Chapter 35
Data-driven approaches

Rebecca Knowles and Nathan Sanders

35.1 Introduction

Data-driven approaches are a natural fit for identifying and quantifying vowel harmony. As vowel harmony is a recurrent pattern with a relatively limited range of variability (potential size of the harmonic domain, number of potential harmonic features, number of potential targets and triggers, etc.), it is amenable to computational techniques commonly used for pattern recognition. However, vowel harmony is often not strictly categorical; some vowels and/or words behave exceptionally, so statistical approaches can more accurately capture a language's tendency towards harmony than is possible with deterministic rules. Data-driven approaches to vowel harmony can be used for tasks like generating hypotheses about the presence of harmony in a language where it was not previously known to exist (though most existing work is proof of concept, tested on known harmonic systems), modelling child language acquisition of harmony systems, or even measuring harmony for comparison across languages or within one language diachronically.

Statistical and computational approaches to vowel harmony vary widely in both their end goals and the techniques they employ (see Chapters 22 and 34, this volume, for discussion of computational methods). Nevertheless, the approaches we discuss in this chapter can all be described generally as consisting of three main components: input, algorithm, and output. The *input* to the model is the data itself, optionally including some form of supervision (such as labeled data or external knowledge). Inputs could be any combination of lexical databases, digitized writings, transcriptions, etc. The *algorithm* is then applied to the input. Typically, this involves computing statistics from the data, which may be examined directly or iteratively used to fit parameters of some model. Finally, the *output* may consist of a measurement (for example, the magnitude of harmony in a language, as in Sanders & Harrison 2012), a clustering of units (such as clusters of harmonic vowels, as in Kodner et al. 2017), a visualization (heat maps, vowel cluster plots, etc.; see Section 35.4 for discussion and examples), and/or other human-interpretable information about the data.

As the starting point for data-driven approaches, the nature of the input is crucial. Different types of inputs pose different challenges, and must be taken into account when designing algorithms and interpreting outputs. We discuss this in Section 35.2. Different models and approaches allow for different representations of vowel harmony systems, so the choice of model has an impact on what is learned and/or missed. Thus, when building a statistical model of vowel harmony, a researcher must not only take into account the nature of the data itself, but also any relevant linguistic principles that may underlie the data. We discuss some important linguistic considerations for model selection in Section 35.3, describing various models along the way. In Section 35.4, we describe ways that models of vowel harmony can visualize their output. We conclude in Section 35.5 with a summary of major successes and challenges of data-driven approaches to vowel harmony, as well as future avenues worth exploring.

35.2 Types of data sources

The choice of data source is often constrained by what is available to the researcher. For example, there may be only limited data available, as for so-called "low-density" or "low-resource" languages (those with little to no online texts, especially common for Indigenous and minority languages), which may be difficult to acquire and/or insufficient to provide reliable results. Further, data that is available may have been collected and preprocessed in a variety of ways that can have a profound impact on the output of any statistical or computational approach. In this section, we discuss some considerations that must be taken into account based on the type of text data source being used as input.[1]

The issues of focus in this section are transcription versus orthography, analysis of harmony across types versus tokens, and anomalous word types. These are often overlooked or dismissed as "preprocessing", but in fact a careful understanding of these issues and how to handle them is necessary for the design of the model and interpretation of its output. This is especially important for approaches that seek to compare vowel harmony across languages, since this requires controlling for differences in datasets between languages.

35.2.1 Transcription versus orthography

Text data for computational study of vowel harmony can be broadly classified into two forms: *transcription* (a symbolic representation of pronunciation with a system like the International Phonetic Alphabet; IPA) and *orthography* (the language's writing system). While these two forms are often related, the relationship can vary, from logograms (with little or no explicit phonological information, as with Chinese hanzi and Japanese kanji), to abjads (ordinarily indicating consonants and only some vowels, as with Arabic and Hebrew scripts), to opaque relationships between writing systems and pronunciation (e.g., due to sound change over time), to languages with near one-to-one mappings between transcription and orthography (such as Finnish).

Since vowel harmony is a phonological phenomenon and transcription provides accurate information about the sounds of a language, transcription would ordinarily be an ideal choice. For a language without an orthography, transcriptions may be the only text data available. However, transcriptions are not without drawbacks. While the IPA is an international standard, many different traditions and conventions exist (Americanist, Teuthonista, pinyin, etc.). Just as transcription practices vary, so may character encodings. Thus, it can be challenging to combine different datasets for one language, let alone compare results across languages, since a variety of transcription systems, levels of detail, and/or character encodings may need to be recognized and accommodated. Finally, transcriptions are time-consuming and expensive to produce; even for fairly broad transcription, it can take trained professionals 15–40 minutes to transcribe a single minute of audio (Demuynck et al. 2002). Creating a sufficiently large dataset for robust analysis requires a combination of time, money, and expertise that may not be available.

[1] While audio data may be a more direct data source for the study of vowel harmony, this chapter and works cited herein focus only on text data. Analysis of audio data faces significant hurdles (such as segmentation and segment classification) largely orthogonal to the issue of vowel harmony.

Because of these drawbacks, texts written in a language's existing orthography are an attractive option. High-density languages already have enormous amounts of text corpora readily available, allowing researchers to pull millions of words from online databases in a matter of seconds, saving significant time and money as compared to audio transcription. Such databases often include corpora like the New Testament (which contains hundreds of thousands of words), commonly used for comparative work due to the large number of languages into which it has been translated. However, in order to be useful for analyzing vowel harmony, the orthography must be relatively transparent, with a one-to-one or nearly one-to-one correspondence with pronunciation. Without this, any conclusions drawn may not accurately reflect the phonology of the language. Even languages with relatively consistent writing systems, such as Spanish and Turkish, may still have unpredictable differences between spelling and pronunciation. Additional challenges for using orthographic data may arise if a language has multiple writing systems and/or an inconsistent writing system, which can introduce noise into the results and require extra steps to clean up the data.

Another option is to approximate a transcription from an orthographic form. For example, Szabó & Çöltekin (2013) convert text data to the IPA using text-to-speech systems. This provides a way to bridge the gap between large data and phonetic data, but it also requires a text-to-speech system for the language of interest which would need to be built if one does not exist.

35.2.2 Types versus tokens

Given a dataset, a critical preprocessing decision is whether to use types or tokens. A *type* is a member of the set of all unique words in a text, while a *token* is any individual instance of a type and may be repeated. Where the only data available is a word list or dictionary, this decision is moot, since the lack of repetition renders type and token equivalent; nevertheless, the researcher must consider whether the word list contains a sufficient range of examples to adequately demonstrate vowel harmony (or the lack thereof).

The choice between types and tokens requires consideration of what it means to quantify vowel harmony in a language. Using tokens to derive statistics places more weight on frequent words; we expect this to be informative about the level of vowel harmony in the text as a whole. Using types places equal weight on frequent and infrequent words, so that numerous rare inflected forms that demonstrate harmony may outweigh a handful of common disharmonic short words; this may tell us more about an underlying harmony which would not be as apparent from a full text. While these two approaches could lead to different conclusions, Szabó & Çöltekin (2013) and Knowles (2012) experiment with both types and tokens, and find similar results between the two.[2] Depending on the corpus, type- or token-based approaches may over- or under-emphasize anomalous words, such as proper nouns (see Section 35.2.3).

[2]Szabó & Çöltekin (2013) experiment with Hungarian and Turkish, with Dutch and English as non-harmonic control languages, while Knowles (2012) uses Finnish, Turkish, Tuvan, and Swahili, with Japanese and Indonesian as non-harmonic controls. We describe both approaches in Section 35.3.1.

35.2.3 Anomalous words

Datasets can contain anomalous words that could influence harmony measurements in ways that may or may not be desirable. First, as a preprocessing data cleaning step, many approaches remove words that contain numeric or certain non-alphanumeric characters, since these may not represent genuine words in the language and, thus, would fall outside the scope of vowel harmony anyway.

Some approaches also remove words containing only one vowel (e.g., Harrison et al. 2004; Szabó & Çöltekin 2013). This may be done to prevent them from skewing overall vowel counts (which are typically used as a baseline to measure vowel harmony) or because the model specifically performs pair-wise measures, though others keep them in precisely to take them into account when calculating baseline vowel counts. Knowles (2012) compares models with and without monosyllabic words, and finds minimal differences; like the type/token distinction, this appears to be a theoretically important distinction that may not always have major practical implications.

Proper nouns can also be a source of anomalies, particularly if they do not have their origin in the language itself. The Bible is commonly used as a corpus for data-driven approaches to computational linguistics and natural language processing due to its many translations, but it is replete with proper nouns. When those proper nouns are left unmodified or simply transliterated during the translation process, they may fail to exhibit the language's harmony system, potentially obscuring its measurement if they are sufficiently frequent.

Loanwords are another potential source of anomalies. They may be borrowed without being fully adapted to the phonology of the recipient language, retaining the donor language's vowel harmony (or lack of harmony), potentially making the recipient language appear more (or less) harmonic than it really is.

These types of anomalous words can be removed from the corpus, though this can be difficult to do precisely, especially in the case of loanwords, where the history of the individual words may not be known. Their influence can instead be controlled through the decision to use types rather than tokens, which would reduce the impact of frequent anomalous words. Alternatively, these words can be simply left in the data, with the understanding that they are part of the language, and thus, should be included in any calculation of vowel harmony.

35.3 Linguistic considerations

35.3.1 Morphological structure

Vowel harmony is often discussed at the word level (Chapter 17, this volume), and as such, most data-driven approaches to vowel harmony consider harmony at the whole word level, defining the word through whitespace (and punctuation) tokenization (see Chapter 19, this volume, for other ways that morphological structure can be relevant). The choice of boundaries naturally constrains the harmony patterns that can be detected algorithmically. For example, Karajá (a Macro-Jê language spoken in Brazil) exhibits vowel harmony across word boundaries (Ribeiro 2002), and in Chamorro (an Austronesian language spoken in Guam and the Mariana Islands), vowel fronting is triggered by certain particles that are written as separate orthographic words (as in *i gima'* 'the house'; cf. *guma'* 'house'; Topping 1968); this harmony pattern would not be captured by an approach that looked for harmony only within orthographic words (Mayer et al. 2010). See Chapter 20, this volume, for discussion of harmony across word boundaries.

Many approaches ignore consonants and consider only vowels, treating vowel harmony as a local process between tier-adjacent vowels rather than a long-distance phenomenon. Harrison et al. (2004) takes a supervised approach to measuring vowel harmony, taking as input a corpus (represented as whole words with consonants ignored) and known harmony patterns and producing as output a measure of how harmonic the language is (relative to a corpus-specific harmony threshold). This is done by first determining, based on the known harmony patterns, whether each word in the corpus is harmonic or not, then comparing the number of harmonic words to the expected number of harmonic words that would occur in a similar corpus with vowels distributed uniformly at random.

Hidden Markov model (HMM) approaches (Baker 2009; Goldsmith & Xanthos 2009; Knowles 2012) apply a common natural language processing technique to the problem of vowel harmony. An HMM includes a "start" state, an "end" state, and some number of connected hidden states. At each point in time, there is a transition with some probability from one hidden state to another, followed by the emission of a particular observable item with some hidden-state-specific emission probability. For vowel harmony, what is being emitted (observed) is the sequence of vowels (typically within one word at a time). The hidden states can be thought of as representing vowel classes (such as a natural class defined by a single value of a binary feature). The probabilities can be initialized randomly, or based on external knowledge (e.g., vowel frequencies). The harmony patterns learned are limited by the configuration of the model: an HMM with two hidden states can learn one two-class harmony system (i.e., both natural classes for a binary feature, such as round and nonround), while learning two separate two-class harmony systems would require at least four hidden states. The parameters of the model (transition and emission probabilities) are fit to a dataset using an iterative algorithm. Intuitively, fitting an HMM to a language with vowel harmony will result in an HMM with a low between-state transition probability (a word is unlikely to contain vowels from more than one harmonic class) and with one state that has high emission probabilities for one class of vowels (the state can be thought to represent that class) while the other state(s) represents the other class(es). HMM-based approaches have been shown to successfully learn harmony patterns for Turkish, Finnish, Hungarian, and other languages. While HMM approaches do model harmony at the whole word level (that is, the start of the word starts at the "start" state, each vowel is emitted by one of the hidden states, and reaching the "end" state indicates the end of a word), they also rely on the Markov assumption: that the probability of the next state depends only on the current state (arguably treating vowel harmony as a local phenomenon between consecutive vowels).

Knowles (2012) also uses word boundaries, and makes an even stricter whole-word harmony assumption with a mixture of unigrams model. That model can be thought of as a special case of the HMM, one where there is zero probability of transitioning between states. A measure of how strongly the two learned probability distributions in the mixture of unigrams model differ correlates with results from the Vowel Harmony Calculator (Harrison et al. 2004).

Other approaches keep the word boundary intact, but only compute statistics from pairs of vowels. Sanders & Harrison (2012) do this with the goal of comparing the level of harmony across languages (with some level of external knowledge). They extract all tier-adjacent vowel pairs, compute raw pair-wise harmony for each feature of interest (such as height or backness), randomly generate corpora with the same distribution of vowel tokens and lengths, compute that same pair-wise harmony for the feature of interest across all random corpora, and then (making the assumption that the random corpora are distributed normally), compute the z-score of the true corpus from the mean (interpretable as the number of standard deviations from the mean). This allows them to uncover patterns of harmony as well as anti-harmony (less harmony than would be expected), and their results correlated with work in Harrison et al. (2004). Szabó & Çöltekin (2013) take a point-wise mutual information approach, fitting a general linear model where the response variable is the point-wise mutual information for vowel bigram types, and the features are indicator variables representing information about vowel bigram harmony.[3] Goldsmith & Xanthos (2009) produce matrices of vowel co-occurrences, to which they apply spectral decomposition to extract eigenvectors that show clusters of vowels. Mayer et al. (2010) also measure vowel co-occurrences for visualization purposes (see Section 35.4).

At the other extreme, it is possible to ignore word boundaries, building models that take as input a stream of characters without any information about the beginning or end of words. Building statistical models of vowel harmony acquisition, Kodner et al. (2017) motivate the use of character streams without word boundaries by noting that infants' sensitivity to vowel harmony predates their ability to segment continuous speech.[4] They model vowel harmony using tier-adjacent vowel pairs, computing point-wise mutual information (logarithm of the normalized conditional probability of the vowel pair occurrence). They then perform k-means clustering to divide the vowels into two groups. Given additional information about vowel features, they can learn a second harmony system by pairing vowels from the first two groups across one feature dimension, and then performing a second round of k-means clustering. They evaluate across eight languages, including six with harmony (Turkish, Finnish, Hungarian, Uyghur, and Warlpiri), one with historical remnants of harmony (Estonian), and two without (German and English); they successfully learn one harmony system but have mixed results for a second. It is worth noting that there is another statistical and acquisition-based argument for ignoring word boundaries when computing pair-wise harmony measures: vowel pairs that cross boundaries are likely less frequent than those within words (unless the language has very short words), so the statistics should still express the language's harmony. In fact, harmony patterns may help children learn to segment (Mintz et al. 2018), and adults can also use them to segment speech (Suomi et al. 1997; Kabak et al. 2010).

[3]They define their bigrams to be pairs of vowels with one or more consonants between them, so as not to treat diphthongs in the same manner as other tier-adjacent vowel pairs.

[4]They extend this work in Caplan & Kodner (2018), which has a greater focus on models for child language acquisition. For more discussion of vowel harmony acquisition outside of computational approaches, see Chapter 38, this volume.

35.3.2 Phonological considerations

As we note, most approaches to the computational study of vowel harmony focus only on vowels, separating them out in a tier-based approach. Such models are unable to capture the effects of consonants in harmony systems. For example, in Assamese (an eastern Indo-Aryan language spoken in India), regressive ATR harmony can be blocked by nasal consonants, as in [sɛkɔni] 'strainer', which would be *[sekoni] if harmony were not blocked by [n] (Mahanta 2007; for more on Assamese harmony, see Chapters 24 and 56, this volume).

Further, consonants may actually participate in harmony due to assimilation (Chapter 2, this volume), which could facilitate discovering vowel harmony if harmonizing consonants were taken into account in the statistical analysis, since longer sequences of harmonizing consonants and vowels would be even less likely to arise by chance than vowels alone. For example, in Turkish, velars and laterals are palatalized in front-harmonizing contexts (Chapter 59, this volume), as seen in the alternations of the velars and laterals in [gʲøz-lʲer] 'eye-PL' versus [kɯz-lar] 'girl-PL'. While blocking consonants such as Assamese nasals may be captured by a model that includes consonants as input (we discuss this further in Section 35.3.3), the issue of consonant assimilation as in Turkish provides additional challenges, particularly for written corpora. Consonant assimilation is often not captured by the orthography and would only be apparent in sufficiently narrow transcriptions.

A rare example of models with the potential to capture more of these phonological effects are those proposed by Goldsmith & Riggle (2012). Their aim is to examine models of vowel harmony through an information theoretic lens, comparing different models on the basis of the probability that they assign to text. They find that their first attempt to build a model using an autosegmental approach – namely by using a combination of a character bigram model where all vowels are collapsed to a single representative "V" along with a bigram model that operates on the vowel tier – underperforms a standard bigram model. They find more success with a Boltzmann model that incorporates unigram character probabilities, mutual information between adjacent characters, and mutual information between tier-adjacent vowels, which shows that the interaction between consonants and vowels cannot be completely cleanly captured by two fully separate tiers. As they focus on a particular language, Finnish, measuring whole-word harmony on a text corpus, their information-theoretic measures capture model fit, but do not focus on how a learned model could be used by a linguist to provide insight about the particular vowel harmony patterns in a given language. An examination of model probabilities, perhaps through a visualization tool like those described in Mayer (2012), could have the potential to illuminate such patterns.

35.3.3 Transparent and blocking segments

A common feature of vowel harmony systems is the presence of segments that do not participate in the harmony system. These include transparent vowels: even though they may appear disharmonic, the remaining vowels will be harmonic with one another, as though the transparent vowels in the word were not there at all (Chapter 21, this volume).

The same often holds true for consonants (Chapter 2, this volume), which enables the approaches that strip out all consonants in order to treat vowels as tier-adjacent. However, there are instances of opaque vowels and consonants which block harmony. Many approaches have some way of handling or evaluating transparent vowels (e.g., Ozburn 2019), but few provide a way to handle opaque vowels. The frequent design decision to remove all consonants from model consideration implicitly treats all consonants as transparent. Most systems focus on vowels only, making it impossible for the systems to handle or highlight harmony-blocking consonants. As with the other challenges discussed so far, different approaches handle (or ignore) transparent and blocking segments in different ways.

Systems like the Vowel Harmony Calculator (Harrison et al. 2004), which incorporate external knowledge, can easily allow researchers to designate specific transparent vowels as not participating in the harmony process. HMM approaches also handle transparent vowels quite naturally; these typically appear as vowels that can be emitted by a state representing any vowel class (Baker 2009; Goldsmith & Xanthos 2009; Knowles 2012). While a vowel that participates in the harmony system will have high emission probability in one state (the state associated with its harmonic class) and low probability in the other, a transparent vowel will not be marked in such a way. Rather, it will have similar probabilities of being emitted by either state. Capturing opaque vowels may require additional states or a bigram HMM, but this remains an open research question. Using a spectral decomposition approach on Finnish vowel co-occurrence, Goldsmith & Xanthos (2009) find that the second eigenvector clusters the neutral (transparent) vowels into their own group, separate from front and back vowels (see Chapters 18 and 67, this volume, for more about transparent vowels in Finnish). However, they note this visually, and do not provide an automatic way of extracting these clusters from the eigenvector information.

35.4 Visualization of vowel harmony

Many approaches discussed so far have the goal of providing a linguist with a form of hypothesis generation – a hint to examine a language's data for evidence of harmony – or hypothesis confirmation of suspected harmony patterns. For many people, an intuitive way to interact with such data is visually. Even approaches like the spectral decomposition approach present their output (the second eigenvector) visually, such that the reader can immediately see vowel clusters on a number line (Goldsmith & Xanthos 2009). HMMs also have an interpretable visual form, though it may require the viewer to read out transition and emission probabilities.

Several works focus specifically on building interpretable visualizations of vowel harmony, intended to be useful to linguists. Mayer et al. (2010) compute vowel succession probabilities and produce a matrix of ϕ coefficients (a value between 0 and 1, representing the association strength between a pair of vowels). They also examine whether each vowel bigram occurs more or less frequently than expected. All of this results in a matrix of numbers which could, on its own, take a long time to interpret. They propose a simplified visual representation of the matrix, using gradients of two colors (one to represent sequences that occur more frequently than expected and one for those that occur less frequently) along with plus and minus signs to distinguish the two when values are close to zero. They automatically sort the matrices to better display apparent clusters of vowels, as shown in our reproduction of a similar style in Figure 35.1, where front vowels and back vowels form clear distinct clusters (indicated by +) with the transparent vowels appearing with vowels from both clusters.[5] Thus, at a glance, one can see whether there appear to be distinct vowel clusters, how strong that clustering is, and whether those vowel pairs appear more or less frequently than they would by chance. They also provide interactive options, allowing a user to hover to get more detailed information from the plot. Mayer (2012) expands on this and provides additional background. In Mayer & Rohrdantz (2013), they present *PhonMatrix*, the online tool that for producing vowel harmony matrices (and more general matrices of phone-pair associations). A strength of this work is that it does not require supervision; the values for the matrix are extracted automatically, as is the matrix ordering.

<INSERT FIGURE 35.1 (MayerStyle.eps) APPROXIMATELY HERE>

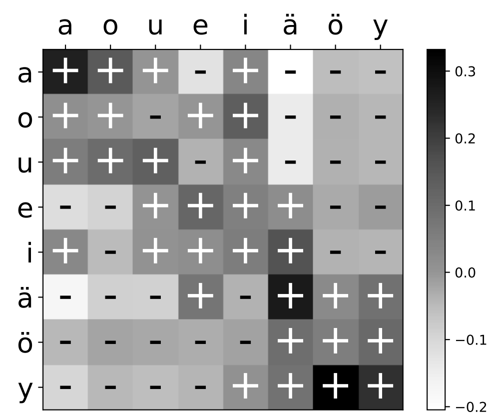
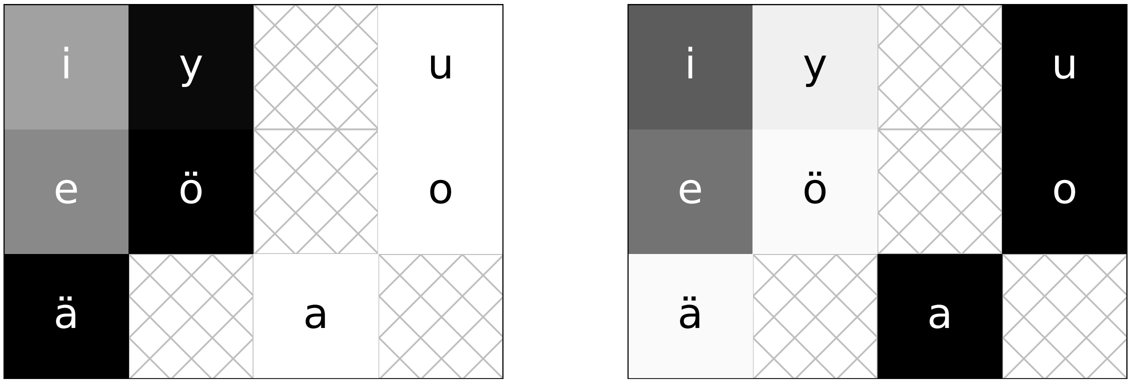


Figure 35.1: Finnish harmony visualization in the style of Mayer et al. (2010). Each square represents how much more/less frequently the vowel in the column follows the vowel in the row than would be expected by chance (+/dark indicates greater than expected frequency, −/light indicates less frequent than expected). Here we use grayscale; the original work uses two colors for clearer visualization.

Requiring slightly more supervision, Knowles (2012) also presents a visualization approach. It requires that the user provide vowels along with their features in a grid (roughly approximating the vowel space), which is then used to produce plots for each of the vowel classes and neutral vowels (if any). By requiring additional supervision, this visualization draws attention to articulatory features of interest, as shown in Figure 35.2. One could imagine combining these to add articulatory features to the Mayer work, or to automatically compare the clusters found in the Mayer work to articulatory features, along the lines of the approach in Szabó & Çöltekin (2013).

<INSERT FIGURE 35.2 (FinnishFrontThesis.eps; FinnishBackThesis.eps) APPROXIMATELY HERE OR BEFORE THIS PARAGRAPH>

[5] All figures use statistics computed from the finnish-gospels.txt file in Harrison et al. (2004) and are generated using Matplotlib (Hunter 2007).



(a) Front class (b) Back class

Figure 35.2: Finnish vowel harmony visualization from Knowles (2012); each subfigure represents a distribution over vowels as learned by the model, with dark colors representing higher probabilities. Comparing them, we see two distinct vowel classes: front and back, with transparent vowels *i* and *e* appearing alongside vowels of either class.

35.5 Concluding remarks

Vowel harmony is often a partial phenomenon, affecting only some vowels in some words, making statistical data-driven approaches a natural fit. Despite many complicating factors, these approaches have been successfully applied to a variety of tasks, including detection of vowel harmony from text corpora, the study of child acquisition of harmony patterns, examinations of models through information theoretic approaches, and visualizations. Many approaches separate vowels into their own tier and compute statistics from vowels alone, but approaches that also incorporate consonants show promise for capturing additional information about vowel harmony patterns. This is a rich and active field of research, but given limited space, we are unable to discuss the full range of important corpus-based studies of vowel harmony, including but not limited to Hayes & Londe (2006), Forró (2013), and Kabak et al. (2008), or models of vowel harmony over time (Harrison et al. 2002; Chapter 43, this volume). Future directions that may prove fruitful include increased incorporation of the role of consonants, as well as visualizations that combine automatic visualization and clustering with external knowledge about articulatory features.

Acknowledgments

We thank the reviewers and editors for their thoughtful feedback, which has helped strengthen this chapter.

References

Baker, Adam C. 2009. Two statistical approaches to finding vowel harmony. Master's thesis, University of Chicago, Chicago.

Caplan, Spencer & Jordan Kodner. 2018. The acquisition of vowel harmony from simple local statistics. In Chuck Kalish, Martina A. Rau, Xiaojin (Jerry) Zhu, & Timothy T. Rogers (eds.), *Proceedings of the 40th Annual Meeting of the Cognitive Science Society*, 1440–1445. Austin, TX: Cognitive Science Society.

Demuynck, Kris, Tom Laureys & Steven Gillis. 2002. Automatic generation of phonetic transcriptions for large speech corpora. In John H. L. Hansen & Bryan L. Pellom (eds.), *Proceedings of the 7th International Conference on Spoken Language Processing*, vol. 1, 333–336. Denver, CO: ISCA.

Forró, Orsolya. 2013. *Ingadozás a magyar elölségi harmóniában: Szempontok a variabilitás szinkróniájának és diakróniájának feltárásához és értelmezéséhez* [Variation in Hungarian backness harmony: Aspects of the investigation and interpretation of the synchrony and diachrony of variability]. PhD dissertation: Pázmány Péter Katolikus Egyetem, Piliscsaba, Hungary.

Goldsmith, John & Jason Riggle. 2012. Information theoretic approaches to phonological structure: The case of Finnish vowel harmony. *Natural Language & Linguistic Theory* 30: 859–896.

Goldsmith, John & Aris Xanthos. 2009. Learning phonological categories. *Language* 85: 4–38.

Harrison, K. David, Mark Dras & Berk Kapicioglu. 2002. Agent-based modeling of the evolution of vowel harmony. In Masako Hirotani (ed.), *Proceedings of the North East Linguistics Society*, vol. 32.1, 217–36. Amherst, MA: GLSA.

Harrison, K. David, Emily Thomforde & Michael O'Keefe. 2004. The Vowel Harmony Calculator. `http://www.swarthmore.edu/SocSci/harmony/public_html/index.html`.

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

Hunter, John D. 2007. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering* 9: 90–95.

Kabak, Barış, Eva Kasselkus, Kazumi Maniwa & Silke Weber. 2008. Vowel harmony has direction and context: A corpus study. Paper presented at the 16th Manchester Phonology Meeting.

Kabak, Barış, Kazumi Maniwa & Nina Kazanina. 2010. Listeners use vowel harmony and word-final stress to spot nonsense words: A study of Turkish and French. *Laboratory Phonology* 1: 207–224.

Knowles, Rebecca. 2012. Vowel harmony: Statistical methods for linguistic analysis. Bachelors thesis: Haverford College, Haverford, PA.

Kodner, Jordan, Spencer Caplan, Hongzhi Xu, Mitchell P. Marcus & Charles Yang. 2017. Case studies in the automatic characterization of grammars from small wordlists. In Antti Arppe, Jeff Good, Mans Hulden, Jordan Lachler, Alexis Palmer, & Lane Schwartz (eds.), *Proceedings of the 2nd Workshop on the Use of Computational Methods in the Study of Endangered Languages*, 76–84. Honolulu: Association for Computational Linguistics.

Mahanta, Shakuntala. 2007. *Directionality and locality in vowel harmony: With special reference to vowel harmony in Assamese*. PhD dissertation: Utrecht University, Utrecht.

Mayer, Thomas. 2012. *The induction of phonological structure*. PhD dissertation: Universität Konstanz, Konstanz.

Mayer, Thomas & Christian Rohrdantz. 2013. PhonMatrix: Visualizing co-occurrence constraints of sounds. In Miriam Butt & Sarmad Hussain (eds.), *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, 73–78. Sofia: Association for Computational Linguistics.

Mayer, Thomas, Christian Rohrdantz, Miriam Butt, Frans Plank & Daniel A. Keim. 2010. Visualizing vowel harmony. *Linguistic Issues in Language Technology* 4: 1–33.

Mintz, Toben H., Rachel L. Walker, Ashlee Welday & Celeste Kidd. 2018. Infants' sensitivity to vowel harmony and its role in segmenting speech. *Cognition* 171: 95–107.

Ozburn, Avery. 2019. *A target-oriented approach to neutrality in vowel harmony*. PhD dissertation: University of British Columbia, Vancouver.

Ribeiro, Eduardo Rivail. 2002. Directionality in vowel harmony: The case of Karajá (Macro-Jê). In Julie Larson & Mary Paster (eds.), *Proceedings of the Twenty-Eighth Annual Meeting of the Berkeley Linguistics Society*, 475–485. Berkeley, CA: Berkeley Linguistics Society.

Sanders, Nathan & K. David Harrison. 2012. Discovering new vowel harmony patterns using a pairwise statistical model. Poster presented at the 20th Manchester Phonology Meeting.

Suomi, Kari, James M. Mcqueen & Anne Cutler. 1997. Vowel harmony and speech segmentation in Finnish. *Journal of Memory and Language* 36: 422–444.

Szabó, Lili & Çağrı Çöltekin. 2013. A linear model for exploring types of vowel harmony. *Computational Linguistics in the Netherlands Journal* 3: 174–92.

Topping, Donald M. 1968. Chamorro vowel harmony. *Oceanic Linguistics* 7: 67–79.