The Acquisition of Vowel Harmony from Simple Local Statistics

Spencer Caplan (spcaplan@sas.upenn.edu)
Department of Linguistics, University of Pennsylvania

Jordan Kodner (jkodner@sas.upenn.edu)
Dept. of Linguistics, Dept. of Computer and Information Science, University of Pennsylvania

Abstract

Vowel harmony denotes a class of phonotactic constraints which limit which vowels can co-occur in words. The characteristics of harmony systems have been well-researched from theoretical, typological, and developmental perspectives. Children are sensitive to harmony very early in their development, as young as seven months, so the mechanisms responsible for harmony acquisition must be able to identify its presence as well as the specifics of individual vowel harmony systems with little input. Prior computational work has sought either to detect the presence of harmony without describing the specific implementation or to describe a specific implementation when the general details are known beforehand. We present a new computational acquisition approach inspired by phonological notions of restrictiveness which succeeds in automatically detecting harmony in some language and describes the gross characteristics of the underlying harmony grammar without prior knowledge about the type of system to expect.

Keywords: linguistics; language acquisition; phonology; computational modeling; vowel harmony;

Introduction

Vowel harmony describes a broad class of phonotactic constraints which partition a language's vowels into two or more classes. Words are restricted from containing vowels from multiple classes regardless of the intervening consonants (Bakovic, 2000). A range of geographically and typologically diverse languages exhibit vowel harmony. It is commonly associated with the Uralic (Finnish, Hungarian, Nenets, etc.), Turkic (Turkish, Uyghur, etc.), and Mongolic (Mongolian, Buryat, etc.) families, among many others (Kaun, 2004), and has occasionally been described for Bantu (kiKongo, Shona, etc.) as well (Beckman, 1997).

Vowel harmony is often identified by the presence of morphological alternations. For example in Turkish (1), the vowel in the plural ending changes according to the vowels within the root that it attaches to. If the root contains front vowels, then the suffix contains an *e*, but if the root contains back vowels, then the suffix contains a corresponding *a*. The roots themselves are subject to the harmony constraint, so they never contain a mix of front and back vowels either.

(1) a. Göz-**ler** ve kulak-**lar**
eye-PL and ear-PL
'Eyes and ears.'

b. kedi-**ler** ve boncuk-**lar**
cat-PL and bead-PL
'Cats and beads'

Canonical vowel harmony applies across all morphemes in whole words. Sometimes, as in Finnish, the results are dramatic. The vowels /i/ and /e/ do not participate in vowel harmony in Finnish, so the harmony in (2) must spread long-distance over these vowels. The opacity of such potentially long distance alternations (Kaun, 2004; Finley, 2009) seems to pose a challenge for learners, yet children do learn them reliably. A large number of laboratory experiments demonstrate that infants as young as seven months are sensitive to vowel harmony alternations (Mintz, Walker, Welday, & Kidd, 2018; Van Kampen, Parmaksiz, Vijver, & Höhle, 2008).

(2) a. kumarreksituteskenteleentuvaisehkolla-ismaisekkuudellisenneskenteluttele**matto**-mammuuksissan**sakaankopahan**

b. epäjärjestelmällistyttä**mättö**myy-dellän**säkäänköhän**

While the phonological representation (Bakovic, 2000; Krämer, 2003) and typological distribution (Aoki, 1968; Kaun, 2004) of vowel harmony are well-researched, an explicit model of acquisition has thus far been lacking. Here we present a computational model for the acquisition of vowel harmony which accounts for a range of known theoretical and experimental facts. It relies on pointwise mutual information (PMI) between pairs of tier-adjacent vowels (i.e., vowels that are adjacent when consonants are disregarded) to correctly identify the presence or absence of vowel harmony (including secondary harmony), to determine which vowels participate in harmony, and to properly categorize those vowels into their harmonizing classes, showing that these broad characteristics of harmony systems can be discovered by a learner with minimal information.

That our learning model operates over local co-occurrence statistics on the vowel tier is consistent with work demonstrating the sufficiency of such 'Tier-based Strictly Local' representations of phonological constraints (Heinz, Rawal, & Tanner, 2011). This restricted computational power is desirable, so as not to over-generate posited generalization—a common problem for the application of many modern statistical learning methods to questions of language acquisition. In the next sections, we review the phonological facts of vowel harmony systems and discuss previous computational approaches for describing their acquisition, and then we introduce the present model and two computational modeling experiments which offer support for our approach. We conclude with a discussion of implications and directions for future work.

Phonological Theories of Harmony

Phonotactic restrictions are *local*. The patterns of constraints that dictate whether a phone can possibly occur in a word

(a) Finnish

(b) Turkish

Figure 1: Venn diagrams showing basic vowel harmony systems of Finnish (left) and Turkish (right). All Turkish vowels harmonize, but the Finnish system includes two 'neutral' vowels which do not respond to or influence vowel harmony.

only operate on that phone and its neighbors. For instance in English, obstruents occupying the same syllable coda must share the same value for voicing. That is why 'cats' is pronounced /kæ**ts**/, but 'dogs' is pronounced /dɔ**gz**/. The *-s* is voiced in the latter but not in the former in order to agree with the preceding obstruent. This pattern is local because it only operates over a fixed distance k from the phone in question ($k = 1$ here). When another phone intervenes between the obstruents, as in /bɑ**k**sɪ**z**/ 'boxes,' the constraint does not apply. That is not to say, however, that all phonotactics apply locally on the surface. They can apply locally on *tiers* as well. For example, vowel harmony applies between vowels and their nearest vowel neighbors regardless of how many intervening consonants there are. These interactions are still local over a fixed k distance because nearest vowels on the surface are adjacent vowels on the vowel-tier.

Like the English example, vowel harmony requires that the relevant phones share some phonological feature, e.g., frontness, height, roundness, ATR (Aoki, 1968). In Finnish for instance, there is a harmony alternation between -FRONT and +FRONT vowels as illustrated in Table 1. For example, in a simple /a,i,u,e,o/ system with hi/non-hi harmony, words can contain only high vowels /i,u/ or non-hi vowels /a,e,o/, but not both. This means that harmonizing sets are not arbitrary. Cross-linguistically, attested vowel harmony alternations are over phonologically defined natural classes.

Finnish	English	Case
pöytä-**nä**	*'table'*	essive
pouta-**na**	*'fine wine'*	essive
koti-**na**	*'home'*	essive
käde-**llä**	*'hand'*	adessive
kesy-**llä**	*'tame'*	adessive
vero-**lla**	*'tax'*	adessive

Table 1: Finnish suffix allomorphy caused by vowel harmony. Data adapted from Baker (2009)

In the simple case, vowel harmony applies to entire words. *All* vowels in a word must share some phonological feature in common. Keeping the ideas of tier-adjacency in mind, it is possible to describe this whole word constraint as a local process: vowel harmony just needs to spread from vowel to adjacent vowel to adjacent vowel in order to express this 'long-distance' process. But this is an over-simplification. Some languages have 'neutral' vowels which do not participate in the harmony process.

In some languages, these neutral vowels block harmony from spreading across them. They are 'opaque' because they sit in the way of the spread of harmony on the vowel-tier and break tier-adjacency. For instance in Fula, ATR harmony normally prevents vowels like /i/ and /ɔ/ from appearing in the same word., but this spreading is blocked by the opaque vowel /a/ (Gafos & Dye, 2011). This is highlighted by the alternation in 3a and 3b, compared with the co-occurrence of /i/ and /ɔ/ when an opaque vowel intervenes as in 3c. In other languages, neutral vowels are said to be 'transparent' because they let harmony pass through them. This can still be expressed locally if the transparent vowels keep track of which harmony class their neighbors are in and spread that feature but then surface without it. For instance, in Finnish there are two 'neutral' vowels /i/ and /e/, which can surface in words with either +BACK and +FRONT vowels. Figure 1 presents a visual representation of the Finnish system with neutral vowels.

(3) a. pɛɛc-ɔn
crack-DIM.PL
'Little cracks'

b. peec-i
crack-PL
'Cracks'

c. bɔɔt-aa-ri
dinner
'Dinner'

In addition to neutral vowels, harmony systems are complicated by the presence of multiple harmony alternations. For example, Turkish not only exhibits primary frontness harmony as in Figure 1, but also a secondary alternation between round and non-round vowels.

Every harmony pattern discussed so far can be expressed locally even when surface alternations may appear over long distances. In fact, it has been argued that all harmony systems can be expressed locally (Heinz & Lai, 2013). This is backed up by typological evidence. So far, none of the many conceivable non-local harmony systems has been attested. For example, one might imagine a 'sour-grapes' system in which a feature only spread at all if every vowel in a guaranteed to harmonize. e.g. if some opaque vowel exists in a word then would-be feature spreading does not occur.

The ideas expressed here imply a certain kind of acquisition model. Vowel harmony spreads vowel-to-vowel to tier-adjacent neighbors, therefore a learner's grammar must be able to operate on that kind of mechanism. A learning model that operates only on tier-adjacent vowels should only be able to identify precisely those kinds of patterns and not over-predict unattested systems. This might be contrasted with a more mathematically powerful model which would need subsequent constraints so as to match the typologically attested distribution of harmony systems. The model we describe here is sufficient for capturing the types of harmony described here. While these are not all the possible harmony systems that exist, these capture the broad patterns that can be learned without specific reference to morphology or a deeper understanding of a language's phonology.

Previous Approaches

Prior computational work on vowel harmony has sought to describe harmony systems or to model the acquisition process when the presence and type of harmony was known beforehand. Harrison, Thomforde, and O'Keefe (2004); Sanders and Harrison (2012) calculate quantitative metrics of harmony over a language. However such a metric for how "harmonic" a language is does not provide an explicit model of acquisition. Similar descriptions and general quantitative metrics of harmony are addressed in Mayer, Rohrdantz, Butt, Plank, and Keim (2010) and Szabó and Cöltekin (2013), but they too are meant as tools for phonological analysis rather than cognitive learning models.

Previous acquisition models, on the other hand, are able to describe harmony systems but need to know up front whether there is harmony in the input and how many levels to represent. The most successful of these (Baker, 2009), uses a two-state Hidden Markov Model (HMM) to characterize harmony status of Turkish, Finnish, with non-harmonizing English and Italian as controls. The approach in Baker (2009) (which builds on Goldsmith and Xanthos (2009); Goldsmith and Riggle (2012)) is restricting, however, in that it requires the researcher to decide upfront how many levels of harmony to search for, with the wrong parameter choice leading in incorrect output. Secondary harmony can only be processed if the HMM is setup with four initial states rather than two. The paper also discusses how a class of models using mutual information (MI) and Boltzmann fields can accurately identify vowel-to-vowel interactions but does not provide an explicit means for describing vowel harmony given the results.

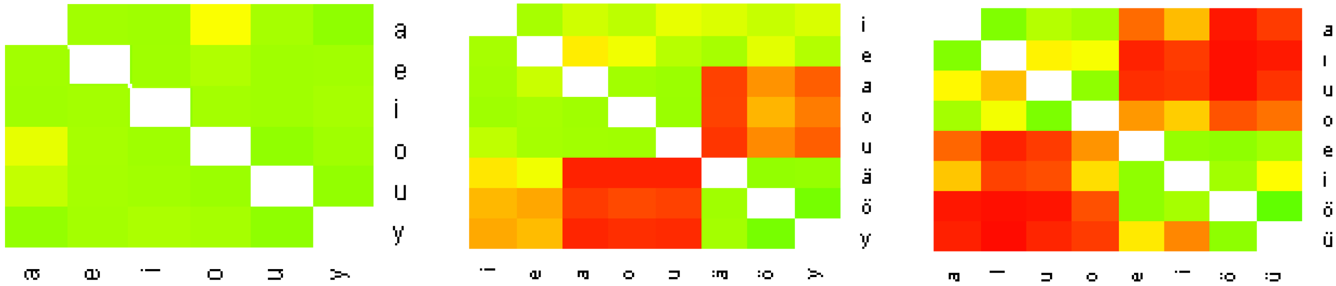
The input data for all previous models has been in the form of word lists. This is important given experimental findings that infants are sensitive to vowel harmony alternations at as young as six or seven months (Mintz et al., 2018). This stage of development predates infants' ability to segment continuous speech into words (Saffran, Aslin, & Newport, 1996), and in fact may serve as a cue to aid in doing so. With this in mind, a full model of vowel harmony acquisition should be able to function taking in either word lists or unsegmented utterances as input. Our acquisition model improves upon previous work by combining the task of detecting harmony with the problem of describing the underlying grammar responsible for harmony. Furthermore, it does away with complex statistical methods in favor of simple local calculations which are amenable to online implementation.

Current Model

The current model takes a stream of segments as input and produces a characterization of the input's harmony system as output. Due to technical limitations, the model was run on standard orthographies instead of phonemic transcriptions. For most languages tested, the orthography maps reasonably well to phonemes, so this may be viewed as a test of robustness rather than a hindrance since an orthographic record necessarily includes some degree of noise or error compared to direct phonemic transcription. As the model performs well over orthographic input, we have confidence that the results generalize to less noisy input schemes.

The model does not have a means of distinguishing between consonants and vowels on its own at this point, so the input must be annotated to indicate which characters are vowels. They may be annotated with a basic set of phonological features as well. The notion that infant learners separate phonological input into distinct consonant and vowel tiers is well established within language acquisition and phonological processing literature (Maye, Werker, & Gerken, 2002; Saffran, Newport, & Aslin, 1996; Aslin, Woodward, LaMendola, & Bever, 1996; Saffran, Aslin, & Newport, 1996; Newport & Aslin, 2004), and all major phonological theories include some notion of basic features. These features are used to pair up harmonizing vowels across harmony sets, but they are not necessary for identifying harmony or the sets themselves, so the model is agnostic to the particular feature values posited for any individual language.

Given the input stream, segments are automatically separated into vowel and consonant tiers, and counts of tier-adjacent vowels are calculated. For example, in the Finnish word *Kalevala*, the tier-adjacent pairs are 'a-e', 'e-a', and 'a-a'. Once the co-occurrences are tabulated, it is possible to compute a simple metric related to *pointwise mutual information* (PMI) between vowels as shown in Eq. 1. The way PMI is applied here has a fairly intuitive interpretation. We would like an estimate how likely two vowels are to co-occur. However, a simple conditional probability is insufficient since



(a) English (b) Finnish (c) Turkish

Figure 2: Heat-maps showing PMI between vowels in English (no harmony), Finnish (harmony with neutral vowels *e* and *i*, and Turkish (harmony for all vowels). Green indicates low PMI (high co-occurance) while red corresponds to a high PMI (low co-occurrence).

it is heavily biased by frequency, so we divide by the probability of the vowel to normalize for that. This is simpler than true PMI as the model does not include any logarithmic transformations. This metric is something which an infant can, in principle, compute. It is simply the ratio between transitional probabilities and raw probabilities—both of which have been experimentally demonstrated to be computed and utilized by infants (Saffran, Newport, & Aslin, 1996; Aslin, Saffran, & Newport, 1998; Pelucchi, Hay, & Saffran, 2009).

$$\mathrm{PMI}(V_1,V_2) = \frac{P(V_1|V_2)}{P(V_1)} \tag{1}$$

This process yields a co-occurrence vector for every vowel. In the absence any phonotactic restrictions, we expect each vector to be fairly uniform. That is, each vowel is expected to co-occur with every other vowel at a more or less uniform rate once the frequencies of both vowels are accounted for. This is precisely the pattern observed in non-harmony languages like English. However, with vowel harmony restricting co-occurrences, those vowels that participate in the alternation should rarely if ever occur with members of the opposite harmony set, yielding skewed distributions. Neutral vowels, on the other hand, should not show this asymmetry. The heatmaps in Figure 2 show a graphical representation of such patterns as captured by the model.

The PMI calculation reveals many characteristics of the system. First, there is no harmony if each vowel has a near-uniform distribution, otherwise there is some kind of harmony. Second, vowels with highly skewed distributions can be contrasted with those with near-uniform distributions in order to separate neutral vowel from harmonizing vowels. The remaining vowels can be grouped according to their PMI vectors to determine to their class memberships.

The model groups the harmonizing vowels into their primary sets by a *k*-means clustering into two groups, but any number of suitable alternatives could achieve similar results. *k*-Means is desirable because it is fast to compute and can be performed online. The typical concern with *k*-means clustering is that the number of clusters needs to be known beforehand, but in the problem at hand, *k* is always 2.

Primary harmonizing classes can be determined without imposing a feature system on the data. The features come into play for identifying secondary harmony and mapping vowels with their counterparts across sets. Vowel distinctions are collapsed over the identified harmonizing feature—i.e. if $\pm$FRONT harmony were identified, then that feature would be removed and vowels with identical features treated as identical. The pairs of vowels that collapse together most closely are those that primary harmony alternates with one another. Then the co-occurrence tabulation is re-run to find skewed distributions that were masked during the first pass.

Evaluation

A series of experiments were run to evaluate the performance of the model. The eight test languages included with five with productive vowel harmony (Finnish, Hungarian, Turkish, Uyghur, Warlpiri), two with secondary harmony (Turkish and Hungarian), two with no harmony (English, German), and one which had vowel harmony in the recent past but no longer has the productive process (Estonian). Of these, Turkish and Uyghur have no neutral vowels in their primary systems, while Finnish, Hungarian, and Warlpiri each have at least one. English and German serve as a control. It is important to confirm that the harmony detection results actually depend on the input. Input data were from MorphoChallenge (Kurimo, Virpioja, Turunen, & Lagus, 2010) when available. The Uyghur and Hungarian were provided for the DARPA LORELEI project, and the Warlpiri data is from (Swartz, 1997). Table 2 summarizes the harmony systems of each input language. The corpora varied considerably in how many harmony violations they have. Turkish was the worst with about a third of words containing at least one violation.

Experiments were performed using two different preparations of the respective data sets. In the first test, the model was fed segmented wordlists from each language with frequency information tabulated from the corpora with extremely infrequent words discarded (those representing less than 0.00001% of the total corpus). The second test, run

Language	Primary	Secondary	Neutral Vowels	Model Correct
Hungarian	frontness	rounding	2/7	Primary only
Turkish	frontness	rounding	(rounding only)	✔
Finnish	frontness	–	2/8	✔
Uyghur	frontness	–	0	✔
Warlpiri	frontness	–	1/3	✔
Estonian	(remnant)	–	–	(✔)
German	–	–	–	✔
English	–	–	–	✔

Table 2: Performance on various test languages. Middle columns indicate the ground-truth system for each language. Checkmarks in the right column indicate that the system was learned exactly. Estonian is indicated in (blue), as the predicted harmony for Estonian learning infants has not yet been tested empirically. Vowel length was removed from Hungarian orthography.

on Turkish, Finnish, Walpiri, English, and German, evaluated the model on whole unsegmented utterances rather than wordlists. This is a more difficult evaluation because, without access to the boundaries between words, the distributional cues required for learning are subject to a much higher degree of noice. This test of robustness was meant to more closely capture the type of input available to infants who are still solving word segmentation for themselves. Performance of the model was the same in both cases: all known productive vowel harmony languages were identified as such, and their vowels properly categorized into the appropriate harmonizing classes, including the Turkish secondary classes. No harmony system was detected for either English or German.

Estonian was included because it lacks *productive* vowel harmony but had it in the past (Harms, 1962). One might expect the language to show the residual fingerprint of its harmony system. That is, the language should show the distributional asymmetries that are typical of its formal vowel harmony system, but that signal should be degraded since it is no longer actively maintained in speakers' grammars. Our learning model discovers exactly such a fingerprint: a tendency towards frontness harmony that is much weaker than its close relative, Finnish, but much stronger than either English or German. Though it is obvious that adult Estonian speakers' grammars do not obey vowel harmony, it is unclear how children react to this signal. Given the experimental finding that infants are sensitive to vowel harmony alternations before word segmentation (Mintz et al., 2018), and crucially, that this distributional sensitivity is able to be elicited even for infants primarily exposed to a non-harmony language such as English, it is conceivable that they do pick it up at some early stage in learning before dropping it later. Experimental evidence on young Estonian learners would be necessary to elucidate this.

A similar note should be made of the model's failure to identify secondary harmony in Hungarian. The phonological facts of Hungarian harmony interactions are complex (Törkenczy, 2016), and it is unknown the stages of acquisition that Hungarian learners may pass through. Without a fuller accounting of early Hungarian phonological acquisition, e.g., whether secondary harmony learned as early as primary harmony or if it relies on later lexical or morphological knowledge, it is unclear whether the model's performance in this case is correct or in error.

Discussion

The structure required to fully describe a harmony system within theoretical phonology requires a number of abstractions on top of acoustic data. Nonetheless, children are adept learners of such complex systems, and experimental evidences confirms that they are sensitive to harmony in broad strokes at an early age. In this paper, we introduced an explicit model of vowel harmony acquisition which is able to account for this process relying only on computationally simple, cognitively plausible tools. This is an improvement over previous quantitative metrics of harmony which either lack an explicit learning model, or require additional parameter specification from the researcher. While the present model does not provide a full specification of a harmony grammar—it does not distinguish opaque from neutral vowels, or identify the direction of spreading, etc.—such early identification of harmony broadly matches the empirical findings on infant sensitivity to harmonic input. It is not clear at what stage of development more rich harmony representation is acquired.

While the experimental results of the model on the majority of test languages were clear, the case of Estonian is worth noting. From one perspective, the model fails to correctly identify the lack of a productive harmony process in modern Estonian. On the other hand, it does uncover a fact about the language. It had productive harmony in the past, and it apparently retains the residual effects of harmony in the present. In this way, the model has something to say about adult linguistic representation.

We also draw attention to the connection between typological facts and models of acquisition. Our acquisition model treats the learning of multiple harmony processes as a fundamentally sequential process. Aksenova and Deshmukh (2018) show that within the typology of multiple harmony systems cross-linguistically, the only attested relations between vowel sets are either disjoint or subset/superset. The only unattested configuration of multiple vowel harmony is one in which harmonizing classes are partially overlapping. Such an overlapping multiple harmony system would require *parallel* rather than *sequential* acquisition. While it requires more thorough investigation, it is promising that typological generalizations in vowel harmony are consistent with independent predictions of an acquisition model. This account of vowel harmony acquisition is intrinsically tied to notions of phonological restrictiveness. The process of tier-adjacent tabulations means that models of this class can formally capture exactly those, and only those, patterns that are consistent with Tier-based Strictly Local formal grammars without any sort of post-hoc or superficial restrictions.

Acknowledgments

We thank the University of Pennsylvania's 2016 LORELEI summer research team for their helpful insights as well as audiences at NECPhon 10 and the Penn Language Development and Language Processing Lab. This research was funded by the DARPA LORELEI program under Agreement No. HR0011-15-2-0023.

References

Aksenova, A., & Deshmukh, S. (2018). Formal restrictions on multiple tiers. *Proceedings of the Society for Computation in Linguistics (SCiL) 2018*, 64–73.

Aoki, H. (1968). Toward a typology of vowel harmony. *International Journal of American Linguistics*, *34*(2), 142–145.

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological science*, *9*(4), 321–324.

Aslin, R. N., Woodward, J. Z., LaMendola, N. P., & Bever, T. G. (1996). Models of word segmentation in fluent maternal speech to infants. *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*, 117–134.

Baker, A. C. (2009). *Two statistical approaches to finding vowel harmony* (Tech. Rep.). Tech. rep., University of Chicago.

Bakovic, E. (2000). *Harmony, dominance and control*. Unpublished doctoral dissertation.

Beckman, J. N. (1997). Positional faithfulness, positional neutralisation and shona vowel harmony. *Phonology*, *14*(1), 1–46.

Finley, S. (2009). *Formal and cognitive restrictions on vowel harmony*. The Johns Hopkins University.

Gafos, A., & Dye, A. (2011). Vowel harmony: transparent and opaque vowels. *The Blackwell companion to phonology*, *4*, 2164–2189.

Goldsmith, J., & Riggle, J. (2012). Information theoretic approaches to phonological structure: the case of finnish vowel harmony. *Natural Language & Linguistic Theory*, *30*(3), 859–896.

Goldsmith, J., & Xanthos, A. (2009). Learning phonological categories. *Language*, *85*(1), 4–38.

Harms, R. T. (1962). *Estonian grammar*. Routledge.

Harrison, D., Thomforde, E., & O'Keefe, M. (2004). *The vowel harmony calculator.*

Heinz, J., & Lai, R. (2013). Vowel harmony and subsequentiality. In *Proceedings of the 13th meeting on the mathematics of language (mol 13)* (pp. 52–63).

Heinz, J., Rawal, C., & Tanner, H. G. (2011). Tier-based strictly local constraints for phonology. In *Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies: short papers-volume 2* (pp. 58–64).

Kaun, A. (2004). The typology of rounding harmony. *Phonetically based phonology*, *87416*.

Krämer, M. (2003). *Vowel harmony and correspondence theory* (Vol. 66). Walter de Gruyter.

Kurimo, M., Virpioja, S., Turunen, V., & Lagus, K. (2010). Morpho challenge competition 2005–2010: evaluations and results. In *Proceedings of the 11th meeting of the acl special interest group on computational morphology and phonology* (pp. 87–95).

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3), B101–B111.

Mayer, T., Rohrdantz, C., Butt, M., Plank, F., & Keim, D. (2010). Visualizing vowel harmony. *Linguistic issues in language technology*, *4*(2), 1–33.

Mintz, T. H., Walker, R. L., Welday, A., & Kidd, C. (2018). Infants' sensitivity to vowel harmony and its role in segmenting speech. *Cognition*, *171*, 95–107.

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance i. statistical learning of non-adjacent dependencies. *Cognitive psychology*, *48*(2), 127–162.

Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Learning in reverse: Eight-month-old infants track backward transitional probabilities. *Cognition*, *113*(2), 244–247.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of memory and language*, *35*(4), 606–621.

Sanders, N., & Harrison, K. D. (2012). Discovering new vowel harmony patterns using a pairwise statistical model. In *Poster presented at the 20th manchester phonology meeting, university of manchester.*

Swartz, S. (1997). *Warlpiri yimi kuja karlipa wangka*. Summer Institute of Linguistics, Australian Aborigines and Islanders Branch, Warlpiri Translation Project.

Szabó, L., & Cöltekin, C. (2013). A linear model for exploring types of vowel harmony.

Törkenczy, M. (2016). *Hungarian vowel harmony*. Oxford University Press.

Van Kampen, A., Parmaksiz, G., Vijver, R. van de, & Höhle, B. (2008). Metrical and statistical cues for word segmentation: The use of vowel harmony and word stress as cues to word boundaries by 6-and 9-month-old turkish learners. *Language acquisition and development: Proceedings of GALA 2007*, *2007*, 313–324.