# Dependencies between morphophonological patterns

Guy Tabachnick (New York University), guyt@nyu.edu

April 30, 2021

## 1 Introduction

One problem that learners face: how to arbitrarily associate *morphophonological patterns* with specific *lexical items*. Examples:

- English past tense (Albright and Hayes, 2003): *reach–reached* but *teach–taught*
- Russian diminutives (Gouskova et al., 2015): [áng$^j$il]–[áng$^j$il-ók] 'angel', but [mónstr]—[mónstr-$^j$ik] 'monster'

When learners need to generate a previously unseen form (past tense, diminutive) of a novel lexical item, how do they do it?

- Typical approaches focus on *phonological* generalizations that speakers can pick up on
  - *-ok* never attaches to stems ending in consonant clusters
- Ackerman et al. (2009); Ackerman and Malouf (2013): inflectional systems are organized in a way to allow speakers to efficiently predict one form given others (solving the Paradigm Cell Filling Problem)

Some models of lexically specific patterns, like the *sublexicon* model (Gouskova et al., 2015; Becker and Gouskova, 2016), allow speakers to pick up on phonological biases.

Today: a case study from Hungarian showing that just looking at the phonology isn't enough:

- nouns take one of two possessive suffixes, *-ɒ* or *-jɒ*
- nouns in a class known as "lowering stems" overwhelmingly take *-ɒ* and only rarely *-jɒ*

Crucially: these correlations are *not* due to shared phonological characteristics between the two classes! Models that capture phonological tendencies in a pattern are not capturing the whole story.

I propose an extension to the sublexicon model to account for these morphological depencies.

# 2 Case study: Hungarian possessives

## 2.1 Possessive allomorphy

In Hungarian, the possessive suffix has two allomorphs (ignoring vowel harmony), *-ɒ/ɛ* and *-jɒ/jɛ*. Nouns select one or the other, with some allowing both (sometimes with a distinction in meaning):

| *word* | *gloss* | *possessive* |
|---|---|---|
| paːr | 'pair' | paːr-**j**ɒ |
| sor | 'line' | ʃor-ɒ |
| ɒblɒk | 'window' | ɒblɒk-(**j**)ɒ |

Table 1: Some Hungarian possessive forms (see Rácz and Rebrus, 2012)

Rácz and Rebrus (2012) summarize: *-jɒ* has spread to become the productive form today (used for loanwords, etc.)

Some basic generalizations about possessive allomorph selection (they describe other non-categorical patterns as well):

| *stem shape* | *possessive allomorphy* | *example* | |
|---|---|---|---|
| V-final | always *-jɒ* | kɒpu-**j**ɒ | 'gate' |
| palatal-final | always *-ɒ* | laːɲ-ɒ | 'girl' |
| sibilant-final | always *-ɒ* | gaːz-ɒ | 'gas' |

Table 2: Rácz and Rebrus (2012): Generalizations about possessive allomorph selection

## 2.2 Lowering stems

They also note: irregular noun classes usually take *-ɒ*. One example: *lowering stems*, which have a low vowel in forms like the plural where most nouns have a mid vowel (see also Rebrus, 2013; Rebrus et al., 2017):[1]

| *stem class* | *example* | *gloss* | *plural* | *possessive* |
|---|---|---|---|---|
| regular | paːr | 'pair' | paːr-o̲k | paːr-**j**ɒ |
| lowering | ɟaːr | 'factory' | ɟaːr-ɒ̲k | ɟaːr-ɒ |

Table 3: Hungarian lowering stems

[1] By default, nouns are non-lowering stems, while all but a handful of adjectives are lowering stems, and some forms behave differently depending on the syntactic environment (Rebrus and Szigetvári, 2018). Setting aside these issues, I limit my inquiry to words that are listed as nouns.

How strong is this tendency? Quite![2]

| *stem class* | *possessive* -ɒ/ɛ | *possessive* -jɒ/jɛ | *possessive* -ɒ/ɛ | *possessive* -jɒ/jɛ |
|---|---|---|---|---|
| regular | 642 | 935 | kaːr-o̲k, kaːr-ɒ 'damage' | paːr-o̲k, paːr-**j**ɒ 'pair' |
| lowering | 104 | 15 | ɟaːr-ɒ̲k, ɟaːr-ɒ 'factory' | ɲaːr-ɒ̲k, ɲaːr-**j**ɒ 'poplar' |

Table 4: Distribution of Hungarian lowering stems and possessive allomorphy

One possibility: lowering stems just look like typical *-ɒ* words (maybe a lot end in sibilants and palatals, for example).

This may be a small factor, for example:

- $470/1575 \approx 30\%$ of regular stems end in palatals or sibilants (which only take *-ɒ*)
- $48/119 \approx 40\%$ of lowering stems do

However, this alone doesn't explain the effect—the morphological effect is independent of the phonological ones!

## 2.3 Phonological and morphological factors influencing possessive allomorphy

I fit two models for predicting a noun's possessive allomorph—one with only phonological factors, one with stem class as a predictor as well.[3]

Not only does adding stem class *improve* the model, but the phonological effect sizes stay steady—the morphological dependency is independent!

Example: the effect of final consonant place. (Here *higher coefficient* → *more likely* to take *-jɒ* relative to the baseline, alveolar. All effects are significant, $p < .0001$.)

| *predictor* | *effect size in model without stem class* | *effect size in model with stem class* | *predicted rates* |
|---|---|---|---|
| final C: alveolar | (baseline) | | |
| final C: labial | –2.05 | –2.25 | tsiːm-jɛ < siːn-jɛ |
| final C: palatal | –9.01 | –9.40 | ryɟ-jɛ < tʃyd-jɛ |
| final C: velar | –3.35 | –3.63 | gøːg-jɛ < tʃøːd-jɛ |
| stem class: regular | (baseline) | | |
| stem class: lowering | — | –3.77 | ɲaːr-jɒ < paːr-jɒ |

Table 5: The effect of final consonant place on Hungarian possessive allomorphy, with and without stem class as a predictor

[2]My Hungarian source is Papp (1969), a morphological dictionary. I consider consonant-final, monomorphemic nouns that are not marked as variable in the possessive. Not included in Table 4 are 23 nouns listed as having variable stem class.

[3]The models were fit using the `lrm` function in R's `rms` package (R Core Team, 2020; Harrell Jr., 2020). See Appendix A for the full models.

Not only is being a lowering stem a fairly strong predictor of selecting *-ɒ*, but it gives information that the phonology of a noun alone does not!

To put it another way: seeing a noun's phonological form is helpful to the learner in figuring out its possessive—but seeing its plural to learn that it's a lowering stem is even more helpful.

# 3 A sublexicon model with morphological dependencies

## 3.1 The basic sublexicon model

The *phonological* effects in Table 5 can be handled with a sublexicon model (Gouskova et al., 2015; Becker and Gouskova, 2016).

Steps for the learner (in the basic sublexicon model):

1. The learner associates a morpheme with a particular morphosyntactic context and creates a rule of exponence linking form to meaning
   - basic rule of exponence for the Hungarian possessive: POSS ↔ jɒ
2. The learner encounters the same morpheme with a different exponent, and splits the rule of exponence into two contextually specific rules marked by diacritics[4]
   - spun-off rules of exponence for the Hungarian possessive:
     - POSS ↔ jɒ / [+j] ___
     - POSS ↔ ɒ / [−j] ___
3. The learner places lexical items into *sublexicons* associated with these diacritics (i.e. their behavior in the given context)
   - [+j] sublexicon = {paːr, kɒpu, ...}
   - [−j] sublexicon = {ɟaːr, gaːz, ...}
4. The learner generates a phonotactic grammar (Hayes and Wilson, 2008) for each sublexicon, which captures generalizations over the phonological shape of its members through phonotactic constraints
   - highly weighted constraints in the [+j] sublexicon: *[palatal]#, *[sibilant]#
   - highly weighted constraints in the [−j] sublexicon: *V#
5. When the learner encounters a new lexeme and has to generate the form in question, they evaluate the stem against each sublexicon's grammar to see which is a better fit and place it accordingly
   - nonce word [oːd] rates better on the [+j] sublexicon
   - the learner marks it with [+j], places it in the [+j] sublexicon, and produces [oːd-jɒ][5]

[4]For the sake of illustration, I mark both new rules of exponence with a diacritic. An alternative is that one rule gets the diacritic and the other is left as a default applying in other cases.

[5]By a regular rule of palatal assimilation, the actual surface form would be [oːɟɒ].

### 3.2 Adding morphological dependencies

Meanwhile, the learner has also noticed that there's variation in the vowel in plural and other forms and has placed words into additional sublexicons based on a [±lower] diacritic:[6]

- PL ↔ ɒk / [+lower] ___ [+lower] = {[ɟaːr], [ɲaːr], ...}
- PL ↔ ok / [−lower] ___ [−lower] = {[kaːr], [paːr], ...}

Recall: being a lowering stem is a very good predictor for taking *-ɒ* as a possessive allomorph ([ɟaːr-ɒ̱k] → [ɟaːr-ɒ]).

In sublexicon terms: if a noun is in the [+lower] sublexicon, it's likely to also be in the [−j] sublexicon.

The basic sublexicon model, however, can't account for this! Let's add a step:

4′. When creating a new sublexicon, add *SUBLEX constraints penalizing membership in already existing sublexicons, and add *SUBLEX constraints for the new sublexicon into all existing sublexical grammars.
  - the new [±j] sublexicons are seeded with *SUBLEX-[+lower] and *SUBLEX-[−lower]
  - by definition: *SUBLEX-[+lower] penalizes lexical items in the [+lower] sublexicon
  - *SUBLEX-[+lower] will be weighted heavily in the [+j] sublexicon grammar: candidates for taking possessive *-jɒ* are penalized for also being a lowering stem with plural *-ɒk*

The addition of *SUBLEX constraints thus extends the sublexicon model to pick up not just on phonological patterns, but also on dependencies between lexically specific patterns!

### 3.3 The organization of morphological systems

Sublexicons can help us conceptualize learning *declension classes*, which are essentially diacritics that appear in multiple contextual rules of exponence:

- in Russian, class II nouns like [komnata] 'room' have [-a] in the nominative, [-u] in the accusative, etc. (Corbett and Fraser, 1993)
- potential learning process:
  - learner starts off with [NOM:a] and [ACC:u], etc. diacritics/sublexicons
  - the two classes contain almost identical nouns and grammars
  - accordingly, the learner merges them into a single [class:II] sublexicon

[6] These rules of exponence are simplified for the purposes of illustration. A given noun's choice of vowel, known as a "linking vowel", is uniform across several morphological categories, of which plural is only one, so there is good reason to think that the selection of vowel is handled in a different rule of exponence from insertion of the plural allomorph.

Relatedly: researchers like Ackerman et al. (2009); Ackerman and Malouf (2013) have argued for treating the inflectional paradigm as a basic unit of linguistic structure based on the ways in which inflected forms are informative about one another.

- some patterns that seem almost unlearnably complex from the perspective of “X contains an exponent of morpheme M” look much more efficient when viewed as “X is a member of paradigm P”

Sublexicons with *SUBLEX constraints allow the formal grammatical model to capture some of this informativity, even under theories of exponence like Distributed Morphology that deny the formal existence of paradigms.

# 4 Summary

Hungarian possessives show:

- many patterns of lexically specified behavior show phonological biases that speakers can pick up on and exploit in generating new forms
- sometimes two such patterns are informative about one another—no reason speakers can’t exploit these as well
- we can model these dependencies using the sublexicon approach with additional *SUBLEX constraints, which give us a way to pick up on biases in the organization of a language’s morphological system

# A Appendix: Full models

This appendix contains the full models calculated for predicting possessive allomorphs among monomorphemic Hungarian nouns without variable possessive forms, with data taken from Papp (1969). The models were calculated using the `lrm` function in R's `rms` package (R Core Team, 2020; Harrell Jr., 2020). The first model (Table 6) contains only phonological factors, while the second (Table 7) also includes the morphological factor of lowering stem class.

I assembled the models using stepwise comparison, where each additional factor significantly improved the model ($p < .0001$) according to the same package's `lrtest` function. For both cases, an ANOVA confirmed that each factor in the final models was significant ($p < .0001$). Furthermore, each addition to the models improved their Akaike Information Criterion (AIC), a metric that rewards model likelihood but penalizes model complexity—that is, the improvement in model fit brought by adding each factor outweighed the cost of having a more complex model.

| | *β coef* | *SE* | *Wald z* | *p* | *predicted rates* |
|---|---|---|---|---|---|
| **Intercept** | **3.03** | **.32** | **9.48** | **<.0001** | **bot-jɒ > bot-ɒ** |
| C Manner (default: plosive) | | | | | |
| **fricative** | **–1.40** | **.39** | **–3.54** | **.0004** | **raːf-jɒ < laːp-jɒ** |
| **sibilant** | **–11.57** | **1.09** | **–10.65** | **<.0001** | **hɒs-jɒ < lɒt-jɒ** |
| **nasal** | **–1.99** | **.28** | **–7.21** | **<.0001** | **oːn-jɒ < toːt-jɒ** |
| **approximant** | **–4.16** | **.31** | **–13.45** | **<.0001** | **tʃɛr-jɛ < tsɛt-jɛ** |
| C Place (default: alveolar) | | | | | |
| **labial** | **–2.05** | **.26** | **–7.95** | **<.0001** | **tsiːm-jɛ < siːn-jɛ** |
| **palatal** | **–9.01** | **1.13** | **–8.01** | **<.0001** | **ryɟ-jɛ < tʃyd-jɛ** |
| **velar** | **–3.35** | **.30** | **–11.30** | **<.0001** | **gøːg-jɛ < tʃøːd-jɛ** |
| Harmony (default: back) | | | | | |
| **front** | **–2.06** | **.19** | **–10.91** | **<.0001** | **køb-jɛ < dob-jɑ** |
| **variable** | **2.51** | **1.10** | **2.27** | **.0230** | **ɒnkeːt-jɒ/jɛ > klɒrineːt-jɒ** |
| V Height (default: mid) | | | | | |
| **high** | **1.91** | **.23** | **8.30** | **<.0001** | **cuːk-jɒ > tʃoːk-jɒ** |
| low | 0.32 | .19 | 1.68 | .0927 | kɒr-jɒ > kor-jɒ |
| V Length (default: short) | | | | | |
| **long** | **1.38** | **.18** | **7.75** | **<.0001** | **boːr-jɒ > bor-jɒ** |
| Coda (default: singleton) | | | | | |
| **geminate** | **2.49** | **.40** | **6.19** | **<.0001** | **ʃikː-jɛ > sik-jɛ** |
| cluster | 0.03 | .21 | 0.12 | .9042 | domb-jɒ > dob-jɒ |
| Syllables (default: monosyllabic) | | | | | |
| **polysyllabic** | **1.20** | **.18** | **6.85** | **<.0001** | **ɛlɛm-jɛ > sɛm-jɛ** |

Table 6: Regression model with phonological predictors of possessive [j], with significant effects bolded

| | *β coef* | *SE* | *Wald z* | *p* | *predicted rates* |
|---|---|---|---|---|---|
| **Intercept** | **3.54** | **.34** | **10.51** | **<.0001** | **bot-jɒ > bot-ɒ** |
| Stem class (default: non-lowering) | | | | | |
| **lowering** | **–3.77** | **.45** | **–8.47** | **<.0001** | **ɲaːr$_{[+\text{low}]}$-jɒ < paːr$_{[-\text{low}]}$-jɒ** |
| undetermined | –0.20 | .27 | –0.72 | .4694 | (no direct comparison) |
| **variable** | **–2.77** | **.69** | **–4.02** | **<.0001** | **taːr$_{[\pm\text{low}]}$-jɒ < tsaːr$_{[-\text{low}]}$-jɒ** |
| C Manner (default: plosive) | | | | | |
| **fricative** | **–0.89** | **.45** | **–1.98** | **.0475** | **raːf-jɒ < laːp-jɒ** |
| **sibilant** | **–11.95** | **1.08** | **–11.08** | **<.0001** | **hɒs-jɒ < lɒt-jɒ** |
| **nasal** | **–2.11** | **.28** | **–7.43** | **<.0001** | **oːn-jɒ < toːt-jɒ** |
| **approximant** | **–4.13** | **.32** | **–13.05** | **<.0001** | **tʃɛr-jɛ < tsɛt-jɛ** |
| C Place (default: alveolar) | | | | | |
| **labial** | **–2.25** | **.27** | **–8.37** | **<.0001** | **tsiːm-jɛ < siːn-jɛ** |
| **palatal** | **–9.40** | **1.15** | **–8.18** | **<.0001** | **ryɟ-jɛ < tʃyd-jɛ** |
| **velar** | **–3.63** | **.31** | **–11.66** | **<.0001** | **gøːg-jɛ < tʃøːd-jɛ** |
| V Height (default: mid) | | | | | |
| **high** | **2.06** | **.24** | **8.53** | **<.0001** | **cuːk-jɒ > tʃoːk-jɒ** |
| **low** | **0.81** | **.21** | **3.86** | **.0001** | **kɒr-jɒ > kor-jɒ** |
| Harmony (default: back) | | | | | |
| **front** | **–2.04** | **.29** | **–7.15** | **<.0001** | **køb-jɛ < dob-jɑ** |
| **variable** | **2.48** | **1.17** | **2.13** | **.0335** | **ɒnkeːt-jɒ/jɛ > klɒrineːt-jɒ** |
| Coda (default: singleton) | | | | | |
| **geminate** | **2.43** | **.42** | **5.86** | **<.0001** | **ʃikː-jɛ > sik-jɛ** |
| cluster | –0.11 | .22 | –0.49 | .6277 | domb-jɒ < dob-jɒ |
| V Length (default: short) | | | | | |
| **long** | **1.30** | **.19** | **6.80** | **<.0001** | **boːr-jɒ > bor-jɒ** |
| Syllables (default: monosyllabic) | | | | | |
| **polysyllabic** | **0.83** | **.19** | **4.41** | **<.0001** | **ɛlɛm-jɛ > sɛm-jɛ** |

Table 7: Regression model with phonological and morphological predictors of possessive [j], with significant effects bolded

# References

Ackerman, F., Blevins, J. P., and Malouf, R. (2009). Parts and wholes: Implicative patterns in inflectional paradigms. In Blevins, J. P. and Blevins, J., editors, *Analogy in Grammar: Form and Acquisition*, chapter 3, pages 54–82. Oxford University Press, Oxford.

Ackerman, F. and Malouf, R. (2013). Morphological organization: The low conditional entropy conjecture. *Language*, 89(3):429–464.

Albright, A. and Hayes, B. (2003). Rules vs. analogy in English past tenses: a computational/experimental study. *Cognition*, 90(2):119–161.

Becker, M. and Gouskova, M. (2016). Source-oriented generalizations as grammar inference in Russian vowel deletion. *Linguistic Inquiry*, 47(3):391–425.

Corbett, G. G. and Fraser, N. M. (1993). Network Morphology: a DATR account of Russian nominal inflection. *Journal of Linguistics*, 29(1):113–142.

Gouskova, M., Newlin-Łukowicz, L., and Kasyanenko, S. (2015). Selectional restrictions as phonotactics over sublexicons. *Lingua*, 167:41–81.

Harrell Jr., F. E. (2020). *rms: Regression Modeling Strategies*. R package version 6.1-0.

Hayes, B. and Wilson, C. (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry*, 39(3):379–440.

Papp, F. (1969). *A magyar nyelv szóvégmutató szótára*. Akadémiai Kiadó, Budapest.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.

Rácz, P. and Rebrus, P. (2012). Variation in the possessive allomorphy of Hungarian. In Kiefer, F., Ladányi, M., and Siptár, P., editors, *Current Issues in Morphological Theory: (Ir)regularity, analogy and frequency: Selected papers from the 14th International Morphology Meeting, Budapest, 13–16 May 2010*, volume 322 of *Current Issues in Linguistic Theory*, pages 51–64, Amsterdam/Philadelphia. John Benjamins.

Rebrus, P. (2013). Miért nincs *j*? In Benő, A., Fazakas, E., and Kádár, E., editors, *„…hogy legyen a víznek lefolyása…”: Köszöntő kötet Szilágyi N. Sándor tiszteletére*, pages 383–401. Erdélyi Múzeum-Egyesület, Kolozsvár [Cluj-Napoca].

Rebrus, P. and Szigetvári, P. (2018). Morphosyntactically conditioned morphophonological variation. Presentation at PhonVar 1, Barcelona.

Rebrus, P., Szigetvári, P., and Törkenczy, M. (2017). Asymmetric variation. In Lindsey, G. and Nevins, A., editors, *Sonic Signatures: Studies dedicated to John Harris*, volume 14 of *Language Faculty and Beyond*, pages 164–187. John Benjamins, Amsterdam / Philadelphia.