2026. Proceedings of the Workshop on Turkic and Languages in Contact with Turkic 10. 86–100.

Vowel cooccurrence as a predictor of perceived vowel similarity in Turkish

Züheyra Tokaç & Jennifer Cole*

Abstract. From an exemplar models perspective, vowel harmony classes emerge from vowel cooccurrence patterns in the lexicon through similarity-based coactivation of exemplars. Although the vowel inventory of Turkish is phonologically symmetrical and Turkish vowel harmony is highly systematic and pervasive, Turkish vowel harmony as it manifests in the lexicon is not exceptionless—Turkish vowel pairs vary in their likelihood of cooccurrence in the lexicon. We hypothesize that Turkish vowel pairs with higher likelihood of cooccurrence will have higher perceived similarity than vowel pairs with lower likelihood of cooccurrence. We test this hypothesis in a Turkish vowel identification experiment with 40 native Turkish listeners. We analyze the perceptual confusion responses as a measure of perceived vowel similarity and find that, contrary to our hypothesis, higher likelihood of vowel cooccurrence yields fewer perceptual confusions indicating lower perceived vowel similarity. In a post-hoc analysis, we find that the relationship between likelihood of vowel cooccurrence and perceived vowel similarity is qualified by vowel peripherality, with non-peripheral vowels driving the unexpected relationship between vowel cooccurrence and perceived similarity. We discuss the results in relation to the role of vowel peripherality in Turkish vowel harmony (cf. Clements & Sezer 1982; Suomi 1983).

Keywords. Turkish vowel harmony; vowel cooccurrence; exemplar models; perceived vowel similarity

1. Introduction. Vowel harmony is traditionally analyzed in terms of abstract rules operating on discrete phonological features that prescribe agreement in the harmony feature for vowels cooccurring within a phonological domain such as the word (e.g., Vago 1973; van der Hulst 2016). However, in at least some harmony languages, harmony is not pervasive across the lexicon, with both harmonic and disharmonic vowel cooccurrence patterns observed. In this paper we ask what consequences, if any, follow from an incomplete application of harmony across the lexicon. This question is prompted by considering harmony systems from the perspective of exemplar theory. In exemplar models, phonological categories such as phonemes or feature classes emerge from the similarity-based coactivation of stored exemplars, where each exemplar represents a unique speech event (Goldrick & Cole 2023; Johnson 1997; Pierrehumbert 2001). The classic example for describing the mechanisms proposed by exemplar models involves vowel perception: each vowel token is stored in memory with its phonetic detail, and each new vowel token is compared to the stored exemplars on the phonetic space, which are coactivated as a function of their phonetic similarity to the new vowel token. The new vowel token is thus identified probabilistically via the similarity-based coactivation of the previously stored exemplars. Exemplars that are phonetically similar and coactivated lead to the emergent representation of a vowel phoneme category. This proposed mechanism can also be applied to lexical exemplars rather than individual phoneme tokens: the production of the word *'big'* will trigger not only

---

[*] We would like to thank the two anonymous reviewers and the audience at Tu+10 for their comments, and Chun Chan for his help with experimental setup. Authors: Züheyra Tokaç, Northwestern University (zuheyratokac@u.northwestern.edu) & Jennifer Cole, Northwestern University (jennifer.cole1@northwestern.edu).

activation of that word, but activation of '*bag', 'bin',* and other words that begin with [b]. The pattern of coactivation here defines an emergent construct of [b] in syllable- and word-onset position. More complex patterns can also emerge. For instance, consider a toy example of a vowel backness harmony system, where the occurrence of a word like 'bibe' with [i] in the initial syllable will trigger coactivation of other words with [i] in the same position, like 'bibi, bime', etc. from which emerges a class of front-harmonic words. In other words, similar to the way in which the vowel tokens are organized according to their phonetic similarity on the phonetic space, on a lexical level of representation, one way in which exemplars can be organized is according to their cooccurrence similarity, whereby exemplars with greater cooccurrence similarity will be coactivated more. In this manner the phonotactic pattern defining harmony can be viewed as an emergent property of the coactivation of vowels across positions in a word, with the coactivated vowels emerging as a harmony class, such as [-back] in the example above (Cole 2009).

We argue here that viewed from the perspective of exemplar theory, the presence of a vowel harmony pattern in the lexicon is predicted to impact perceived vowel similarity. This follows from the bidirectional relationship between coactivation and perceptual similarity, as follows. First, as a central tenet of exemplar theory, coactivation results from the spreading activation of phonetically detailed word instances stored in the exemplar lexicon (e.g., Pierrehumbert 2001). Second, evidence from categorical perception shows that in at least some instances, coactivation leads to an increase in perceived similarity. Specifically, phonetic distinctions between sounds that are perceived as instances of the same phoneme category may go undetected by listeners, as shown in classical tests of perceptual discrimination (e.g., Kuhl 1991). In the case of vowel harmony, lexical coactivation based on cooccurrence similarity gives rise to emergent classes of harmonic vowels, as sketched above, and as a (weak) form of categorical perception, predicts an increase in the perceived similarity of vowels belonging to the same harmony class. Partial support for this hypothesis is provided in a cross-linguistic perceived vowel similarity rating study (Terbeek 1977). Terbeek found that listeners rated nonnative vowel pairs to be perceptually less similar if a vowel pair is phonemically contrastive in their native language compared to listeners of languages in which the vowel pair is not contrastive. In addition, across ratings of vowel pairs that are contrastive in listeners' native languages, Turkish listeners rated nonnative vowel pairs to be perceptually less similar than other listeners if the vowel pair consisted of a [+back] vowel and a [-back] vowel, suggesting that vowels are perceived to be more similar within a harmony class than across. In other words, these results suggest that phonological knowledge warps the perceptual vowel space and perceived vowel similarity is influenced by phonological factors beyond the phonetic similarity of vowel sounds. Similar to the phonetic distance between two vowel sounds being perceptually enhanced if the sounds are phonemically contrastive, experience with vowel harmony might warp the perceptual vowel space and perceptually enhance the phonetic distance between vowels across harmony classes. Stated more generally, two vowels that have a high likelihood of cooccurrence in the same word will have higher levels of coactivation and are therefore predicted to have higher perceived similarity. The present study tests this prediction, examining whether the likelihood of vowel cooccurrence in the Turkish lexicon, which reflects vowel harmony, is a predictor of vowel similarity as perceived by native speakers.

One can argue that the hypothesis from exemplar theory is counterintuitive; if vowels that frequently cooccur in words are perceived to be more similar to one another beyond their phonetic similarity predicts, that might lead to contrast neutralization for these vowels over time, which would make vowel harmony counterproductive to maintaining lexical contrasts. Relatedly, an alternative hypothesis to consider is that vowels that cooccur more frequently in words will become perceptually more dissimilar rather than more similar over time. This hypothesis follows from the reasoning that vowel harmony helps to establish that the cooccurring vowels are contrastive rather than allophonic, because vowels belonging to the same harmony class are more likely to have overlapping distributions as a result of their phonotactic similarity. For instance, a lexical contrast between the words 'bibe' and 'bibi' in the toy example above would arguably help establish a phonemic contrast between /e/ and /i/ and hence lead to enhanced perceived *dis*-similarity between these vowels. In this sense, greater likelihood of vowel cooccurrence might facilitate rather than compromise contrast preservation. To anticipate our discussion below, we find that the relationship between vowel cooccurrence and perceived vowel similarity is qualified by one particular index of phonetic similarity, such that our results support both the hypothesis from exemplar theory (greater cooccurrence similarity leads to greater perceived similarity) *and* the contrast preservation hypothesis, although the evidence for the latter is stronger.

1.1. TURKISH VOWEL PHONOLOGY. The most prominent property of Turkish vowel phonology is that Turkish has backness harmony and rounding harmony (e.g., Erguvanlı-Taylan 2015), where backness harmony is noted as being more systematic and pervasive, with all Turkish vowels participating as both triggers and targets. The Turkish vowel system of 8 vowels is defined by three vowel features, namely [±back], [±round], and [±high] (Table 1), and is phonologically symmetrical when considered from the perspective of a rule-based account of Turkish vowel harmony, though not on the basis of Turkish vowel phonetics: phonetically, Turkish vowels are continuously distributed along the F2 dimension and group into three rather than two F1 levels (e.g., Kopkallı-Yavuz 2010). In contrast, descriptively, Turkish vowel harmony involves agreement in [±back] word-internally and with the stem-final vowel in suffixation, and agreement in [±round] with the vowel in the immediately preceding syllable for the [+high] vowels. The descriptive account of Turkish vowel harmony is visually represented in Table 2, where successive vowel pairs are assigned discrete harmonic (+) or disharmonic (–) labels. Thus, the phonological vowel system in Table 1 emerges based on the Turkish vowels' participation in the two types of vowel harmony.

	front		back	
	rounded	unrounded	rounded	unrounded
high	y	i	u	ɯ
nonhigh	œ	e	o	ɑ

Table 1. Turkish vowel inventory

	V2							
V1	ɑ	ɯ	o	u	e	i	œ	y
ɑ	+	+	+	–	–	–	–	–
ɯ	+	+	+	–	–	–	–	–
o	+	–	+	+	–	–	–	–
u	+	–	+	+	–	–	–	–
e	–	–	–	–	+	+	+	–
i	–	–	–	–	+	+	+	–
œ	–	–	–	–	+	–	+	+
y	–	–	–	–	+	–	+	+

Table 2. Rule-based description of Turkish vowel harmony. + represents vowel pairs that are backness- and rounding-harmonic, whereas – represents vowel pairs that are disharmonic.

Despite the discrete, rule-based description of Turkish vowel harmony summarized above, neither type of vowel harmony in Turkish is exceptionless as it manifests in the lexicon. For instance, Kabak et al. (2008) show that about 1/3 of disyllabic Turkish words are disharmonic, and among these, certain harmonic vowel pairs as well as some disharmonic vowel pairs are more common than other vowel pairs. Studies computationally modeling vowel cooccurrence patterns and vowel-to-vowel transition probabilities calculated over the Turkish lexicon confirm that Turkish vowels with high likelihood of cooccurrence define emergent groups that align with the Turkish vowel harmony classes (e.g., Baker 2008; Caplan & Kodner 2018). In comparison, these studies find that vowel cooccurrence and transition probabilities in the lexicons of languages without vowel harmony such as English reveal no systematic vowel cooccurrence patterns. Rather, in languages such as English, vowel pairs have much less variability in their likelihoods of cooccurrence compared to Turkish where certain (harmonic) vowel pairs have high likelihood of cooccurrence whereas other (disharmonic) vowel pairs have low likelihood of cooccurrence.

One way in which a vowel pair's likelihood of cooccurrence can be quantified is Pointwise Mutual Information (PMI), calculated as in (1), where *a* and *b* are two vowels, by taking the logarithm to the base 2 of the total of bigram frequencies of the two vowels cooccurring in disyllabic words divided by the multiple of the unigram frequencies of each of the vowels occurring in disyllabic words:

(1) $$PMI = \log_2 \frac{P(ab)+P(ba)}{P(a)\ x\ P(b)}$$

The resulting PMI value indicates whether the two vowels *a* and *b* cooccur in disyllabic words more frequently than what would be predicted if the two vowels are independently distributed in the disyllabic words of the lexicon: a positive PMI value denotes high likelihood of cooccurrence, which is expected for harmonic vowels in a language like Turkish, a PMI value of 0 denotes independent distribution of vowels, which is expected for vowels that are not subject to harmony or other cooccurrence constraints, and a negative PMI value denotes a low likelihood of cooccurrence, which is expected for disharmonic vowels.

To obtain a non-discrete description of Turkish vowel harmony as it manifests in the lexicon, we calculated PMI over 18,166 disyllabic Turkish words in the 1998 Turkish dictionary from the Turkish WordNet corpus (Bakay et al. 2021), as described above. Figure 1 shows the resulting PMI values for each Turkish vowel pair as a heatmap, where purple cells represent vowel pairs with positive PMI and hence high likelihood of cooccurrence and orange cells represent vowel pairs with negative PMI and hence low likelihood of cooccurrence. As such, comparable to Table 2 which represents whether a vowel pair is harmonic or disharmonic in discrete terms, the PMI heatmap represents whether a vowel pair is likely to cooccur or not. Yet, the PMI heatmap additionally represents the gradience in likelihood of vowel cooccurrence in the Turkish lexicon. The shading of cells in the heatmap represent the varying degrees of likelihood of vowel cooccurrence, with darker purple denoting higher likelihood of cooccurrence than lighter purple, and darker orange denoting lower likelihood of cooccurrence than lighter orange. Thus, unlike the discrete representation of Turkish vowel harmony in Table 2, the heatmap of likelihood of vowel cooccurrence in the Turkish lexicon represents how much Turkish vowel pairs vary in their likelihood of cooccurrence.

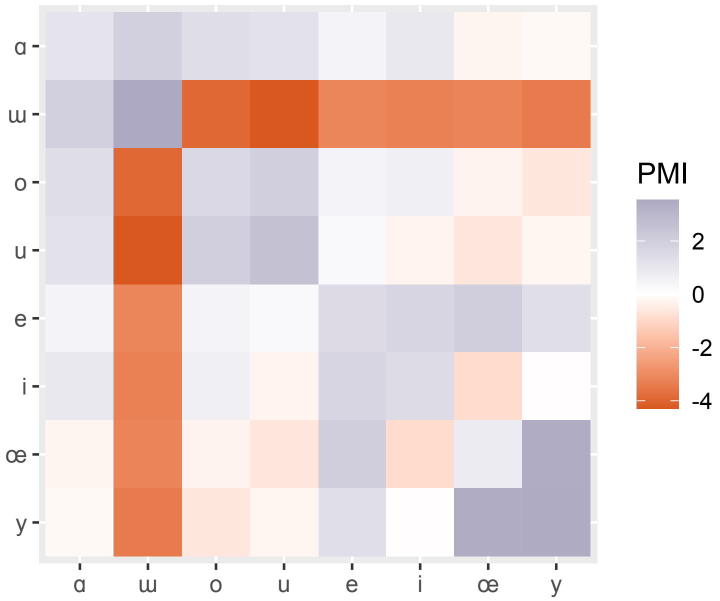

Figure 1. Heatmap of PMI values for the Turkish vowel pairs in the disyllabic Turkish words, indicating the likelihood of vowel cooccurrence in the Turkish lexicon.

As described above, we argue that a hypothesis of exemplar models is that greater cooccurrence similarity such as in the case of harmonic vowels would lead to greater perceived similarity beyond what is predicted by the phonetic similarity of vowels. When we apply this understanding of similarity-based coactivation to the heatmap in Figure 1, our hypothesis is that not only will harmonic vowel pairs be perceived to be more similar to each other than disharmonic vowel pairs, but also vowel pairs with higher PMI values and hence higher likelihood of cooccurrence will be perceived to be more similar to each other than vowel pairs with lower PMI values and hence lower likelihood of cooccurrence. We test this hypothesis in an 8-alternative forced choice (8-AFC) Turkish vowel identification experiment. We take perceptual confusions in vowel identification as a measure of perceived vowel similarity, and model perceptual confusion responses as a function of PMI values in addition to other phonetic and phonological predictors of vowel identification. We predict that vowel pairs with higher PMI values will be perceptually confused with one another more than will be vowel pairs with lower PMI values.

2. Methods. Forty native Turkish listeners (24 women, 1 other) participated in a vowel perception study which included the 8-AFC vowel identification experiment of interest here. Participants were monolingually raised Turkish speakers aged 18-61 (M_{age} = 26.35) residing in Istanbul, Turkey, at the time of their participation. Thirty-two participants reported that they were fluent in English as a second language and one participant reported that they were fluent in German as a second language. Participants completed a total of 160 vowel identification trials. Stimuli were 160 naturally produced monosyllabic /bVb/ nonwords, with 20 nonword tokens per each of the 8 Turkish vowels in the /b_b/ context produced by 20 native Turkish speakers in a separate production study (Tokaç 2025). The native Turkish speakers (12 women) were aged 19-56 (M_{age} = 28.45), monolingually raised in Turkey, fluent in English, and residing in Evanston, Illinois, at the time of their participation in the production study. Each vowel identification trial began with a play button visually presented on screen. When the participant pressed the play button, a randomly selected auditory stimulus containing a monosyllabic nonword token was auditorily presented once, and the participant was asked to identify the vowel in the nonword stimulus from among the 8 Turkish vowels visually presented on screen in their orthographic representations (Figure 2). The response options were presented in the same alphabetical order in every single trial in a circular array with the 4 back vowels to the left of the array and the 4 front vowels to the right of the array. Participants made a response by clicking on one of the 8 response options, after which a new trial automatically began with the play button visually presented on screen.

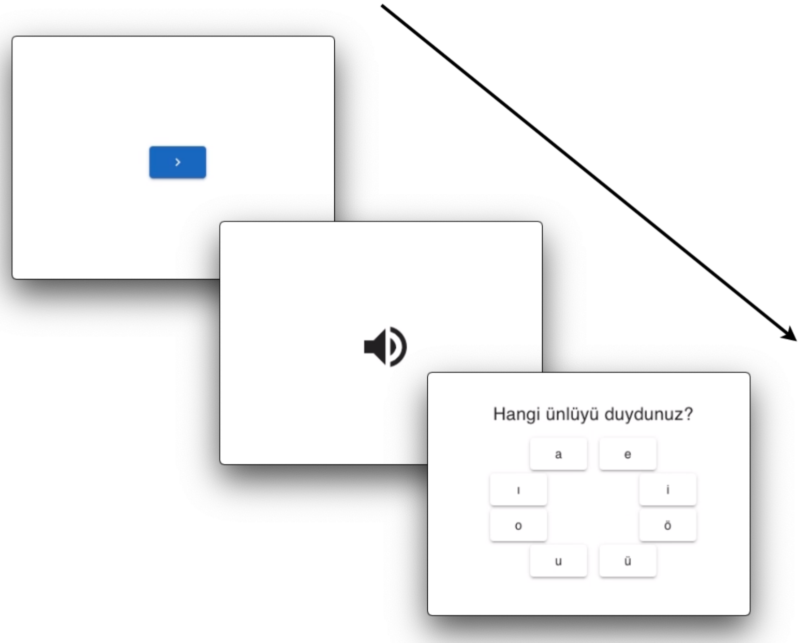

Figure 2. Flow of a trial in the 8-AFC Turkish vowel identification experiment.

3. Results. Out of the total aggregate of 6400 vowel identification trials, participants had a total of 670 responses in which a stimulus vowel was perceptually confused with another vowel, yielding an accuracy rate of 89.57%. Mean response time measured from stimulus onset was 1626.19 ms. A linear mixed effects regression modeling response times with accuracy as the predictor and random intercepts by participant and item revealed that accurate vowel identification responses were 196.18 ms faster than perceptual confusion responses ($SE = 12.68$, $t = -15.47$, $p < 0.0001$), suggesting that in trials that received inaccurate responses, listeners might have perceived the stimuli to be harder to identify and phonemically ambiguous. Figure 3 depicts participant responses as a heatmap of response proportions to each stimulus vowel category, where the dark blue cells in the diagonal represent the accurate vowel identifications and the remaining 56 cells represent the perceptual confusions, with darker shades representing a higher proportion of perceptual confusion. A binomial logistic regression modeling vowel identification accuracy with stimulus vowel category as the predictor (with /ɑ/ as the reference level) and random intercepts by participant and item revealed that accuracy rate varied by stimulus vowel category. The stimulus vowel categories /ɯ, œ, y/ exhibited significantly lower accuracy relative to /ɑ/ (/ɯ/: $\hat{\beta} = -3.30$, $SE = 0.61$, $z = -5.43$, $p < 0.0001$; /œ/: $\hat{\beta} = -2.15$, $SE = 0.61$, $z = -3.51$, $p = 0.0005$; /y/: $\hat{\beta} = -1.77$, $SE = 0.62$, $z = -2.87$, $p = 0.004$). In addition, the stimulus vowel category /e/ was marginally less accurate ($\hat{\beta} = -1.23$, $SE = 0.63$, $z = -1.94$, $p = 0.052$) whereas the stimulus vowel category /i/ was marginally more accurate ($\hat{\beta} = 1.37$, $SE = 0.76$, $z = 1.81$, $p = 0.07$) relative to /ɑ/. We return to a potential explanation of this pattern of results in the discussion section.

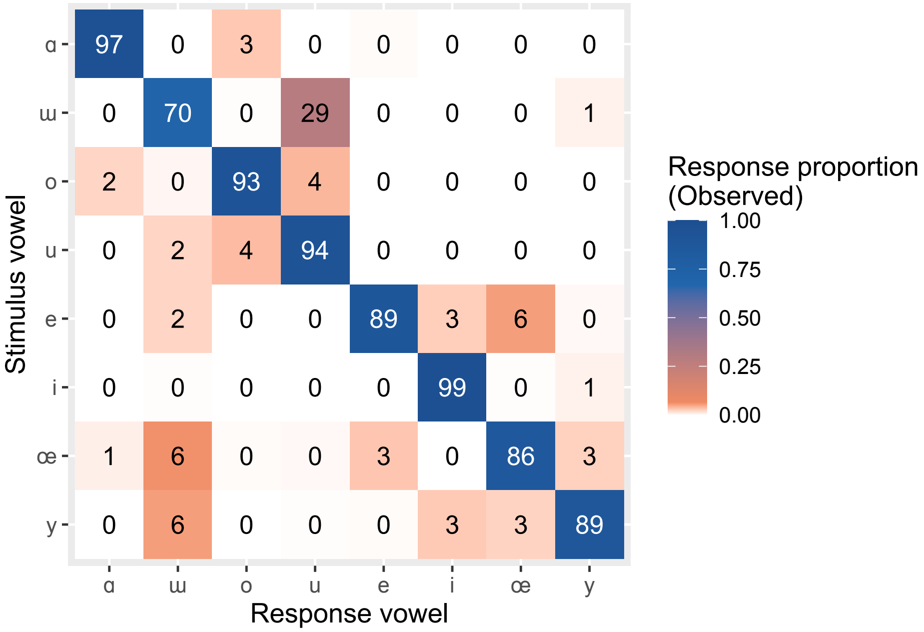

Figure 3. Heatmap of observed participant response proportions in vowel identification.

Participant responses were analyzed as count data to examine whether the likelihood of cooccurrence influences perceived vowel similarity as measured by perceptual confusions in vowel identification. For each of the 160 stimulus tokens, the number of times each of the perceptual confusion responses was made was calculated. Out of the 56 possible perceptual confusions, 28 were not observed in the data (see the white cells in Figure 2). To account for the zero counts in the data when modeling the count data, we ran a zero-inflated Poisson regression (pscl package; Jackman 2024; Zeileis et al. 2008). There were three sets of predictors in the model: phonetic descriptors of the stimulus token, phonetic vowel similarity measures, and phonological vowel similarity measures (Table 3). The phonetic similarity measures that involve vowel categories were calculated over the full set of monosyllabic nonword productions from which the stimuli in the vowel identification experiment were selected, which had 10 tokens per each of the 8 Turkish vowels by 27 monolingually raised native Turkish speakers (Tokaç 2025; see above). The categorical variables were sum coded and the continuous variables were scaled and centered. Note that although predictor variables such as Euclidean distance and Pillai score or feature dissimilarity ratio and PMI are conceptually related, all predictor variables included in the final model were examined for collinearity (all *VIF*s < 3.7), suggesting that the predictors are independent. The results from the Poisson regression are reported in Table 4. The results from the zero-inflated binomial logistic regression were mainly in line with the results from the Poisson regression and are omitted here for the sake of brevity.

Predictor type	Predictor	Notes
Phonetic descriptor of the stimulus token	Stimulus token F1 Stimulus token F2 Speaker gender Stimulus token duration	
	F1 : speaker gender	Interaction term (dropped due to collinearity)
	F2 : speaker gender	Interaction term
Phonetic similarity of the stimulus token to the response vowel category	Euclidean distance	Measure of phonetic distance between stimulus token and response vowel category centroid on F1-F2 space
	Pillai score	Measure of phonetic overlap between stimulus and response vowel categories on F1-F2 space
Phonological similarity of the stimulus token to the response vowel category	Feature dissimilarity ratio	Measure of phonological feature dissimilarity between stimulus and response vowel categories over the 3 Turkish vowel features
	PMI	Measure of likelihood of cooccurrence between stimulus and response vowel categories

Table 3. Predictors in the zero-inflated Poisson regression.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.01	0.15	0.05	0.96
F1	-0.02	0.09	-0.23	0.82
F2	0.47	0.13	3.68	< 0.001***
Speaker gender	0.09	0.10	0.93	0.35
Duration	-0.29	0.06	-5.16	< 0.001***
Euclidean distance	-1.38	0.15	-9.37	< 0.001***
Pillai score	-0.11	0.12	-0.91	0.36
Feature dissimilarity ratio	-0.34	0.07	-4.93	< 0.001***
PMI	-0.22	0.06	-3.68	< 0.001***
F2 : speaker gender	-0.07	0.14	-0.52	0.60

Table 4. Poisson regression results.

The significant main effects of stimulus vowel F2 and duration indicate that stimulus vowels with higher F2 and shorter vowel duration exhibit more perceptual confusions. Significant main effects of Euclidean distance and the phonological feature dissimilarity ratio were in the expected directions. Greater Euclidean distance between stimulus vowel token and response vowel category predicts fewer perceptual confusions, suggesting that vowels that are closer in phonetic space also have greater perceptual similarity. Similarly, greater vowel feature dissimilarity ratio between stimulus and response vowel categories predicts fewer perceptual confusions, suggesting that vowels with more similar phonological features also have greater perceptual similarity. Lastly, there was a significant main effect of PMI, suggesting that independent from their phonetic and/or phonological similarity, stimulus and response vowels that are more likely to cooccur in the Turkish lexicon are perceptually confused to a lesser degree (Figure 4). In other words, contrary to our hypothesis, we do not find evidence that greater cooccurrence similarity predicts greater perceived similarity. This finding regarding cooccurrence similarity challenges our hypothesis, calling to question the prediction from exemplar models that cooccurrence should lead to coactivation and by extension to greater perceived similarity.

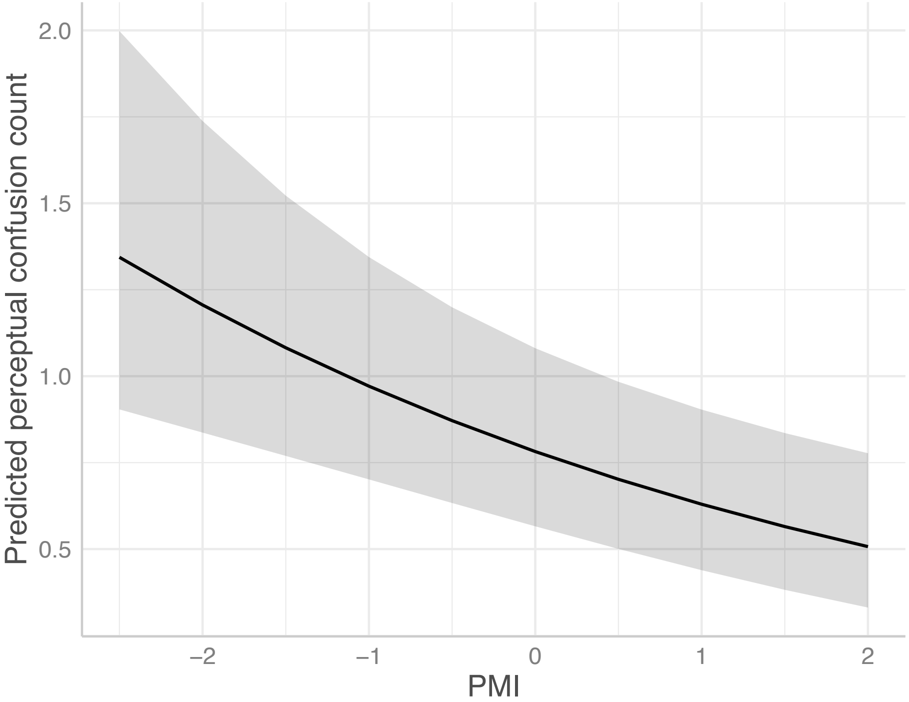

Figure 4. Perceptual confusion counts for stimulus and response vowel pairs as predicted by the PMI values of the stimulus and response vowel pairs.

4. Discussion. Our analysis of the perceptual confusion responses in the Turkish vowel identification experiment revealed that, contrary to our hypothesis, perceptual confusions are *less* likely between vowels that frequently cooccur in lexical items compared to the likelihood of confusions between vowels that cooccur with lower frequency. This result is more compatible with the contrast preservation hypothesis and is unexpected under the exemplar theory view that similarity-based coactivation increases the likelihood of perceptual confusion, because vowels that belong to the same harmonic class (e.g., [+back] vs. [-back]) have greater phonetic, phonological and phonotactic similarity than do vowels that are not from the same harmonic class. Yet as suggested above, on further consideration, this prediction from exemplar theory seems at odds with the survival of a harmony system over generational time because it predicts perceptual confusions among just those vowels that cooccur most frequently in lexical items. A higher rate of perceptual confusion among harmonic vowels would predict a greater likelihood of contrast neutralization among vowels belonging to the same harmonic class. Taken to the extreme, this could result in the collapse of a harmonic class to single contrast, e.g., a single "[+back]" vowel contrasting with a single "[-back]" vowel.

Counter to this line of reasoning, it has in fact been argued that vowel harmony is perceptually motivated in facilitating the identification of vowels that are insufficiently distinct in phonetic space (Kaun 1995; Suomi 1983). Suomi (1983) argues that for vowels of similar height (on the same F1 level), those vowels with more extreme F2 values have a perceptual advantage over neighboring vowels with less extreme F2 values. Applied to Turkish, /ɑ, e, i, o, u/ are the vowels with the more extreme F2 values, occupying peripheral regions of the vowel space, while /ɯ, œ, y/ are non-peripheral in the F2 dimension. Our results show that the non-peripheral vowels, /ɯ, œ, y/, are in fact more susceptible to perceptual confusions than the neighboring peripheral vowels. Suomi proposes that Turkish vowel harmony helps the perceptually confusable /ɯ, œ, y/ vowels to be accurately identified by keeping vowel backness and rounding constant throughout a word in much of the lexicon, increasing the contextual predictability of successive vowels. Suomi's claim that Turkish vowel harmony is perceptually beneficial for /ɯ œ y/ is also in line with Clements and Sezer's (1982) observation that in Turkish, /ɯ, œ, y/ only rarely occur in disharmonic roots, whereas the vowels /ɑ, e, i, o, u/ freely occur in disharmonic roots.

Following the analyses by Suomi and Clements and Sezer, we consider the hypothesis that our results regarding PMI a predictor of perceptual confusion among vowels might be driven by the perceptually confusable 'non-peripheral' vowel set /ɯ, œ, y/. There are two parts to this hypothesis. First, is that due to their inherent perceptual confusability, vowel harmony might constrain non-peripheral vowels to cooccur with vowels with which they are perceptually less confusable. Second, the negative relationship between PMI and the likelihood of perceptual confusion might be enhanced for non-peripheral vowels compared to peripheral vowels. We test the second part of this proposal in a post-hoc zero-inflated Poisson regression modeling perceptual confusion counts with vowel peripherality (sum coded, non-peripheral = -1, peripheral = 1) and the interaction of PMI and vowel peripherality as additional predictors (Table 5).

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	-0.12	0.17	-0.68	0.50
F1	0.02	0.10	0.25	0.81
F2	0.55	0.13	4.20	< 0.001***
Speaker gender	0.13	0.10	1.33	0.18
Duration	-0.32	0.06	-5.45	< 0.001***
Euclidean distance	-1.45	0.16	-8.86	< 0.001***
Pillai score	0.04	0.12	0.33	0.74
Dissimilarity ratio	-0.42	0.07	-5.93	< 0.001***
PMI	-0.39	0.07	-5.90	< 0.001***
Peripherality	-0.30	0.13	-2.35	0.02*
F2 : Speaker gender	-0.24	0.14	-1.68	0.09
PMI : Peripherality	0.61	0.10	6.15	< 0.001***

Table 5. Post-hoc Poisson regression results.

As in the previous model, all predictor variables were examined for collinearity, suggesting that the predictors including vowel peripherality are independent (all *VIF*s < 3.7; $VIF_{peripherality}$ =

1.34). The results from the post-hoc Poisson regression replicate the results from the previous Poisson regression with significant main effects of stimulus vowel F2, duration, Euclidean distance, dissimilarity ratio, and PMI. In addition to the main effects that are replicated, there is a significant main effect of vowel peripherality (Figure 5) whereby peripheral vowels have a lower rate of perceptual confusion than non-peripheral vowels. This finding supports Suomi's (1983) speculation that non-peripheral vowels are inherently perceptually more confusable. Moreover, there is a significant interaction of PMI and vowel peripherality whereby non-peripheral vowels with high PMI had fewer perceptual confusion responses than non-peripheral vowels with low PMI. In contrast, peripheral vowels with high PMI had more perceptual confusion responses than peripheral vowels with low PMI. The significant interaction of PMI and vowel peripherality confirms our post-hoc hypothesis that the negative relationship between PMI and perceptual confusion counts were driven by non-peripheral vowels. On the other hand, this significant interaction suggests that for peripheral vowels, the relationship between PMI and perceptual confusion counts is in line with our initial hypothesis based on exemplar theory, namely, that greater cooccurrence similarity leads to greater perceived similarity, although this effect is considerably weaker compared to the effect observed in non-peripheral vowels.

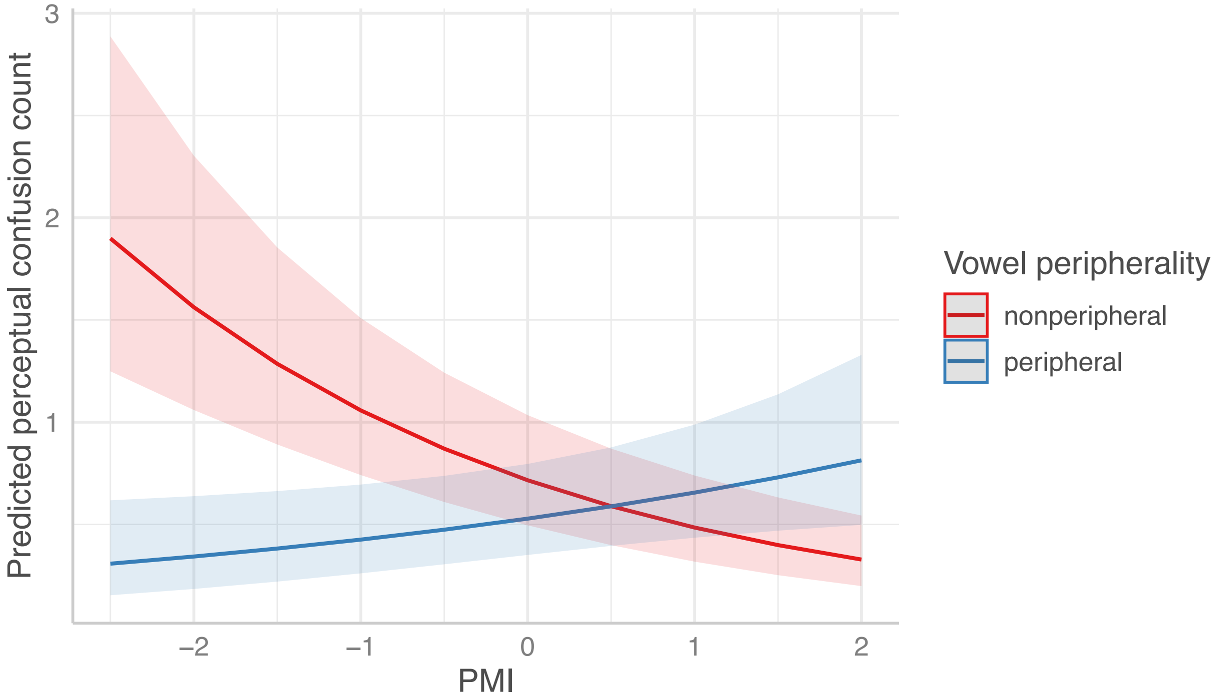

Figure 5. Perceptual confusion counts for stimulus and response vowel pairs as predicted by the interaction of stimulus vowel peripherality and PMI values of the stimulus and response vowel pairs.

Together, these results suggest that the relationship between cooccurrence similarity and perceived similarity is qualified by the inherent phonetic confusability of vowels. That is, for the vowels that are peripheral, i.e., phonetically more extreme and therefore more distinct from one another, cooccurrence similarity enhances perceived similarity, whereas for the non-peripheral vowels that are phonetically less extreme and hence inherently perceptually more confusable, the opposite is true. We suggest that where phonetic and hence perceptual similarity and confusability are already high, greater cooccurrence similarity does not further enhance perceptual similarity, and thus does not increase the likelihood of perceptual confusion in a way that would compromise contrast preservation.

Recall that our main hypothesis in this work stemming from exemplar models and the alternative hypothesis guiding the post-hoc analysis differ in two aspects. First, the main hypothesis posits that there is a positive relationship between the likelihood of vowel cooccurrence and perceived vowel similarity in Turkish whereas the alternative hypothesis posits that there is a negative relationship. Second, the main hypothesis posits that the likelihood of vowel cooccurrence predicts perceived vowel similarity whereas the alternative hypothesis posits that perceived vowel similarity predicts the likelihood of vowel cooccurrence. We argued above that the negative relationship between PMI and perceptual confusion rates might be a result of non-peripheral vowels being constrained by vowel harmony to cooccur with vowels with which they are less confusable. If vowel harmony is perceptually motivated to facilitate the accurate identification of non-peripheral vowels as described above, then we expect lexical cooccurrence patterns that position non-peripheral vowels in words that have vowels which are least likely to be perceptually confused with the non-peripheral vowels. To test this new post-hoc hypothesis, that higher perceptual confusability with a vowel predicts lower likelihood of cooccurrence with that vowel for the non-peripheral Turkish vowels, we ran a linear regression modeling the PMI values of stimulus-response vowel pairs with their perceptual confusion counts in the vowel identification experiment, stimulus vowel peripherality, and their interaction as predictors.

The results from the linear regression (Table 6) reveal a significant main effect of perceptual confusion counts in vowel identification suggesting that vowel pairs with more perceptual confusion responses have lower PMI values. There was also a significant main effect of stimulus vowel peripherality suggesting that peripheral vowels have higher PMI values than non-peripheral vowels. The significant main effect of vowel peripherality is consistent with the observations that in the Turkish lexicon, peripheral vowels are subject to a lesser degree of cooccurrence constraints than non-peripheral vowels (Clements & Sezer 1982). Lastly, there was a significant interaction of perceptual confusion and stimulus vowel peripherality such that higher perceptual confusion predicts lower PMI values for non-peripheral vowels (Figure 6). This significant interaction supports Suomi's (1983) claim that Turkish vowel harmony is perceptually motivated in that non-peripheral vowels are constrained in the Turkish lexicon to cooccur with vowels with which they are perceptually less confusable.

	Estimate	**Std. Error**	**t value**	**Pr(>|t|)**
(Intercept)	-0.84	0.09	-9.03	< 0.001***
Perceptual confusion count	-0.34	0.06	-5.29	< 0.001***
Peripherality	1.00	0.12	8.47	< 0.001***
Perceptual confusion count : Peripherality	0.54	0.14	3.93	< 0.001***

Table 6. Post-hoc regression results.

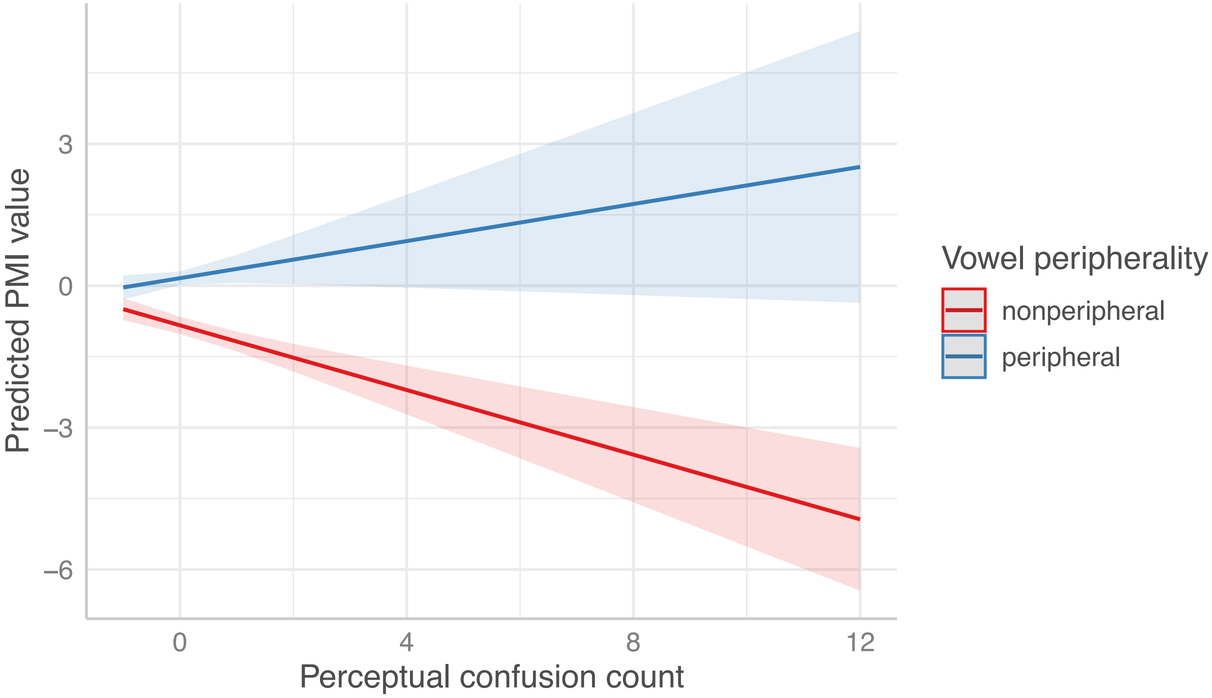

Figure 6. PMI values as predicted by the interaction of stimulus vowel peripherality and perceptual confusion counts.

Together, these results suggest that although the positive relationship hypothesized by exemplar models holds for Turkish vowels classified as peripheral, this relationship does not generalize to all Turkish vowels, suggesting that harmony classes cannot be fully explained as emergent from the cooccurrence similarity-based coactivation of exemplars. In contrast, the results here provide clearer support for the contrast preservation hypothesis: we show that there is an overall negative relationship between the likelihood of vowel cooccurrence and perceived vowel similarity in Turkish and that this relationship is driven by non-peripheral vowels with weaker perceptual contrasts. Following Clements and Sezer (1982) and Suomi (1983), we argue that vowel cooccurrence in the Turkish lexicon is sensitive to the perceptual distinctiveness of vowels and is motivated to facilitate vowel perception by restricting vowel cooccurrence.

5. Conclusions. In this work, we set out to test a hypothesis stemming from an exemplar theory understanding of vowel harmony, that greater pairwise vowel cooccurrence would lead to greater perceived vowel similarity. Our analysis of the Turkish vowel identification data here reveals that for vowels that are peripheral in the phonetic dimension related to harmony (F2), greater cooccurrence similarity does lead to greater perceived similarity, although this relationship is weak and does not generalize to all Turkish vowels. On the other hand, the opposite pattern is true for non-peripheral vowels, which we argue can be attributed to non-peripheral vowels being inherently susceptible to perceptual confusions. We propose that the negative relationship between cooccurrence similarity and perceived similarity for non-peripheral vowels reflects the effect of vowel harmony on the Turkish lexicon, constraining non-peripheral vowels to cooccur with vowels with which they are perceptually less confusable, as previously argued by Suomi (1983) and Kaun (1995). In a post-hoc analysis, we showed that the perceptual confusions in our Turkish vowel identification data predict vowel cooccurrence patterns in the Turkish lexicon, in particular with non-peripheral vowels being more likely to cooccur with vowels with which they are less perceptually confusable. Hence, although we cannot confirm that greater cooccurrence similarity leads to greater perceived vowel similarity for all Turkish vowels as predicted by exemplar models, we find evidence in support of previous accounts suggesting Turkish vowel harmony is perceptually motivated to increase the identification accuracy of non-peripheral vowels and hence help contrast preservation.

References

Bakay, Özge, Özlem Ergelen, Elif Sarmış, Selin Yıldırım, Atilla Kocabalcıoğlu, Bilge Nas Arıcan, Merve Özçelik, Ezgi Sanıyar, Oğuzhan Kuyrukçu, Begüm Avar, et al. 2021. Turkish WordNet KeNet. In Piek Vossen & Christiane Fellbaum (Eds.), *Proceedings of the 11th global wordnet conference*, 166–174. Global WordNet Association. https://doi.org/978-946402731-0

Baker, Adam C. 2008. *Two Bayesian approaches to finding vowel harmony*. University of Chicago technical report. https://newtraell.cs.uchicago.edu/files/tr_authentic/TR-2009-03.pdf

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. Cognitive Science Society.

Clements, George N., & Engin Sezer. 1982. Vowel and consonant disharmony in Turkish. In Harry van der Hulst & Norval Smith (Eds.), *The structure of phonological representations* (Part II). 213–255. Foris Publications. https://doi.org/10.1515/9783112423325-007

Cole, Jennifer. 2009. Emergent feature structures: Harmony systems in exemplar models of phonology. *Language Sciences, 31*(2–3), 144–160. https://doi.org/10.1016/j.langsci.2008.12.004

Erguvanlı-Taylan, Eser. 2015. *The Phonology and Morphology of Turkish.* Boğaziçi University Press.

Goldrick, Matt, & Jennifer Cole. 2023. Advancement of phonetics in the 21st century: Exemplar models of speech production. *Journal of Phonetics, 99*, 101254. https://doi.org/10.1016/j.wocn.2023.101254

Jackman, Simon. 2024. *pscl: Classes and methods for R developed in the Political Science Computational Laboratory*. https://github.com/atahk/pscl/

Johnson, Keith. 1997. Speech perception without speaker normalization: An exemplar model. In Keith Johnson & John W. Mullennix (Eds.), *Talker Variability in Speech Processing*, 145–165. Academic Press.

Kabak, Barış, Eva Kasselkus, Kazumi Maniwa, & Silke Weber. 2008. *Vowel harmony has direction and context: A corpus study*. The 16th Manchester Phonology Meeting paper presentation.

Kaun, Abigail R. 1995. *The typology of rounding harmony: An Optimality Theoretic approach*. University of California, Los Angeles dissertation. https://doi.org/doi:10.7282/T3R49PM2

Kopkallı-Yavuz, Handan. 2010. The sound inventory of Turkish: Consonants and vowels. In Seyhun Topbaş & Mehmet Yavuz (Eds.), *Communication Disorders in Turkish*, 27-47. Multilingual Matters. https://doi.org/10.21832/9781847692474-007

Kuhl, Patricia K. 1991. Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Perception & psychophysics, 50*(2), 93–107. https://doi.org/10.3758/BF03212211

Pierrehumbert, Janet. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In Joan L. Bybee & Paul J. Hopper (Eds.), *Frequency and the emergence of linguistic structure*, 137–157. John Benjamins Publishing Company. https://doi.org/10.1075/tsl.45.08pie

Terbeek, Dale. 1977. *Working Papers in Phonetics, No. 37: A cross-language multidimensional scaling study of vowel perception*. UCLA: Department of Linguistics working paper. https://escholarship.org/uc/item/3nx4g138

Tokaç, Züheyra. 2025. *Mapping the Turkish vowel space: A comprehensive vowel production study*. The 118th Meeting of the Acoustical Society of America, New Orleans, LA, USA poster presentation. https://doi.org/10.1121/10.0038218

Vago, Robert M. 1973. Abstract Vowel Harmony Systems in Uralic and Altaic Languages. *Language*, *49*(3), 579. https://doi.org/10.2307/412352

van der Hulst, Harry. 2016. Vowel Harmony. In *Oxford Research Encyclopedia of Linguistics*. Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.38

Zeileis, Achim, Christian Kleiber, & Simon Jackman. 2008. Regression models for count data in R. *Journal of Statistical Software, 27*(8), 1–25. https://doi.org/10.18637/jss.v027.i08