**The role of vowel harmony during visual word recognition: Evidence from Turkish**

Vowel harmony is a phonological phenomenon in which vowels within a word share certain phonetic features (e.g., frontness vs. backness harmony, roundness harmony, among others). This phenomenon is present in several families of languages (e.g., Turkic Languages like Turkish and Kazakh; Finno-Ugric Languages like Finnish and Hungarian], among others; see Goldsmith, 1990; Rose & Walker, 2011, for reviews). In these languages, the vowels that constitute a word require similar mouth and lip movements, potentially facilitating speech production (Khalilzadeh, 2010) and speech segmentation (Ketrez, 2013). Furthermore, vowel harmony also affects morphological processes (e.g., suffixes sharing the vowel harmony with the root) and word formation (Van der Hulst, 2016).

An open question is whether vowel harmony, being a phonological phenomenon, also affects word recognition in the visual modality. In the present paper, we examine this issue in Turkish. In Turkish, there are four front vowels (i, ö, ü, e; pronounced as /i/, /ø/, /y/, and /e/, respectively]) and four back vowels (ı, o, u, a, pronounced as /ɯ/, /o/, /u/, and /a/, respectively). When the initial syllable of a word begins with a front vowel, the following syllables should only contain front vowels (e.g., *güven* [trust], *ödül* [award]). Conversely, when the first syllable starts with a back vowel, the subsequent syllables should only contain back vowels (e.g., *karar* [decision], *yakın* [near]). Historically, all words in old Turkic were harmonious. However, the evolution of the Turkic languages, together with the influence of other cultures across time (e.g., via words from Arabic, Persian, and, more recently, Western languages), has reduced this proportion (Bıyıklı, 2020; Harrison et al., 2002). Nonetheless, modern Turkish still displays a well-defined pattern of vowel harmony: around 75% of the words follow this principle (Harrison et al., 2002). (Footnote 1: Other Turkic languages have essentially lost vowel harmony (e.g., only 53% of words in Modern Uzbek are harmonious; Harrison et al., 2022). A similar decreasing pattern of vowel harmony has also occurred in languages from other families such as standard Estonian (Kiparsky & Pakusalu, 2003) and Korean (Sohn, 2001)) Before introducing the rationale of the experiments, we briefly review the literature that has examined the role of vowel harmony in the auditory and visual domains.

In the auditory domain, it has been shown that infants can rapidly pick up vowel harmony patterns by segmenting the syllables into proto-words in a sequence of continuous speech (i.e., vowel harmony provides a salient cue for word segmentation). Critically, this pattern occurs even the infants have never been exposed to a language with vowel harmony (e.g., English; see Mintz et al., 2018). Vowel harmony also plays a role in child-directed speech. For instance, Ketrez (2013) examined a corpus-based study that compared two harmonious languages (Turkish and Hungarian) and two non-harmonious languages (Polish and Persian). In the two harmonious languages, within-word sequences in child-directed speech were more likely to be harmonious than disharmonious, suggesting that languages with vowel harmony provide learners with extra cues for word segmentation during continuous speech (i.e., sequences with a different pattern of vowel harmony would be considered as separate words). Vowel harmony is also a useful segmentation cue for adult speakers. Suomi et al. (1997) asked adult Finnish participants to identify target trisyllabic nonsense words (e.g., *hymy /ˈhymy/*) embedded within spoken sequences such as *pyhymy /ˈpyhymy/* (harmonious string) and *puhymy /ˈpuhymy/* (disharmonious string; /*u*/ is a back vowel and /*y*/ is a front vowel). Accuracy was higher, and response times were faster when the targets occurred in disharmonious than harmonious strings (see also Vroomen et al., 1998, for parallel evidence). Vowel disharmony play a segmentation cue for the Finnish participants (see Kabak et al., 2010; for similar results in Turkish).

Vowel harmony also plays a role in the visual domain. In an eye movement study during sentence reading in Finnish, Bertram et al. (2004) included two types of compound words: the constituents could share the same vowel harmony (e.g., *satuolento* [fairytale creature] the red color is for illustrative purposes to signal the two constituents]), or not (e.g., *sähköasentaja* [electricity expert]; front vowel harmony in the first constituent and back vowel harmony in the second). Bertram et al. (2004) found shorter gaze durations on harmonically dissonant compound words, suggesting that lack of vowel harmony across the constituents would help segmentation in compound words. Recently, Perea et al. (2022; Experiment 3) examined whether vowel harmony could modulate lexical processing in Finnish using monomorphemic words using a lexical decision task. As vowel disharmony only occurs in a very small subset of loan words in Finnish, Perea et al. (2022) only focused on the role of vowel harmony for pseudowords. They compared harmonious pseudowords (e.g., *HÖPEÄ*) and harmonious pseudowords (i.e., containing both front and back vowels, *VÖURIO*). They found faster and more accurate “no” responses to disharmonious than harmonious pseudowords and suggested that vowel disharmony would make the items less wordlike. One limitation of this experiment is that, because of the characteristics of Finnish words, their manipulation of vowel harmony only involved pseudowords (i.e., all words were harmonious). To reach firm conclusions on the role of vowel harmony during visual word recognition, it is necessary to compare harmonious and disharmonious words in a language that, while showing vowel harmony, also has a non-negligible number of disharmonious words. Turkish is an ideal language in this respect; as stated earlier, around one in four words are disharmonious, thus allowing us to select well-matched sets of harmonious and disharmonious words.

The main goal of the present study was to examine the impact of vowel harmony on lexical access in visually presented Turkish words. To that end, we employed the most common task in visual word recognition (lexical decision; see Balota et al., 2006; Dufau et al., 2012). It has been hypothesized that visual word recognition involves the integration of various codes, including both orthographic and phonological codes (e.g., phonological coherence hypothesis; see Frost, 1998; Frost, Katz, & Bentin, 1987; Ratcliff et al., 2004; Van Orden & Goldinger, 1994). In the context of a lexical decision task, it has been suggested orthographic and phonological codes from the visual stimuli would combine into a single measure of *coherence (*Van Orden & Goldinger, 1994). The amount of coherence can be translated into a parameter of quality of evidence in mathematical models of lexical decision (e.g., diffusion model; Gomez, 2012; Ratcliff et al., 2004). The higher degree of coherence elicited by the orthographic and phonological codes of a printed word, the higher the values of quality of evidence for a “yes” response. Under the assumption that a phonological constraint such as vowel harmony contributes to the formation of coherent phonological codes (see Berent et al., 2001, for evidence of phonological constrain in lexical decision experiments), harmonious words would reach a higher level of coherence than disharmonious words—for the latter there could be some cues of mismatching phonological cues (see Van Orden & Goldinger, 1994, for a similar reasoning with English irregular words).

In Experiment 1, we selected two types of monomorphemic Turkish words: harmonious (i.e., containing only front or back vowels; e.g., *temel* [basis] front and *orman* [forest] back) and disharmonious (i.e., containing both front and back vowels; *kalem* [pen] or *vakit* [time]) matched in word frequency and orthographic neighborhood. The predictions are clear. If vowel harmony reflects the internalization of phonological rules in the mental lexicon of Turkish speakers, the phonological codes of the words following vowel harmony would resonate more as “words” than those from the words that do not follow this principle. In this scenario, we would expect faster word identification times for harmonious than for disharmonious words. Alternatively, if vowel harmony does not serve as a marker for word recognition in Turkish, given that non-negligible proportion of disharmonious Turkish words (around one of four), one would expect very similar response times and error rates for harmonious and disharmonious words. For comparison purposes, we employed the same manipulation for the set of pseudoword foils (i.e., pseudowords could be harmonious or disharmonious). While the predictions for pseudowords in lexical decision are less precise (see Dufau et al., 2012, for discussion on the mechanisms for “no” decisions), one could argue that disharmonious pseudowords would resonate less than harmonious pseudowords. If so, one would expect faster “no” responses (or more errors) for disharmonious than harmonious pseudowords—note that this pattern would replicate the experiment of Finnish conducted by Perea et al.’s (2022; Experiment 3). As Experiment 2 was designed after knowing the findings of the initial experiment, we prefer to introduce its rationale later.

**Experiment 1**

**Method**

The sampling, method, procedure, and analysis strategy were pre-registered at <https://osf.io/ep3gx>.

*Participants*

Thirty-six native speakers of Turkish participants took part in the experiment (M = 26.9 years, SD = 7.72)—this yielded 2,556 observations per condition, which is above the criteria suggested by Brysbaert and Stevens (2008) for small-sized effects. The participants were recruited from online settings (via social networks) and had normal/corrected vision. None of them reported having any speech/reading problems. They all signed an informed consent form. The Experimental Research Committee of the University of Valencia approved the experiment.

*Materials*

We selected 142 nouns of 4 to 6 letters (M = 4.97, SD = 0.81) from an online blog corpus in Turkish [WordLex, Gimenes & New, 2015]. Half of these nouns were harmonious (all vowels were from the same type, front or back, for instance, *sanat* [art]). They had an average frequency of 135,84 occurrences per million words (MeanZipf = 4.941, SDZipf = 0.413) and a mean OLD20 distance of 1.604 (SDOLD20 = 0.284). The other half of words were disharmonious (i.e., words contained both front and back vowels, for instance, *zafer* [victory]), with an average frequency of 112,66 occurrences per million words (MeanZipf = 4.845, SDZipf = 0.416) a mean OLD20 value of 1.654 (SDOLD20 = 0.257). Both sets of words were matched in both word-frequency and orthographic neighborhood (both *p*s > .25). The manipulation for the target nonwords was the same (harmonious pseudowords vs disharmonious pseudowords). These pseudowords were generated via Wuggy (Keuleers & Brysbaert, 2010). The set of stimuli is presented in the OSF link XXXXX.

*Procedure*

The experiment was implemented by PsychoPy (Peirce et al., 2019), and Pavlovia.org (Open Science Tools, Nottingham, UK) was used as the experiment platform. The typical lexical decision task procedure was followed. Participants were asked to determine whether the string was a meaningful word. There was a brief practice session, consisting of 16 trials, before the experiment. Each trial began with a fixation cross with 500ms. While participants had unlimited time to respond in practice rounds, in experimental trials, after the presentation of the target (either word/pseudoword with harmony/disharmony), they had 2 seconds to respond. Every 120 trials, participants were asked to take a short break before continuing the experiment.

*Data analysis*

The inferential analyses were conducted using Bayesian linear mixed-effects models with the *brms* (Bürkner, 2017, 2018, 2019) package in RStudio (R Core Team, 2022). Separate analyses were carried out for words and nonwords. In both cases, the only fixed factor was the type of word (harmonious vs. disharmonious), and the random-effect structure was the maximal allowed by the design (i.e., random intercepts and slopes of Type of word for subjects, and random intercepts of Type of word for items). The exgaussian family was utilized to model reaction time and the Bernoulli family was employed for the binary, accuracy data. The model involved four chains and 10,000 iterations (the warm-up phase consisted of 2000 iterations).

**Results and Discussion**

Reaction times shorter than 250 ms were removed from the latency data (7 observations, 0.06% of trials). A response deadline of 2000 ms was set, and the program automatically classified any response exceeding this timeframe as an error (37 observations, 0.33% of trials). No participants' accuracy fell below the pre-registered threshold of 75%.

Table X provides the descriptive statistics associated with response times and accuracy for both words and nonwords within the vowel harmony conditions.

**Table x.** Means and standard errors (inside of the brackets) of reaction time and accuracy for word conditions and vowel harmony conditions.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | *RT* | *Accuracy* |
| *Word* | |  |  |
|  | Harmonious | 604 (3.2) | 0.974 (0.003) |
|  | Disharmonious | 623 (3.3) | 0.973 (0.003) |
| *Pseudoword* | |  |  |
|  | Harmonious | 685 (3.5) | 0.946 (0.004) |
|  | Disharmonious | 686 (3.6) | 0.943 (0.005) |

*Words*

The reaction times were, on average, 19 ms faster for harmonious than disharmonious words (*b* = 10.36, 95% CrI [2.79, 17.87]). We found no evidence of a vowel harmony effect in the accuracy data (*b* = -0.27, 95% CrI [-0.92, 0.33]).

*Pseudowords*

As shown in Table 1, the reaction times and accuracy were similar for disharmonious pseudowords and harmonious pseudowords (response times: *b* = 0.79, 95% CrI [-8.80, 10.35], accuracy: *b* = -0.13, 95% CrI [-0.66, 0.42]).

Following the preregistration protocol, we also conducted exploratory analyses that examined the dynamics of the effect vowel harmony over time via delta plots (see De Jong et al., 1994). The logic is that, for words, a measure of coherence in a lexical decision task could be interpreted as quality of evidence in a diffusion model (drift-rate parameter; Ratcliff et al., 2004, Van Orden & Goldinger, 1994). In this scenario, the response time distribution of harmonious words would be less skewed (due to the geometry of the diffusion model) than the response time distribution of disharmonious word. If so, the effect of vowel harmony for words should increase across quantiles—as also happens with variables like word-frequency (Ratcliff et al., 2004). To examine this pattern, we computed the reaction times for five quantiles (.1, .3, .5, .7, and .9) per participant and condition. Then, we averaged these reaction times across participants to obtain "vincentiles”, for each quantile, we computed the difference (disharmonious – harmonious) across quantiles for those vincentilized response times. We then plotted these points, with the reaction times on the x-axis and the deltas on the y-axis. As an illustration, a flat line around 20 ms would indicate a shift in reaction time distributions—typically signaling a difference in headstart processing, whereas an ascending trend across quantiles would suggest differences in the quality of evidence rates across conditions (see Fernández-López et al., 2022; Gomez et al., 2013).

A graph showing the difference between words and nonwords

Description automatically generated**Figure X.** Delta plots (differences between disharmonious and harmonious condition in 1. 3. 5. 8. and .9 quantile) for both words and pseudowords.

For words, the delta plot in the left panel of Figure XXX reflects an increasingly larger effect in the higher quantiles. This pattern is consistent with the idea that vowel harmony affects the quality of information during lexical decision (see Gomez et al., 2013). For nonwords, the delta plot was approximately centered RT around zero.

In sum, readers can identify faster harmonious than disharmonious words in Turkish, revealing that this vowel harmony affects lexical access. While the critical finding in this experiment corresponds to the lexical items (i.e., the words), we believe that it is relevant to examine the potential reasons for the apparent discrepancy between the lack of an effect of vowel harmony for pseudowords in the present experiment and in the experiment reported by Perea et al.’s (2022) in Finnish. There were two basic differences. On the one hand, the set of words in the Finnish experiment was always harmonious. Thus, lack of vowel harmony in the Finnish would have signalled that the stimulus was not a word, leading to faster and more accurate “no” responses to disharmonious pseudowords. On the other hand, in Turkish, around 25% of words are not harmonious, which is dramatically smaller in Finnish for which disharmonious words are extremely low. Thus, Turkish readers may not use “vowel disharmony” as a reliable marker for “no” responses. To tell apart these two explanations, we designed Experiment 2. As in the Perea et al. (2022) experiment, all words were harmonious. The set of pseudowords was the same as in Experiment 1, except that all words were now harmonious. If the cause of the discrepancy between the vowel harmony effect for pseudowords between the current experiment and Perea et al.’s (2022) Experiment 3 was only the experimental design, one would expect an effect of vowel harmony for pseudowords when only using harmonious words in the set of stimuli. Alternatively, if vowel disharmony is not used as a marker for “no” responses in Turkish, one would expect similar response times and error rates for harmonious and disharmonious pseudowords in Experiment 2.

**Experiment 2**

**Method**

This sample procedure and analysis were pre-registered at <https://osf.io/25wsh>.

*Participants*

A new sample of thirty-six Turkish participants from the same population as Experiment 1 participated in the experiment (Mage = 26.94, SDage = 8.74). None of them reported having any speech/reading problems.

*Materials*

We employed the same set of pseudowords as in Experiment 1. As for the word stimuli, we kept the 71 harmonious nouns from Experiment 1 and replaced the disharmonious nouns with other harmonious nouns of similar word frequency from the same corpus. Overall, the mean frequency of the words in Experiment 1 and 2 were similar (Zipfexp1 = 4.845, Zipfexp2 = 4.906, *t* = 0.302, p > .30).

*Procedure*

The procedure is the same as Experiment 1.

*Data analysis*

Reaction times shorter than 250 ms were disregarded (148 observations, 1.27% of trials). There were no participants whose accuracy fell below the threshold of 75%.

The data analysis method is the same as Experiment 1.

*Results*

Error responses removed from the latency and accuracy analyses(612 observations, 5.25% of trials). Table X provides the descriptive statistics associated with response times and accuracy for the nonwords within the vowel harmony conditions.

**Table x.** Means and standard errors (inside of the brackets) of reaction time and accuracy for pseudowords and vowel harmony conditions.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | *RT* | *Accuracy* |
| *Pseudoword* | |  |  |
|  | Harmonious | 658 (3.51) | 0.950 (0.004) |
|  | Disharmonious | 652 (3.58) | 0.957 (0.004) |

Note: For Words: mean RT: 581 ms, standard error; 2.22, Accuracy: 0.955

The analyses of the latency data showed slightly faster responses (around 6 ms, on average) for disharmonious pseudowords than harmonious pseudowords (effect; 6 ms, *b* = -9.06, 95% CrI [-17.73, -0.35]). The accuracy rate did not show any trends of an effect. the same for disharmonious pseudowords compared to harmonious pseudowords (effect; 0.007, *b* = 0.19, 95% CrI [-0.22, 0.60]).

*Delta plots*

**Figure X.** Delta plot (differences between disharmonious and harmonious condition in 1. 3. 5. 8. and .9 quantile) for pseudowords.



In this case, the delta plot showed a mean reaction time (RT) that was centred slightly below zero, thus revealing that while in a much smaller degree than the parallel experiment in Finnish (749 ms for harmonious pseudowords vs 695 ms for disharmonious pseudowords; Perea et al., 2022), disharmonious pseudowords responded to slightly faster than harmonious pseudowords.

In summary, we observed a vowel harmony effect on pseudowords in the same direction as the Finnish experiment (Perea et al., 2022). While for Finnish readers vowel disharmony is a clear signal for “no” responses, this scenario is quite different for Turkish readers for who vowel disharmony is not salient marker for “no” responses.

**General Discussion**

Vowel harmony is a phonological phenomenon for which the vowels in a word share some properties (e.g., either front vowels or back vowels [backness-frontness harmony]) that occurs in several families of languages (e.g., Turkish [Turkic languages], Finnish [Finno-Ugric languages], among others). Previous research has consistently shown that vowel harmony can be used to help word segmentation in auditory and visual modalities. In the present lexical decision experiments, we investigated whether vowel harmony also exerts a role during visual word recognition.

We chose Turkish as the target language because, while showing vowel harmony, it contains a non-negligible proportion of disharmonious words (around one of four words; Harrison et al., 2002). We hypothesised that if vowel harmony contributes the formation of stable orthographic-phonological codes, harmonious words (e.g., *iğne* [needle], both vowels are front vowels) would lead to a higher degree of coherence than disharmonious words (e.g. *eşya* [stuff] contains a back vowel and a front vowel), thus eliciting faster “yes” responses than disharmonious words. Results from Experiment 1 effectively showed faster response times for harmonious than disharmonious words. Notably, the effects of vowel harmony were negligible for pseudowords, suggesting that lack of vowel harmony is not a relevant index to signal “no” responses in lexical decision in Turkish—we only found a small effect of vowel harmony for pseudowords when the set of word stimuli only contained harmonious words (Experiment 2)

The presence of faster lexical decision responses to harmonious than disharmonious words aligns well with the global coherence hypothesis (Van Orden, 1987). In this theory, the more coherent the phonological codes the higher the quality of information driving the lexical decision responses. Assuming that “word” responses are driven by a measure of degree of match between the visual input and the lexical entries in memory that reflect the quality of evidence (Ratcliff et al., 2004), harmonious words would elicit a greater degree of match than disharmonious words. Notably, in this type of models, the predictions apply not only to the mean response times, but also to the distribution of the response times. Specifically, the response time distribution of the words with a higher quality of evidence (i.e., harmonious words) would reflect less skewed distributions than the response times distributions of the words with a lower degree of evidence. This is exactly the pattern we found for word stimuli in Experiment 1 (see Figure XXXX). The effect of vowel harmony increased across quantiles, reflecting a more skewed distribution for the disharmonious pseudowords. Thus, vowel harmony appears to affect the degree of “wordness” in lexical decision. An exploratory analysis using the diffusion model (Ratcliff et al., 2024) with the D\*M package (van den Bergh et al., 2020) confirmed this pattern: drift-rates (a measure of quality of evidence) was higher for harmonious than for disharmonious words (see Supplementary Materials XXXX).

The effect of vowel harmony was much minimal or null for pseudowords—we only found a very small 6-ms advantage for disharmonious than harmonious pseudowords in Experiment 2, when all words were harmonious. While this latter pattern was the expected by the global coherence hypothesis (i.e., vowel disharmony would signal a “no” response to a larger degree than vowel harmony), the effect was much smaller than a parallel experiment in Finnish (Perea et al., 2021)—note that unlike Turkish, vowel disharmony in monomorphemic words is extremely reduced to a small set of loan words. The dissociation of the effects of vowel harmony for words and pseudowords in lexical decision favor the idea that “no” responses in the lexical decision task may be driven by a variety of mechanisms (see Dufau et al., 2012) and they may not reflect the same mechanisms for “yes” and “no” responses.

Clearly, the sizeable effect of vowel harmony during the recognition of visually present words suggests a complex interplay between orthographic and phonological processes on how the brain encodes lexical information. While word recognition times in the present study provide information on the nature of the effect of vowel harmony, showing an increasing effect with time consistent with the idea of resonance due to stable phonological codes in the word recognition system, further studies should investigate the impact of vowel harmony using single-word identification techniques that rely on the time dynamics such as the recording of event-related potentials (Grainger & Holcomb, 20XX). Another important avenue is to test the role of vowel harmony in more ecological setting, embedding harmonious and disharmonious words as target words in sentences and registering the readers’ eye movements (e.g., Rayner et al., 2013).

In summary, vowel harmony plays a role on the identification of visually presented words (i.e., harmonious words are identified more rapidly than disharmonious words) in a task that does not have an overt phonological component (lexical decision) in Turkish. At a theoretical level, our findings align with those models that assume that orthographic and phonological codes are integrated during lexical access (Stone & Van Orden, 20XX). At a methodological level, the presence of sizeable effects of vowel harmony for words, but not for pseudowords suggest that “yes” and “no” lexical decision responses may be driven by partly different mechanisms (see Dufau et al., 2012).

(This is a footnote if anything…Addition to our previous discussion about the differ between Finnish experiment (Perea et al., 2022, Experiment 3) and current Experiment 2, there is one methodological difference. As it is not possible to manipulate the vowel harmony on words in Finnish, Perea et al manipulated the pseudowords by words. That means that, they transform words to pseudowords in order to manipulate vowel harmony with more word-likely pseudowords (e.g., *HÖPEÄ—*harmoniousvs*. HOPEÄ—*disharmonious [baseword: *HÄPEÄ*]). Whereas we generate pseudowords by Wuggy, and we control by

**References**

Bertram, R., Pollatsek, A., & Hyönä, J. (2004). Morphological parsing and the use of segmentation cues in reading Finnish compounds. *Journal of Memory and Language*, *51*(3), 325–345. <https://doi.org/10.1016/j.jml.2004.06.005>

Bıyıklı, M. (2020). Turkish glossary production and meaning in the inherit of Turkish culture: Comparative analysis of seventeen Turkish language dictionaries. Manas Sosyal Araştırmalar Dergisi, 9(4), 2051-2066.

Bürkner P (2017). “brms: An R Package for Bayesian Multilevel Models Using Stan.” Journal of Statistical Software, 80(1), 1–28. doi:10.18637/jss.v080.i01.

Bürkner P (2018). “Advanced Bayesian Multilevel Modeling with the R Package brms.” The R Journal, 10(1), 395–411. doi:10.32614/RJ-2018-017.

Bürkner P (2021). “Bayesian Item Response Modeling in R with brms and Stan.” Journal of Statistical Software, 100(5), 1–54. doi:10.18637/jss.v100.i05.

Gimenes, M.. & New. B. (2015). Worldlex: Twitter and blog word frequencies for 66 languages. Behavior research methods, 1 – 10.

Goldsmith, J. A. (1990). *Autosegmental and metrical phonology.* Blackwell.

Harrison, D. K., Dras, M., & Kapicioglu, B. (2002). Agent-Based Modeling of the Evolution of Vowel Harmony. *North East Linguistics Society*, *32*(1).

Inhoff, AW, Solomon, M., Radach, R. ve Seymour, BA (2011). Temporal dynamics of the eye–voice span and eye movement control during oral reading. *Journal of Cognitive Psychology*, 23(5), 543–558. <https://doi.org/10.1080/20445911.2011.546782>

Kabak, B., Maniwa, K., & Kazanina, N. (2010). Listeners use vowel harmony and word-final stress to spot nonsense words: A study of Turkish and French. *Laboratory Phonology*, *1*(1). https://doi.org/10.1515/labphon.2010.010

Katz, L., & Frost, R. (1992). Chapter 4 the reading process is different for different orthographies: The orthographic depth hypothesis. *Advances in Psychology*, 67–84. <https://doi.org/10.1016/s0166-4115(08)62789-2>

Ketrez, F. N. (2013). Harmonious cues for speech segmentation: A cross-linguistic corpus study on child-directed speech. Journal of Child Language, 41(2), 439–461. https://doi.org/10.1017/s0305000912000724

Keuleers, E., Brysbaert, M. Wuggy: A multilingual pseudoword generator. Behavior Research Methods 42, 627–633 (2010). <https://doi.org/10.3758/BRM.42.3.627>

Kiparsky, P. & Pakusalu, L. (2003). Towards a typology of disharmony. The Linguistic Review, 20, 217-241. <https://doi.org/10.1515/tlir.2003.009>

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. Behavior Research Methods, 51(1), 195-203. <https://doi.org/10.3758/s13428-018-01193-y>

Perea, M., Hyönä, J., & Marcet, A. (2022). Does vowel harmony affect visual word recognition? evidence from Finnish. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *48*(12), 2004–2014. <https://doi.org/10.1037/xlm0000907>

Rose, S., & Walker, R. (2011). Harmony systems. In J. Goldsmith, J. Riggle, & A. Yu (Eds.), Handbook of phonological theory (2nd ed., pp. 240–290). Blackwell.

Sohn, H. M. (2001). The Korean language. Cambridge University Press.

van den Bergh, D., Tuerlinckx, F., & Verdonck, S. (2020). DstarM: an R package for analyzing two-choice reaction time data with the D∗ M method. Behavior research methods, 52, 521-543.

Van der Hulst, H. (2016). Vowel harmony. In *Oxford research encyclopedia of linguistics*.

Van Orden, G. C., & Goldinger, S. D. (1994). Interdependence of form and function in cognitive systems explains perception of printed words. Journal of Experimental Psychology: Human Perception and Performance, 20(6), 1269.

Vroomen, J., Tuomainen, J., & de Gelder, B. (1998). The roles of word stress and vowel harmony in speech segmentation. *Journal of Memory and Language*, *38*(2), 133–149. <https://doi.org/10.1006/jmla.1997.2548>

**Supplementary Materials**

Diffusion model

We employed Ratcliff diffusion model (REF) with D star M package (REF). We let the parameters drift rate (*v*) and encoding time (*Ter*) free between 4 conditions (harmonious vs disharmonious x word vs pseudoword). All other parameters kept constant across conditions.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | *a* | *v* | *Ter* | *z* | *sz* | *sv* | *st* |
|  |  | 1,322 |  |  | 0,548 | 0,396 | 0,934 | 0,171 |
| Words | |  |  |  |  |  |  |  |
|  | Harmonious |  | 3,646 | 0,341 |  |  |  |  |
|  | Disharmonious |  | 3,281 | 0,346 |  |  |  |  |
| Pseudowords | |  |  |  |  |  |  |  |
|  | Harmonious |  | -3,388 | 0,384 |  |  |  |  |
|  | Disharmonious |  | -3,363 | 0,381 |  |  |  |  |

According to the diffusion model results, the difference in *Ter* between vowel harmony conditions for both words and pseudowords were similar. As expected, harmonious words has higher drift rate than disharmonious words (36.5 ms), which is fully consistent with the global coherence hypothesis and all other results we presented before. This means that evidence accumulation from harmonious words were faster and clearer than disharmonious words. As can be seen, there is almost no difference between the drift rate of harmonious and disharmonious pseudowords.

**Rebuig**

In the study of Kabak et al. (2010), they asked French and Turkish speakers to detect the nonword targets presented in written form (e.g., *paVO*) inside of the 5-syllable auditory nonsense spelling letter combinations (harmoniously matched [*goLUshopaVO*] or mismatched [*gøLYshøpaVO*]). It was observed that when the target and syllable are mismatched harmoniously, Turkish speakers demonstrate more accurate detection. Unlike French speakers, Turkish speakers were more sensitive to vowel sequences in word segmentation.