Testing variability effects in Hungarian vowel harmony

Fanni Patay[1], Ágnes Benkő[1,2], Ágnes Lukács[3], Péter Rebrus[1] & Miklós Törkenczy[1,2]
[1]MTA Research Institute for Linguistics / [2]Eötvös Loránd University / [3]Department of Cognitive Science, Budapest University of Technology and Economics, MTA-BME Lendület Language Acquisition Research Group

Hungarian backness harmony shows various degrees of transparency and variation, but the empirical testing of these variability effects in corpora is problematic because of data sparseness. We have created an experiment using harmonically mixed stems and four different harmonic suffixes, and collected information about the variants from native speakers in the form of a sentence completion task. We show that there are significant differences between stem types, and that the harmonic suffix can also affect the behaviour of the stem. Our results confirm that native speakers can learn unnatural patterns and that they obey the Law of Frequency Matching (Hayes et al. 2009).

Keywords: vowel harmony, variation, vacillation, neutral vowels, experimental phonology

1. Harmony

The backness harmony system of Hungarian (henceforward HVH) is both determined and underdetermined by phonology. This manifests itself in variation where some stems that are identical in their phonological properties relevant in harmony behave in more than one way harmonically. This may be lexical variation, where different stems of the same phonological shape belong to different harmonic classes (e.g. **hɑvɛr-ok** 'mate-PL' vs. **kontsɛrt-ɛk** 'concert-PL'), or vacillation, where the same stem may take front or back alternants of the same suffix (e.g. **fotɛl-ok** / **fotɛl-ɛk** 'armchair-PL'). In this paper, we focus on the latter type of variation. Vacillation is confined to a phonologically identifiable 'zone of variation' (Hayes & Cziráky Londe 2006): it can occur after harmonically mixed stems whose final portion consists of a back-vowelled syllable followed by one

https://doi.org/10.1075/atoh.16.05pat
© 2020 John Benjamins Publishing Company

or more syllables whose vowel is a neutral one, schematically [BN$^+$].[1] Vacillation is motivated by two effects which modify the basic pattern of neutral-vowel transparency in HVH. While the neutral vowels /i/ and /iː/ are totally transparent ([Bi]B, e.g. **forint-nak** 'florin-DAT'), the other neutral(-like) vowels are variably transparent: [Be]B/F (e.g. **arzeːn-nak/nɛk** 'arsenic-DAT') and [Bɛ]F/B (e.g. **hotɛl-nɛk/nak** 'hotel-DAT'). This is called the Height Effect. An additional effect obtains when two neutral vowels follow the back vowel: [BNN]. In this case either variable transparency or non-transparency occurs: e.g. [Bii]B/F (**aspirin-nak/nɛk** 'aspirin-DAT') or [Biɛ]F (**sanitɛr-nɛk** 'sanitary ware-DAT'). This is called the Count Effect. The two effects are shown schematically in (1), where x < y means: 'y is more front than x'.

(1)[2] a. Height Effect: [Bi(ː)] < [Beː] < [Bɛ]
papiːr-ok < **taːɲeːr-ok, sateːn-ɛk/ok** < **havɛr-ok, fotɛl-ɛk/ok, kontsɛrt-ɛk**

b. Count Effect: [BN] < [BNN$^+$]
parti-ra < **alibi-ra/re, horribiliʃ-ra/rɛ**
taːɲeːr-ra < **klarineːt-rɛ/ra, proteːziʃ-rɛ/ra**

These effects were known in the literature, and were typically impressionistically described based on intuition/introspection (e.g. in Vago 1980, Siptár & Törkenczy 2000; see Törkenczy 2016 for an overview) before they were named by Hayes et al. (2006, 2009), but the latter are the first empirical studies[3] based on extensive corpus-based research and psycholinguistic experiments. Hayes and Cziráky Londe (2006) has shown that the Height Effect and the Count Effect (i) manifest themselves in type frequency, measured in the ratio of back/front suffixed forms to all harmonically suffixed forms in a corpus; and that (ii) native speaker reaction, based on wug testing, matches the results of the corpus study. That is, they obey what Hayes et al. (2009) call the "Law of Frequency Matching" and define as the state of affairs when "[s]peakers of languages with variable lexical patterns respond

1. B, F and N are back, front and neutral vowels, respectively; [and] mark the edges of stems. Some vowels that are neutral in roots and invariable suffixes may also occur in the front alternants of harmonically alternating suffixes; in this last case we have encoded them as front (F), e.g. dative **-nɛk** (**-nak**), adessive **-neːl** (**-naːl**). Unless otherwise indicated, in order to avoid unnecessary clutter we suppress the consonants, the non-final portions of stems and the length mark 'ː' in formulas throughout the paper.

2. Glosses: **papiːr-ok** 'paper-PL', **taːɲeːr-ok** 'plate-PL', **sateːn-ɛk/ok** 'satin-PL', **kontsɛrt-ɛk** 'concert-PL', **parti-ra** 'party-SUBL', **alibi-ra/re** 'alibi-SUBL', **horribiliʃ-ra/rɛ** 'horrible-SUBL', **klarineːt-rɛ/ra** 'clarinet-SUBL', **proteːziʃ-rɛ/ra** 'prosthetic-SUBL'.

3. Ringen & Kontra (1989) is a notable exception in the earlier literature.

stochastically when tested on such patterns. Their responses aggregately match the lexical frequencies" (p. 826). However, Hayes and Cziráky Londe (2006) presents a simplified picture in several respects (e.g. they examined word forms with one and the same harmonic suffix (the dative) only). There is every indication that the patterning of variation and its conditioning in the zone of variation in Hungarian is far richer and is conditioned by many and varied factors, which include syllable count ([BN] vs. [BBN] stems, cf. Ringen & Kontra 1989), the place and manner of articulation of stem-final consonants (cf. Hayes et al. 2009), the phonological shape of the harmonic suffix (C-initial vs. V-initial, cf. Rebrus & Törkenczy 2013), paradigm uniformity ([[BN]N] vs. [BNN] stems, cf. Rebrus & Törkenczy 2015, Rebrus & Szigetvári 2016), the fuzziness of harmonic domain boundaries due to the gradience of morphological complexity (cf. Rebrus & Törkenczy 2017) and the multiple application of the Height Effect (stems in which the Height Effect and the Count Effect can combine: [BN_xN_y] vs. [BN_zN_w] stems, cf. Rebrus & Törkenczy 2016). In this study, we take a closer look at the last one of these factors.

It has not been sufficiently explored what the *combined effect* of the Count and the Height Effects is for those stems that end in a back vowel followed by two neutral vowels (BNN-stems). Rebrus and Törkenczy (2015, 2016) quantify these two effects in terms of a measure of variability, the Frontness ratio (F-ratio).[4] The F-ratio of a stem class is the ratio of the number of front suffixed forms to the number of all harmonically suffixed forms whose stems belong to the relevant stem-class. The F-ratio is measured in type frequency, i.e. the number of different word-forms (as opposed to tokens) is counted:

(2) The Frontness Ratio

$$\text{F-ratio} = \frac{\text{number of front suffixed forms}}{\text{number of front suffixed forms} + \text{number of back suffixed forms}}$$

The F-ratio increases (i.e. the transparency decreases) between the relevant forms as defined by the Height and the Count Effects. Because of the Height Effect, the transparency of the neutral vowel in [Bi] stems is greater than in [Be] stems, and the same holds between [Be] and [Bɛ] stems; expressed in F-ratios, this is [Bi] < [Be] < [Bɛ]. In accordance with the Count Effect, transparency decreases in [BNN^+] stems compared to [BN] stems. Minimally,[5] this means that if one of the

4. Compare Hayes & Cziráky Londe (2006) for a similar measure.

5. Since we do not know the relative strengths of these effects, we cannot tell whether the Count Effect applies independently of the identity of the neutral vowels involved or not, i.e. what the relationship is between [BNN^+] stems and [BN] stems when they do not share a neutral vowel, e.g. [Be] vs. [Bii].

neutral vowels is fixed while the other varies, transparency decreases in accordance with the Height Effect, e.g. [Bi] < [Bie] (fixed N_1) and e.g. [Bɛ] < [Beɛ] (fixed N_2).

BNN-stems are a context for the combined application of the Height Effect and the Count Effect. Given the three neutral vowels /i(ː)/, /eː/ and /ɛ/, there are nine types of BNN-stems, shown in (3).[6]

(3) Types of BNN-stems[7]

	i(ː)	eː	ɛ
i(ː)	[Bii] **alibi**	[Bie] **klarineːt**	[Biɛ] **kabinɛt**
eː	[Bei] **proteːziʃ**	[Bee] **ateːneː**	[Beɛ] **konteːnɛr**
ɛ	[Bɛi] **bakɛlit**	[Bɛe] **suvɛreːn**	[Bɛɛ] **kompɛtɛnʃ**

Rebrus & Törkenczy (2016) argue that in BNN-stems, which are subject to the Height Effect and the Count Effect, there are two additional effects, Cumulativity and Locality.

Cumulativity means that since we have two neutral vowels in BNN-stems, the Height Effect applies twice: (i) for the second neutral vowel with a fixed quality of the first one, and (ii) for the first neutral vowel with a fixed quality of the second one. Cumulative interaction between the two neutral vowels in a [BN_1N_2] environment is defined as in (4) (where *x, y, z* are neutral vowels). In the definition, the ordering "≤" (which allows equality (or near-equality) of F-ratios) is used instead of the strict ordering "<" because some BNN-classes have F-ratios that are very close and nearly equal to 1 (which is the maximal possible value of an F-ratio).

(4) Cumulativity
  i. Height Effect for N_2: if [Bx] ≤ [By] then [Bzx] ≤ [Bzy]
  ii. Height Effect for N_1: if [Bx] ≤ [By] then [Bxz] ≤ [Byz]
  iii. transitivity: if [Bx_1x_2] ≤ [By_1y_2] and [By_1y_2] ≤ [Bz_1z_2] then [Bx_1x_2] ≤ Bz_1z_2]

The ordering relation defined in (4) yields 27 different ordered pairs of the 9 possible BNN sequences (where only the Ns vary between the 3 different values). This is shown in (5), where ordered pairs are connected by arrows, the direction of an arrow corresponds to the ordering ≤, and ordering by transitivity (4iii) is left unindicated to avoid clutter.

6. No distinction is made between long iː and short i. We only consider [BNN^+] stems with exactly two neutral vowels since those with longer N sequences are extremely rare (e.g. **kompatibiliʃ** 'compatible').

7. Glosses: **alibi** 'alibi', **klarineːt** 'clarinet', **kabinɛt** 'cabinet', **proteːziʃ**, 'prosthetic', **ateːneː** 'Athena', **konteːnɛr** 'container', **bakɛlit** 'bakelite', **suvɛreːn** 'sovereign', **kompɛtɛnʃ** 'competent'.

(5) Ordering by Cumulativity

$$\begin{tikzcd}
{[Bii]} \arrow[r] \arrow[d] & {[Bie]} \arrow[r] \arrow[d] & {[Biɛ]} \arrow[d] \\
{[Bei]} \arrow[r] \arrow[d] & {[Bee]} \arrow[r] \arrow[d] & {[Beɛ]} \arrow[d] \\
{[Bɛi]} \arrow[r] & {[Bɛe]} \arrow[r] & {[Bɛɛ]}
\end{tikzcd}$$

In (5), the stem classes that are greater in F-ratio are those which are to the right and/or down, and those that are smaller are to the left and/or up. The other pairs (i.e. those that are to the right and up or are to the left and down) are not defined to be in relation by Cumulativity.

Rebrus and Törkenczy (2016) has found on the basis of the frequencies (F-ratios of BNN-stems) in the Szószablya webcorpus (Halácsy et al. 2004) that Cumulativity holds true of the interaction of the three neutral vowel qualities.[8]

Rebrus and Törkenczy (2016) propose that another effect, Locality, applies to BNN-stems. Locality is the dominance of the Height Effect of N_2, the neutral vowel closer to the suffix in a $[BN_1N_2]$ environment. It is defined as in (6), (where *x, y* are neutral vowels):

(6) Locality
If [B*x*] ≤ [B*y*] then [B*yx*] ≤ [B*xy*]

Locality introduces three further orderings: [Bɛi] ≤ [Biɛ], [Bɛe] ≤ [Beɛ], [Bei] ≤ [Bie]. As can be seen in (7), the F-ratios of BNN-stems in the webcorpus reflect the ordering by Locality in the first two cases, but not in the last one: [Bɛi] (0.726) ≤ [Biɛ] (0.987), [Bɛe] (0.864) ≤ [Beɛ] (1.000) but [Bei] (0.674) ≰ [Bie] (0.579):

(7) F-ratios of BNN-stems

	i(:)	e:	ɛ
i(:)	**0.560**	**0.579**	**0.987**
e:	**0.674**	**0.931**	**1.000**
ɛ	**0.726**	**0.864**	0.996≈**1.000**

However, a closer look at the F-ratios of BNN-stems shows that this 'anomaly' is related to the fact that (in contrast to the other ones) the [Bie] stem type is

8. There are two irrelevant exceptions [Bee] vs. [Bɛe] and [Beɛ] vs. [Bɛɛ]; see Rebrus and Törkenczy (2016) for details and explanation.

not internally harmonically consistent. Internal harmonic consistency means that given a division of a stem type into subtypes, the F-ratios of the subtypes are not significantly different. Given a division of BNN-stems into consonant-final [BNNC] and vowel final [BNN#][9] subtypes, the internal harmonic consistency of a BNN-stem type can be defined as in (8):

(8) Internal harmonic consistency: [Bxy#] ≈ [BxyC]
(where X ≈ Y means that the F-ratio of X is not significantly lower/higher than that of Y)

The F-ratios in (9) show that the stem type [Bie] is indeed inconsistent (non-homogeneous) as opposed to the other BNN types, three of which are shown for contrast. Note also that the non-homogeneity of [Bie] results in the fact that the F-ratio of its vowel-final subtype [Bie#] is rather low; i.e. stems that belong to this subtype have a preference for the *back* alternants of harmonic suffixes.

(9) The internal harmonic consistency of the [Bie] type
i. homogeneous
[Bii#] ≈ [BiiC] 0.63 vs. 0.51
[Bei#] ≈ [BeiC] 0.61 vs. 0.70
[Bɛi#] ≈ [BɛiC] 0.77 vs. 0.70
ii. non-homogeneous
[Bie#] ≉ [BieC] 0.33 vs. 0.75

Although the corpus study in Rebrus and Törkenczy (2016) has found these effects, the question arises whether native speakers indeed observe the Law of Frequency Matching in this case (shown for the Height Effect and the Count Effect *separately* by Hayes et al. 2006, 2009).[10] Furthermore, the empirical testing of these variability effects in corpora is problematic because of data sparseness: these classes represent stems whose harmonically suffixed forms can be extremely rare. These are the main motivations for psycholinguistic testing, i.e. for collecting information about the variants from native speakers directly. In this paper, we report on our findings based on the psycholinguistic experiment we conducted. We wanted to find

9. We use the string of symbols "N#]" to indicate that the stem ends in a neutral vowel.

10. There are some important differences: Hayes et al. (2006, 2009) counted forms with the dative suffix -nak/-nɛk only whereas Rebrus & Törkenczy (2016) considered singly suffixed forms containing any harmonically alternating monosyllabic suffix. Hayes and Cziráky Londe (2006) did examine the application of the Height Effect in BNN-stems but (over)simplified the effect. They assumed that the ordering of the N_1N_2 sequences solely depends on the last neutral vowel (N_2) in the stem: the more transparent N_2 is according to the Height Effect, the more transparent the sequence is.

answers to the following questions: In which cases are (i) the Height Effect, (ii) the Count Effect, (iii) the Cumulativity Effect and (iv) the Locality Effect satisfied or violated? Furthermore: (v) Are the stem classes homogeneous in their harmonic behaviour: do consonant-final and vowel-final stems behave in the same way?

2. Experiment

2.1 Participants

21 adults participated in the experiment (14 women (mean age: 33 years, range 19–66 years; 7 men (mean age: 44 years, range 26–67 years). All participants were native speakers of Hungarian, 19 currently living in Budapest (10 born & raised there).

2.2 Stimuli

We set up 9 classes of real monomorphemic stems representing the relevant groups. For BN-stems there are 2 bisyllabic stem classes [Be] and [Bɛ]. We did not test class [Bi] because it shows no variability: all such stems take back suffixes. We included two trisyllabic stem classes in the experiment: [BBe] and [NBe]. For BNN-stems, each neutral vowel quality in each position is represented, except for [Bee], which is practically empty; and [BNɛ] stems, which do not show variation in the corpus, all such stems always take front suffixes: [Bii], [Bei], [Bie], [Bɛi], [Bɛe]. The number of stems in each class roughly corresponds to the real size of the class (all the stems in the class)[11] and in each one, we have a balanced sample, with both consonant-final and (different) vowel-final stems in each class where relevant. Consider the table of comparisons in (10) below, where the columns and the rows are stem subtypes and a cell at the intersection of a row and a column is a comparison of two subtypes, i.e. a potential ordering between them. The cells show the orderings imposed by the effects discussed in Section 1 (cf. (1), (4), (6), and (8)). The notation, where each symbol represents an ordering between the subtypes compared by these effects, is as follows:

- Height Effect: $<_{HE}$
- Count Effect: $<_{CE}$

11. We wanted to include as many stems as possible, hence the difference in stem tokens across types. Testing fewer stems of the bisyllabic types would not have provided us a detailed picture of the variation observed in and across different types.

- the transitive corollary of the Height Effect and the Count Effect: (<)
- Cumulativity: ≤
- the transitive corollary of Cumulativity: (≤)
- Locality: $\leq_{LOC}$
- internal harmonic consistency: ≈

The shaded cells represent comparisons where these effects do not impose an ordering on the subtypes of stems.

(10) A table of comparisons between stem subtypes

	Be	Bɛ	Bii	Bei	Bie	Bɛi	Bɛe
(Bi = 0)	$<_{HE}$	$<_{HE}$	$<_{CE}$	$<_{CE}$	$<_{CE}$	$<_{CE}$	(<)
Be	≈	$<_{HE}$		$<_{CE}$	$<_{CE}$	(<)	$<_{CE}$
Bɛ		≈				$<_{CE}$	$<_{CE}$
Bii			≈	≤	≤	≤	(≤)
Bei				≈	$\leq_{LOC}$	≤	(≤)
Bie					≈		≤
Bɛi						≈	≤
Bɛe							≈

Altogether, 104 stems were tested in the 9 classes. Each stem appeared with 4 different consonant-initial suffixes (dative **-nak/-nɛk**, instrumental **-val/-vɛl**, elative **-boːl/-bøːl** and allative **-hoz/-hɛz**), yielding 416 experimental sentences altogether. See (11) for the distribution of stems across stem classes in the experiment.

(11) The distribution of the number of stems across stem classes in the experiment[12]

Stem type	Number of stems	Example
Be	37	taːɲeːr, sateːn
Bɛ	37	havɛr, fotɛl
Bii	6	aspirin, kolibri
Bie	6	matineː, klarineːt
Bei	3	proteːziʃ, poeːziʃ

12. Glosses: **kolibri** 'hummingbird', **matineː** 'matinee', **poeːziʃ** 'poetry', **ʃpagɛtti** 'spaghetti', **sutɛreːn** 'basement', **indoneːz** 'Indonesian', **finaːleː** 'finale', **majoneːz** 'mayonnaise', **karanteːn** 'quarantine'.

Bɛi	6	bakɛlit, ʃpagɛtti
Bɛe	3	suvɛreːn, sutɛreːn
NBe	3	indoneːz, finaːleː
BBe	3	majoneːz, karanteːn

The experimental sentences were complemented by 208 filler sentences. Filler sentences were 4–6 words long, they did not contain mixed stems or any of the target suffixes. Experimental sentences were sorted into two batches, yielding two versions of the experiment (A and B). Each participant heard every target stem with two different suffixes, an a/ɛ (dative or instrumental) suffix and a non-a/ɛ suffix (allative or elative) out of the four (208 target sentences) and 208 filler sentences. Filler sentences were the same in the two batches.

2.3 Method and procedure

We collected data from adult participants in an elicited production task disguised in the form of a sentence repetition task. Each target word+suffix combination was presented acoustically, as part of a digitally prerecorded sentence. The target inflections in each sentence (and sometimes another syllable in the sentence) were masked by a carefully inserted cough that prevented the participant from hearing the inflection, but not the stem or the remaining portions of the sentence, as illustrated below (where strikethrough represents the cough):

Valamiért sosem voltam híve az aszpirin~~nak/nek~~.
'For some reason I have never been devoted to aspirin-~~DAT~~.'
A trópusi kolibri~~nak/nek~~ kék a tollazata.
'The tropical hummingbird-~~DAT~~ (has) blue plumage.'

The audible parts of the sentence make it clear which inflection is missing, but provide no cues to the frontness of the actual suffix alternant. After hearing the sentence, participants were asked to repeat the sentence. This design was modeled after Warren's (1970) phoneme restoration procedure. Restoration with the same procedure works at the morpheme level e.g. for affixes in Hungarian (Dankovics & Pléh 2001), and has been successfully used as an elicited production method for suffixes with children (Lukács et al. 2009). In this design, participants are usually unaware that the inflections are missing, which allows us to examine the differences in variability in production without relying on metalinguistic awareness and conscious decision about the front/back variants. Crucially, it also allows us to collect data for stem+suffix combinations that are rarely or never attested in corpora, thus providing new sets of data for systematically testing the above hypotheses.

2.4 Results

The dependent variable in all cases was the frontness of the inflection the participants produced. To test the Count Effect, the Height Effect and the Cumulativity Effect, the effect of stem type and stem subtype was tested on mean percentages of front answers. Since Type and Subtype were within-subject factors, the results were analyzed by a Repeated measures ANOVA (analysis of variance), and were further tested by post-hoc pairwise comparisons.

2.4.1 *Results by generalized type – Count Effect*

First, we tested the Count Effect by comparing mean percentages of front answers in a repeated measures GLM with Type as a 4-level factor (BN, BBN, NBN, BNN). A significant main effect of Type was observed ($F(3, 60) = 44{,}66$; $p<0.001$). Post hoc pairwise comparisons showed that BN-stems were significantly different from all others ($p<0.001$), and, as expected by the Count Effect, BNN-stems were associated with a significantly higher ratio of front answers than stem types with a single neutral vowel between the back vowel of the stem and the suffix (i.e. BN, BBN, NBN stems).[13] BBN and BN stems differed significantly ($p<0.05$). No other pairwise differences were significant. The results are shown in Figure 1.

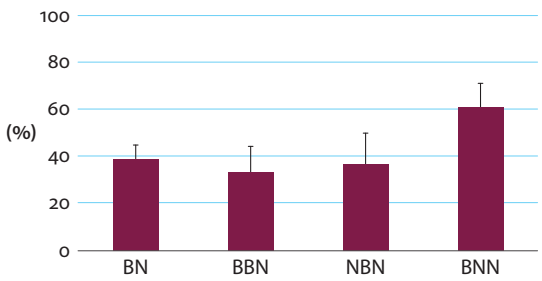

Figure 1. Mean percentage of front answers across participants by stem type. Error bars indicate standard deviations (SDs)

2.4.2 *Height Effect*

To test the Height Effect, we compared mean percentages of front answers in a repeated measures GLM with Subtype as a 4-level factor ([Be], [BBe], [NBe], [Bɛ]),

13. Note that the difference would have been even greater if we had included [Bi] and [BNɛ] stems.

which revealed a significant main effect of Subtype (F(3, 60) = 184.46; p<0.001). Post hoc pairwise comparisons showed significant differences between all pairs of subtypes (p<0.001) except for [NBe] vs. [BBe], which did not differ statistically. The standard deviation is also higher in these subtypes,[14] and, compared to BN types, the number of roots tested is lower (3 vs. 37).

The results illustrating the Height Effect by subtype are shown in Figure 2.

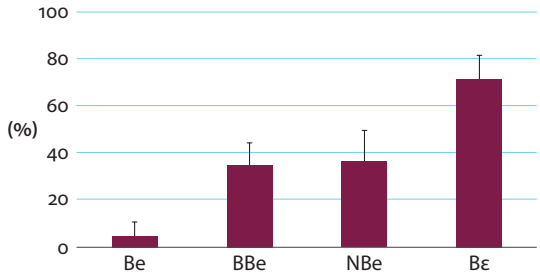

Figure 2. Mean percentage of front answers across participants by stem subtype. Error bars indicate SDs

2.4.3 *Cumulativity 1*

The Cumulativity hypothesis was first tested by comparing the results for [Bii], [Bei], [Bie] and [Bɛi] in a repeated measures GLM with Subtype as a 4-level factor, which revealed a significant main effect of Subtype (F(3, 60) = 26.485; p<0.001). Post hoc pairwise comparisons showed significant differences between all pairs of subtypes ([Bie] < [Bɛi], [Bei], [Bii] at p<0.001; [Bii] < [Bɛi] and [Bei] < [Bɛi] at p<0.01) except for [Bei] vs. [Bii], which did not differ statistically. The results by subtype are shown in Figure 3.

14. The F-ratios of the stems belonging to the [BBN] type were the following: **paraːde** 'parade' 0.7%, **karanteːn** 'quarantine' 8%, **majoneːz** 'mayonnaise' 93%; and that of the stems in the [NBN] type were: **finaːleː** 'finale' 0%, **diadeːm** 'diadem' 31%, **indoneːz** 'Indonesian' 73%. Presenting the averaged F-ratios of these types hides the extent of variation within the types, although this might provide an explanation for the lack of significant difference between the two types. It seems that the two types are also not harmonically consistent (similarly to the type [Bie], see (9) above). That is, consonant-final and vowel-final stems do not behave in a uniform way.

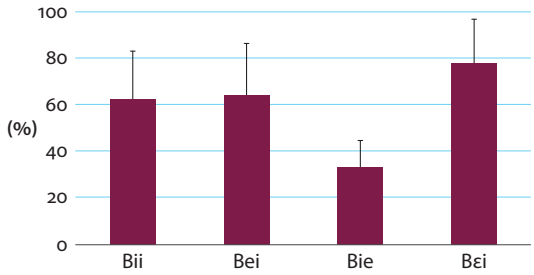

Figure 3. Mean percentage of front answers across participants by stem subtype. Error bars indicate SDs

Our hypotheses expected [Bei] ≤ [Bie] to hold, but since [Bie] stems have significantly lower F-ratios than all other BNN-stems, they violate both the Cumulativity and the Locality Effects. This is consistent with the corpus results shown in (7) above, but the difference is more pronounced in our results.

2.4.4 *Cumulativity 2*

The Cumulativity hypothesis was also tested by comparing the results for [Be], [Bei], [Bie] and [Bɛe] stems in a similar repeated measures GLM with Subtype as a 4-level factor, which revealed a significant main effect of Subtype ($F(3, 60) = 128.471$; $p<0.001$). Post hoc pairwise comparisons showed significant differences between all pairs of subtypes ([Bɛe] < [Bei] at $p<0.05$; all others at $p<0.001$). The results by subtype are shown in Figure 4.

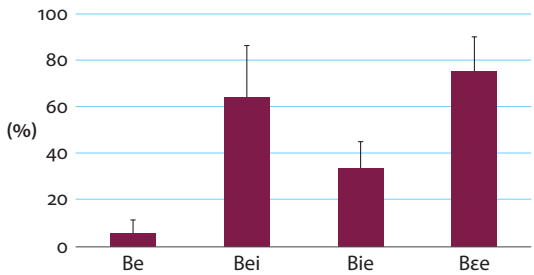

Figure 4. Mean percentage of front answers across participants by stem subtype. Error bars indicate SDs

Figures 3 and 4 show that the order and the quality of the neutral vowels in a BNN-stem can cause significant differences in their F-ratios, and that the harmonic behaviour of a stem is affected by both N_1 and N_2 (i.e. the last vowel is not the only one responsible for harmonic behaviour, *contra* Hayes & Cziráky Londe

2006, Bowman 2013). According to the experiment, native speakers follow the general trends found in Rebrus and Törkenczy's (2016) corpus results, but the differences can be more marked.

2.4.5 *Comparisons of C-final vs V-final roots*

A corpus study (Rebrus and Törkenczy 2016) has shown that consonant-final BNN-stems ([BNNC]) and vowel-final BNN-stems ([BNN#]) are not necessarily homogeneous in their harmonic behaviour in all subtypes. In (12), we show the F-ratios of [BNNC] and [BNN#] stems, comparing experimental data with word forms in the webcorpus that contain the same stem+suffix combinations that we tested in the experiment. According to the Cumulativity Effect, the F-ratios should be the following: [Bii] ≤ [Bie] ≤ [Bɛi]. As can be seen in (12), this only holds for [BieC] ≤ [BɛiC], [Bii#] ≤ [Bɛi#] and [Bie#] ≤ [Bɛi#] in the webcorpus and not for the other three ([BiiC] ≰ [BieC], [BiiC] ≰ [BɛiC], [Bii#] ≰ [Bie#]) of the theoretically possible six pairs.[15] The results of the experiment are a close match for the corpus frequencies. The two pairs violating Cumulativity are [BiiC] ≰ [BieC] and [Bii#] ≰ [Bie#]; the others all conform to the ordering by Cumulativity. Note that with one exception ([BiiC] ≰ [BɛiC] in the webcorpus), the pairs violating ordering by Cumulativity all involve [Bii] vs. [Bie] comparisons.

The difference between the F-ratios of [BiiC] stems (webcorpus: 0.94, experiment: 0.50) is due to the fact that the webcorpus contains only one stem out of the three which were tested (**aspirin** 'id.'). However, this stem in the experiment has an F-ratio of 0.81, which fits in with the general patterns observed in the webcorpus.

(12) Internal consistency: front-ratios of the C# and V# roots used in the experiment in the webcorpus and the experiment

	Webcorpus			**Experiment**			
	C#	**V#**	**diff.**	**C#**	**V#**	**diff.**	***p***
[Bii]	0.94	0.81	**0.13**	0.50	0.75	**–0.25**	0.002
[Bie]	0.43	0.20	**0.23**	0.46	0.21	**0.25**	0.000
[Bɛi]	0.88	0.91	**–0.03**	0.73	0.86	**–0.13**	0.058

As can be seen in (12), Cumulativity holds for consonant-final (C#) stems in the experiment; the difference between [BiiC] and [BieC] (0.50 vs. 0.46) is not

15. Compare (7) where (i) all available BNN-stems were counted with all available harmonic suffixes, and (ii) consonant-final and vowel-final stems were not distinguished.

significant. However, with vowel-final (V#) stems, Cumulativity is violated by the [Bie] subtype. Regarding internal harmonic consistency, it has already been noted that [Bie] stems are not consistent (cf. (9ii)). The experimental results also show that [Bii#] and [BiiC] stems differ significantly, but in the other direction: [Bii#] stems are more likely to take a front suffix than [BiiC] ones, while [Bie#] stems are more likely to take back suffixes than [BieC] stems.

If we look at all of the subtypes and their harmonic behaviour according to the webcorpus results (taking the same stem+suffix combinations that we tested), and compare them to the experimental data in (13) below, we can see that there are three marked differences ([Bii], [Bei], [Bɛe]). In the case of [Bii] stems, the difference may be due to the fact that out of the six stems we tested, only three were found in the webcorpus, yielding an F-ratio of 0.86, while the experimental result is 0.63. [Bei] stems show a 31% difference (webcorpus: 0.95, experiment: 0.64), and [Bɛe] stems show a 24% difference in their F-ratios (webcorpus: 0.99, experiment: 0.75). The differences here are probably due to the fact that the words tested in these subtypes are very infrequent. Nevertheless, it is clear that the general trends in harmonic behaviour in the corpus are a close match for our results, i.e. native speakers do observe the Law of Frequency Matching.

(13) Harmonic behaviour of BNN-stems in the webcorpus and in the experiment

Types			Webcorpus			Experiment		
Bii	**Bie**	Biɛ	0.86	0.30	((0.99))	0.63	0.33	–
Bei	Bee	Beɛ	0.95	–	((1.00))	0.64	–	–
Bɛi	**Bɛe**	Bɛɛ	0.89	0.99	((1.00))	0.79	0.75	–

3. Conclusion

Previous studies of variability in Hungarian vowel harmony were based on corpus data, which has its known limitations (e.g. data sparseness on rare combinations), and/or wug testing using a single alternating suffix, which may oversimplify the dimensions and the range of variation. Our experiment provided direct data from native speakers, which made it possible to examine a larger set of stem+suffix combinations, and thus shed light on several lesser examined areas of HVH. The data provided new insights about the combined workings of the Height and the Count Effect, the behaviour of different harmonically alternating suffixes, and the harmonic behaviour of consonant-final and vowel-final stems. The experiment has also allowed us to confirm previous

hypotheses that were formed based on the basis of corpus data and to identify areas/directions where more research is needed.

We have examined the harmonic behaviour of existing words combined with four different suffixes, based on an experiment with native speakers, and we provided a statistical analysis. Our results confirm the Height Effect and the Count Effect previously observed in corpus studies (Hayes & Cziráky Londe 2006 (based on Google searches), Rebrus & Törkenczy 2016 (based on the Szószablya webcorpus)) and wug tests (Hayes & Cziráky Londe 2006): [Be] stems are more transparent (i.e. less front) than [Bɛ] stems, and the frontness of BNN-stems is significantly higher than that of BN-stems. Furthermore, our results also confirm the Cumulativity and Locality Effects in native speaker behaviour: the quality of N_1, the quality of N_2, and the order of the two neutral vowels, are *all* important in determining a [BN_1N_2] stem's harmonic behaviour, contrary to the simplified view in Bowman 2013 and Hayes & Cziráky Londe 2006, who assume that only the last neutral vowel N_2 is relevant. Internal consistency effects were also observed in the experiment: the stem types [Bie] and [Bii] are not internally homogeneous. The stems in these types show different harmonic behaviour depending on whether their final segment is a consonant or a vowel, albeit with an opposite harmonic bias (i.e. [Bie#]: back bias; [Bii#]: front bias). The experiment has also confirmed the violation of Cumulativity found in the corpus involving the stem type [Bie#].

Interestingly, harmonic suffixes do not show a uniform behaviour either – contrary to virtually all analyses of HVH, which typically assume that harmonically alternating suffixes are uniform in their harmonic behaviour (but see Rebrus & Törkenczy 2013). Further research is needed to map out the exact way in which different suffixes behave when attached to BN and BNN-stems, but we can see the general patterns in Figure 5. We can observe differences in F-ratio between the harmonic suffixes examined in every stem type.

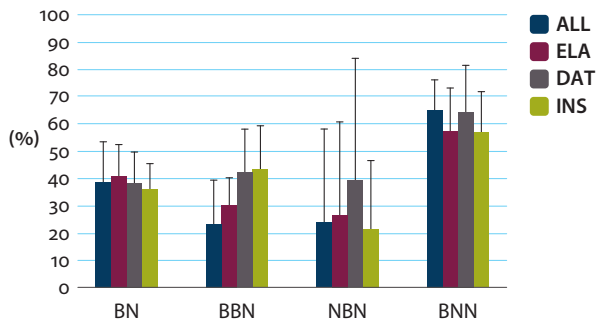

Figure 5. Mean percentage of front answers across participants by stem type+suffix. Error bars indicate SDs

Even with a relatively small sample of roots and a completely different methodology of data collection, our results correspond to the tendencies observed in the webcorpus. However, we have to emphasize that the conditioning of harmonic variation may be even more fine-grained. For instance, it is itself a generalization that we treat different stems of the same harmonic pattern (i.e. stems that have the same neutral vowels following a back vowel) as belonging to the same subtype (e.g. **bakɛlit** 'bakelite' and **ʃpagɛtti** 'spaghetti' belonging to the subtype [Bɛi]). However, the frequency of the stems themselves and their harmonic behaviour may differ from the generalized subtype. This means that even though the six stems in the [Bii] subtype are labelled by a certain number that represents their F-ratio (in this case, 0.63), the individual stems themselves may have different harmonic behaviour (F-ratios); e.g. **aspirin** 'id.' and **kolibri** 'hummingbird' do not behave exactly the same way.

It is a further generalization that we treat the experimental and webcorpus data as the same. We have pointed out above that our results conform to the Law of Frequency Matching. Nevertheless, we have found some interesting differences between the frequencies gained from the corpus and the native speaker reactions in our experiment. The difference in F-ratios between the webcorpus and the experiment may indicate that the comparison of written and experimental data is not straightforward, as token and type frequencies taken from the corpus are merely a simulation of gaining data from participants.

Acknowledgements

We would like to thank Katalin Mády and Bertalan Polner for their help with the experiment. This work has been supported by the National Scientific Grant NKFI-119863 'Experimental and theoretical investigations of vowel harmony patterns'.

References

Bowman, Samuel R. 2013. Two arguments for a positive vowel harmony imperative. Ms. Stanford [ROA 1181].

Dankovics, Natália & Csaba Pléh. 2001. Hangrestaurációs jelenségek és alaktani feldolgozás a magyarban: Azt halljuk-e, amit várunk? [Sound restauration and morphological processing in Hungarian: Do we hear what we expect to hear?] In Csaba Pléh & Ágnes Lukács (eds.), *A magyar morfológia pszicholingvisztikája* [*The psycholinguistics of Hungarian morphology*], 55–83. Budapest: Osiris.

Halácsy, Péter, András Kornai, László Németh, András Rung, István Szakadát & Viktor Trón. 2004. Creating open language resources for Hungarian. In *Proceedings of Language Resources*

and Evaluation Conference (LREC04). LREC, 203–210. http://szotar.mokk.bme.hu/szoszablya/searchq.php

Hayes, Bruce & Zsuzsa Cziráky Londe. 2006. Stochastic phonological knowledge: The case of Hungarian vowel harmony. *Phonology* 23. 59–104. https://doi.org/10.1017/S0952675706000765

Hayes, Bruce, Kie Zuraw, Péter Siptár & Zsuzsa Londe. 2009. Natural and unnatural constraints in Hungarian vowel harmony. *Language* 85. 822–863. https://doi.org/10.1353/lan.0.0169

Lukács, Ágnes, Laurence B. Leonard, Bence Kas & Csaba Pléh. 2009. The use of tense and agreement by Hungarian-speaking children with language impairment. *Journal of Speech, Language and Hearing Research 52*/1. 1–22.

Rebrus, Péter & Péter Szigetvári. 2016. Diminutives: Exceptions to harmonic uniformity. *Catalan Journal of Linguistics* 15. 101–119. https://doi.org/10.5565/rev/catjl.186

Rebrus, Péter & Miklós Törkenczy. 2013. Magánhangzó-diszharmónia [Vowel disharmony]. *Talk delivered at the Research Institute for Linguistics*, Hungarian Academy of Sciences, Budapest, 22 October.

Rebrus, Péter & Miklós Törkenczy. 2015. Monotonicity and the typology of front/back harmony. *Theoretical Linguistics* 41(1–2). 1–61. https://doi.org/10.3765/amp.v2i0.3769

Rebrus, Péter & Miklós Törkenczy. 2016. A non-cumulative pattern in vowel harmony: A frequency-based account. In Gunnar Ólafur Hansson, Ashley Farris-Trimble, Kevin McMullin & Douglas Pulleyblank (eds.), *Proceedings of the 2015 Annual Meeting on Phonology*. Washington, DC: Linguistic Society of America. https://doi.org/10.3765/amp.v3i0.3692

Rebrus, Péter & Miklós Törkenczy. 2017. Gradient harmonicity in compounds. In Karen Jesney, Charlie O'Hara, Caitlin Smith & Rachel Walker (eds.), *Proceedings of the 2016 Annual Meeting on Phonology*. Washington, DC.: Linguistic Society of America. https://doi.org/10.3765/amp.v4i0.4007

Ringen, Catherine & Miklós Kontra. 1989. Hungarian neutral vowels. *Lingua* 78. 181–191. https://doi.org/10.1016/0024-3841(89)90052-1

Siptár, Péter & Miklós Törkenczy. 2000. *The phonology of Hungarian*. Oxford: Oxford.

Törkenczy, Miklós. 2016. Hungarian vowel harmony. In Mark Aronoff (ed.), *Oxford bibliographies in linguistics*. New York: Oxford University Press. https://doi.org/10.1093/obo/9780199772810-0134

Vago, Robert. 1980. *The sound pattern of Hungarian*. Washington: Georgetown University Press.

Warren, Richard M. 1970. Perceptual restorations of missing speech sounds. *Science*, 167. 392–393. https://doi.org/10.1126/science.167.3917.392