteratively updated. In each iteration, both a left and a right rule are established based on the current tier, and the rule with higher accuracy is used for the update. Based on the current tier, the model examines all (UR, SR) pairs and incrementally removes all projected adjacent segments that cannot account for the alternation from the tier. After each iteration, the updated tier is passed to the next iteration. The iteration ends when the rule can pass the Tolerance Principle (Yang, 2016) or no updates can be made to the tier. The whole rule and tier updating procedure is applied separately to assimilation and dissimilation, and the model ultimately selects the rule that yields higher accuracy.

The default rule is established at the very last step, and is only built when there are cases with no adjacent segment after tier projection. If those cases do not alternate, then the surface form is taken as the default form.

2.2 *The Current Revised Model* As shown in Figure 2, based on the D2L model, the current model has three key adjustments: (1) It further distinguishes between transparent segments and inert blockers. (2) It updates the tier and rule condition individually. (3) It builds rule pairs where each pair includes a main rule and its corresponding default rule at the very beginning of the rule construction step.

First, as shown in the initialization stage, in addition to the tier set T and deletion (transparent) set D, a blocker set B is also created, which is an empty set. The tier and the rule condition can then be distinguished based on the initial setting. The tier T is the completion set of the deletion set D, and the rule condition set C is the difference between the tier T and the blocker set B. In other words, the rule condition set contains all segments on the tier except for the inert blockers, since inert blockers are considered on the tier but not in the rule condition. At the beginning, both T and C are full sets.

Second, in the rule-building step, the constructed rule form and the rule application are slightly different from the original D2L model. In the current model, since both the tier and rule condition are initialized, the rule is established based on both the current tier and rule condition. Then, for rule application, in the original model, the main rule always applies unless there is no segment adjacent to the target on the tier. However, in the current model, since the rule condition is not equal to the adjacent segment after tier projection, after tier projection, the main rule applies only if the adjacent segment is in the rule condition set.

Third, a default rule is also created in the rule-building step, rather than in the last step of output. The rule-building module always creates rule pairs that each include a main rule and its corresponding default rule, and the best rule pair is then selected by comparing the combined accuracy of the main rule and its default rule. The main rule is the assimilation rule or dissimilation rule, the same as the rule in the original model; the default rule is the rule that assigns the target a feature value when the main rule doesn't apply. The default rule needs to be set at the beginning, because the default value will be used to identify blockers in later steps.

Then the question is how to determine the default value at the beginning. As shown in the top-right box in Figure 2, there are two possible ways to determine the default value. If the input (UR, SR) pairs include cases where the main rule doesn't apply, then the most frequent form among those cases is taken as the default. If there is no case where the main rule doesn't apply, the overall most frequent form is taken as the default.

Fourth, since the tier and the rule condition are distinguished, updating procedures are also adjusted, as shown in the bottom-right box in Figure 2. In the updating process, if a projected adjacent segment cannot account for the alternation by assimilation or dissimilation, the model then checks whether the target takes the default form. If so, the adjacent segment is added to the blocker set; otherwise, it is added to the deletion(transparent) set. If a segment appears in both sets, it is removed from the blocker set. This situation would occur when a segment is adjacent to both default and non-default target forms, and its inconsistent behavior indicates transparency.

The tier is then updated as the complement of the deletion set, and the rule condition is defined as the tier with blockers excluded, though blockers themselves remain on the tier. This updated tier is used for the next round of tier projection and rule condition updates. The process repeats until neither the tier nor the rule condition changes further.

Additionally, there is a "reset" checkpoint, depending on whether the default rule is updated. Before the tier and condition update, if the current default rule differs from that of the previous iteration, the tier and condition are reset to their initial states. This reset is necessary because an incorrect default value could exclude true blockers from the tier, which is an error that cannot be corrected through further iteration.

3 Evaluation

3.1 *Data* The testing data is a right-spreading vowel harmony toy dataset designed based on the Khalkha Mongolian vowel harmony pattern, as demonstrated in (1). In this dataset, the vowel inventory is the same as Khalkha Mongolian (Nevins, 2010), which contains height, backness (rounding) and ATR constrasts, including three high vowels /i, u, ʊ/, two mid vowels /e, o/ and two low vowel /a, ɔ/[4]. Only backness (rounding) harmony is considered, and it only applies to mid vowels. Therefore, the two alternating segments, i.e., the targets, are /e/ and /o/, which alternate between [±back]([±round]). Only non-high vowels, /e, o, a, ɔ/, can be the triggers of vowel harmony. The high front vowel /i/ is transparent, which is invisible to vowel harmony, and the two high back vowels /u, ʊ/ are blockers, which always make the target take the default form /e/. The summary of the harmony pattern is demonstrated in the Table 1 below.

Vowel type	Trigger	Target	Transparent	Blocker
Vowel	/e, o, a, ɔ/	/e, o/	/i/	/u, ʊ/

Table 1: Vowel harmony pattern of the toy data.

The word items in the data set have the structure "stem (+neutral affix) + target affix". They always have one stem, which contains the trigger, and one word-final affix, which contains the target. There are three types of stems: [-back] and [+back] trigger stems and neutral stems. The number of stems containing the [-back] and [+back] triggers is equal: Three stems each of /e/, /a/, /o/, and /ɔ/ were included. Another type of neutral stems, including one high vowel each (/i/, /u/, and ʊ), is also included. There is only one word-final affix, which is represented as /Mt/, where M represents a mid vowel not specified for [back]. For each stem+affix combination, there are five tokens without an in-between affix, two tokens with a transparent affix (containing the transparent vowel /i/), and two tokens with a blocker affix (containing the blocker vowel /u/ or /ʊ/.

In total, there are 135 UR-SR pairs: 15 (stems; 3*4 non-high vowels + 1 transparent + 2 blocker) * 9 (in-between affix; 5 Ø + 2 transparent + 2 blocker) * 1 (final affix) = 135. Among the pairs, 93 cases realize the target as /e/ (default form) and 42 cases take the /o/ form.

3.2 *Results* For the toy dataset, the original D2L model learned the correct rule type, i.e., assimilation, and the correct alternating features. However, it excluded both high front and high back vowels, i.e., both transparent vowels and blockers, from the tier and thus failed to learn the correct grammar. Additionally, the wrong rule learned by the original model passed the Tolerance Principle by an accuracy of 89%. In contrast, the current model successfully learned the rule by having the high back vowels on the tier but excluding them from the rule condition and reached the accuracy of 100%.

[4] /ɔ/ is considered as a low vowel here, because in Khalkha Mongolian, /ɔ/ alternates with /a/. The difference between /a, ɔ/ and /e, o/ involves the [ATR], in that /a, ɔ/ are [-ATR] and /e, o/ are [+ATR]. For simplification, in the toy data set, we only include the /e, o/ alternation and ignore the [ATR] harmony.

4 Discussion

4.1 *Comparison with the Original D2L model* The results on the toy data set have shown that the original D2L model fails to learn the correct rule, since the tier induction mechanism doesn't distinguish transparent segments and blockers. Both transparent segments and blockers don't assimilate with the target, but the blockers consistently predict the target form, and thus should be maintained on the tier, but excluded from the assimilation rule condition. In contrast, the current model uses an explicit mechanism to detect the blockers and distinguish them from transparent segments, and thus enables separate tier induction and rule condition induction. The blocker detection procedure requires the default form to be specified at the very beginning, but also allows updating of the default form. This mechanism demonstrates a trial-and-error learning process.

In fact, Belth (2024) has tested the D2L model's ability in two potential blocking cases, but neither of them is actually blocking, as in the Khalkha Mongolian data. The first case is the toy data in Finley (2011)'s first experiment, which examined the learning of non-local /s/-/ʃ/ harmony using an artificial language learning paradigm. The training data of this experiment consists of sibilant harmony cases across intervening vowels, i.e., items of CVSV-SV structure. However, in the testing phase, items that have intervening vowels and nonsibilant consonants, i.e., CVSV-SV and SVCV-SV, were both provided. In the second condition, human participants didn't predict the harmony pattern, so the nonsibilant consonants can be considered as some kind of "blockers". The D2L model's learning results matched the human responses. However, this doesn't mean the learner learned the blocking. It is because those nonsibilant blockers never occurred between the trigger and the target that the D2L learner didn't exclude them from the tier.

The second case is the Latin liquid dissimilation, where the "blockers" are in fact more adjacent triggers, and are not inert blockers. In the Latin adjectival -alis/-aris affix, /l/ is considered as the default form of the liquid, and it dissimilates to [r] when preceded by /l/ from a long distance, as in *lun-aris* 'lunar'. This dissimilation can be blocked by intervening /r/ and non-coronal consonants (Cser, 2010), as in *flor-alis* 'floral' and *leg-alis* 'legal'. Belth (2024) set the feature [lateral] as the alternating feature and [l] as the only [+lateral] segment. As a result, the model learned a [lateral] dissimilation rule on the consonant tier, which maintains /r/ and non-coronal consonants on the tier. However, in this case, the intervening /r/ and non-coronal consonants are not actually considered as "blockers", but also triggers of the dissimilation, which dissimilate the targets to [+lateral].

4.2 *Alternative blocker detection mechanisms* Among the works about phonotactic learning, Kim (2022)'s RTL model also offered a solution to detect blockers by adding an extra evaluation step to the existing IPL model. The original IPL model (Gouskova & Gallagher, 2020) uses trigram constraints $^*X[\,]Y$ to identify forbidden substrings, and searches for appropriate tier projection by examining the natural classes that include both X and Y. However, this model cannot identify blockers. Further, the RTL model introduced a blocker detection procedure by replacing the placeholder with each visible segment on the "evaluation" tier, and then checks whether there is a blocker and which segment is the blocker. This mechanism relies on the preset "evaluation" tier, which is either a consonantal or vocalic tier. Since the learner can only get access to the consonant or vowel tiers during evaluation, this model cannot deal with cases where consonant blockers exist in vowel harmony, which has been attested (Rose & Walker, 2011).

In comparison, my current model doesn't rely on any preset tiers to detect the blockers. Rather, it learns the blocking patterns by checking the prediction consistency of the adjacent segments, which is purely data-driven. Therefore, it can handle cases of consonant blockers in vowel harmony.

4.3 *Future directions* One limitation of the current paper is that only toy data has been tested. To extend the model to more realistic data, one major issue is how to distinguish between exceptions and blocking cases. The original D2L adopted the Tolerance Principle (Yang, 2016) to deal with the sparsity and exceptions in naturalistic data, and the incorrect rule learned on the current toy data indeed passed the Tolerance Principle. It is not clear whether this mechanism can distinguish between exceptions and blocking cases in realistic data, since there might not be enough blocking cases and thus could be ignored (Gouskova & Gallagher, 2020). Future research will hopefully expand the current learner to capture more realistic data and test exception-filtering mechanisms on data involving blocking.

Additionally, similar to the baseline D2L model, the current model has only considered (dis)harmony processes. A broader framework should also allow the learning of phonological processes that don't rely on the feature (dis)agreement between the condition and the target, such as deletion, epenthesis, etc.

5 Conclusion

When learning long-distance phonological processes, it is necessary to distinguish between transparent segments and opaque but inert blockers. The former should be excluded from the tier, while the latter should remain in the tier but not be included in the rule context.

One possible approach that could be adopted by learners is to distinguish between blockers and transparent segments by checking whether the non-triggering adjacent segments consistently predict the default target form, and thus distinguishing between blockers and transparent segments. The default form is established and updated via a trial-and-error process, updated together with the tier and rule condition. By adding these procedures to the D2L model (Belth, 2024), the current model successfully learned vowel harmony involving blocking, and thus provides a more comprehensive learning procedure of phonological tiers.

References

Belth, Caleb (2023). Towards a learning-based account of underlying forms: A case study in turkish. Hunter, Tim & Brandon Prickett (eds.), *Proceedings of the Society for Computation in Linguistics 2023*, Association for Computational Linguistics, 332–342, URL `https://aclanthology.org/2023.scil-1.29`.

Belth, Caleb (2024). A learning-based account of phonological tiers. *Linguistic Inquiry* 1–37, URL `https://doi.org/10.1162/ling`$_{a0}$0530.

Clements, G. N. (1987). Toward a substantive theory of feature specification. Paper presented at the UCLA Symposium on Segment Structure, October 1987.

Cser, András (2010). The -alis/-aris allomorphy revisited. Rainer, Franz, Wolfgang U. Dressler, Dieter Kastowsky, Hans Christian Luschützky & Elisabeth Peters (eds.), *Variation and Change in Morphology: Selected Papers from the 13th International Morphology Meeting, Vienna, February 2008*, John Benjamins, Amsterdam, 33–52.

Finley, Sara (2011). The privileged status of locality in consonant harmony. *Journal of Memory and Language* 65:1, 74–83.

Goldsmith, John (1979). *Autosegmental Phonology*. Garland, New York.

Goldsmith, John & Jason Riggle (2012). Information theoretic approaches to phonological structure: The case of finnish vowel harmony. *Natural Language and Linguistic Theory* 30, 859–896, URL `https://doi.org/10.1007/s11049-012-9169-1`.

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

Hayes, Bruce & Colin Wilson (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39:3, 379–440, URL `https://doi.org/10.1162/ling.2008.39.3.379`.

Jardine, Adam & Kevin McMullin (2017). Efficient learning of tier-based strictly k-local languages. Drewes, Frank, Carlos Martín-Vide & Bianca Truthe (eds.), *Language and Automata Theory and Applications*, Springer International Publishing, vol. 10168 of *Lecture Notes in Computer Science*, 64–76.

Kim, Seoyoung (2022). *Restrictive Tier Induction*. Phd thesis, University of Massachusetts Amherst, URL `https://scholarworks.umass.edu/server/api/core/bitstreams/a2905554-06b4-4bd7-8ed9-024`

McCarthy, John (1979). *Formal Problems in Semitic Phonology and Morphology*. Doctoral dissertation, MIT, Cambridge, MA.

Nevins, Andrew (2010). *Locality in Vowel Harmony*. MIT Press, Cambridge, MA.

Rose, Sharon & Rachel Walker (2011). Harmony systems. Goldsmith, John A., Jason Riggle & Alan C. L. Yu (eds.), *The Handbook of Phonological Theory*, John Wiley & Sons, Ltd, 240–290.

Vergnaud, Jean-Roger (1977). Formal properties of phonological rules. Butts, Robert E. & Jaakko Hintikka (eds.), *Basic Problems in Methodology and Linguistics*, Reidel, Dordrecht, 299–318.

Yang, Charles D. (2016). *The Price of Linguistic Productivity: How Children Learn to Break the Rules of Language*. The MIT Press.