Hungarian speakers use morphological dependencies in inflecting novel forms

Word count: 14,943

Guy Tabachnick
University of Nova Gorica
guyt@ung.si

Abstract Theories of morphology must account for lexicalized variation: lexical items that differ unpredictably in their inflection must be memorized individually and differ in their stored representation. When tested on such cases, adult speakers usually follow the "law of frequency matching" (Hayes et al. 2009), extending gradient phonological patterns from the lexicon. This paper looks at lexicalized variation in the Hungarian possessive: first, I show that a noun's choice of possessive is partially predicted by its plural form as well as its phonological shape. Then, using a novel nonce word paradigm, I show that Hungarian speakers productively apply this cooccurrence pattern between the plural and possessive. I handle lexicalized variation with diacritic features marking lexical entries and propose that Hungarian speakers have learned a gradient cooccurrence relation between diacritic features indexing their plural and possessive forms, extending the sublexicon model of Gouskova et al. (2015). In this proposal, morphological knowledge is distributed across rules in a generative grammar, individual lexical items indexed for their morphological properties, and pattern-matching grammars storing generalizations over those indexed lexical items.

Keywords: frequency matching; diacritic features; productivity; wug test; Hungarian

1 Introduction

One aspect of linguistic knowledge is arbitrary associations between words and the patterns of word formation that they follow. For example, Hungarian speakers know that the words [paːr] 'pair' and [kaːr] 'damage' arbitrarily take distinct possessive suffixes, which I call *-jV* ([paːr-jɒ]) and *-V* ([kaːr-ɒ]). Speakers can also extend patterns productively: given a novel noun, like the English borrowing [baːr], speakers use one of the existing suffixes to form its possessive (in this case, [baːr-jɒ]). This productivity is the object of this paper: What patterns have speakers learned about their language? How do they generalize the arbitrary patterns of known lexical items to unknown ones?

Previous work on morphological productivity has focused on *phonological* factors: for example, nouns ending in [t], like [gaːt] 'dam', are more likely to take *-jV* ([gaːt-jɒ]) than those ending in [r]. In this paper, I focus on another type of generalization: correlations between arbitrary associations of lexical item and pattern (which I call *morphological dependencies*). While [paːr] and [kaːr], like most words, take the plural suffix *-ok*, a small number of words like [ɟaːr] 'factory' instead have plural *-ɒk*; nouns with this plural greatly prefer *-V* (e.g. [ɟaːr-ɒ]). These correlations drive the organization of complex morphological systems (see e.g. Wurzel 1989; Finkel & Stump 2007; Halle & Marantz 2008; Ackerman et al. 2009; Ackerman & Malouf 2013), but speakers' knowledge of them has not been systematically tested. In a nonce word study, I find that speakers learn and productively apply the morphological dependency between plural and possessive alongside phonological generalizations about the distribution of possessives. I show that theoretical tools used to capture speakers' knowledge of phonological generalizations can neatly be applied to morphological dependencies as well, and that these tools complement with a syntactic, piece-based approach to morphological derivations, responding to arguments that the latter ignores morphological dependencies.

1.1 How to infer unknown forms

Linguists typically test speakers' productive use of lexical patterns with nonce word studies: speakers are asked to inflect a made-up word; since they cannot have stored associations between these words and the patterns they take, they must fall back on broader generalizations. When tested on lexically variable patterns through nonce word studies (e.g. Albright & Hayes 2003; Ernestus & Baayen 2003; Gouskova et al. 2015) and artificial language studies (e.g. Hudson Kam & Newport 2005), adults usually follow what Hayes et al. (2009) call the Law of Frequency Matching:

(1) *Law of Frequency Matching (Hayes et al. 2009: 826)*
Speakers of languages with variable lexical patterns respond stochastically when tested on such patterns. Their responses aggregately match the lexical frequencies.

While there are exceptions to this "law" (e.g. Pertsova 2004; Becker et al. 2011), it describes experimental results across a wide range of languages and phenomena. To date, these experiments have generally studied how speakers stochastically generate novel forms according to their *phonological* characteristics. For example, Hayes & Londe (2006) show that Hungarian speakers assign back or front harmony to nonce nouns with ambiguous harmony stochastically according to the particular vowels in the stem.

How can these phonological generalizations be encoded grammatically? A common approach to lexically arbitrary allomorphy is to mark words following a certain pattern with diacritic features: for example, nouns like [gaːt] 'dam' that take possessive *-jV* would be marked with a feature [+j]. Speakers can then extract generalizations over nouns that share a feature, as proposed by the sublexicon model of phonological learning (Allen & Becker 2015; Gouskova et al. 2015; Becker & Gouskova 2016) described in Section 6. Thus, the preference of words like [gaːt] for *-jV* can be expressed as a cooccurrence relation between phonological features and [+j]:

(2) *A phonological generalization over the distribution of Hungarian possessives*
Nouns ending with [coronal,–continuant,–nasal] (that is, alveolar stops) tend to have [+j] (e.g. [gaːt-jɒ] 'her dam')

The phonological properties of a Hungarian noun are not the only source of information about its possessive. Wurzel (1989) notes that a word's form in one cell of its morphological paradigm can be predictive of its behavior in other cells. While Ackerman et al. (2009), Ackerman & Malouf (2013), Parker & Sims (2020), and others have studied the information contained within morphological paradigms, they have not tested whether and how speakers actually use this information. Thus, Bonami & Beniamine (2016) write: "It should be stressed that this paper only established that speakers are exposed to relevant information and that this information is helpful; the next step, of course, is to establish experimentally that speakers do indeed rely on [certain correlations between paradigm cells] when addressing predicting the form of unknown words." This paper shows that Hungarian speakers do, in fact, learn a correlation between paradigm cells (the plural and possessive) from their lexicon.

As I show in Section 6, the sublexicon model's cooccurrence relations between features can also account for morphological dependencies. For example, if nouns like [ɟaːr] with plural *-ɒk* are marked with the diacritic feature [+lower], then the dependency between plural and possessive can be expressed as follows (Halle & Marantz (2008) make a similar proposal for Polish):

(3) *A morphological generalization over the distribution of Hungarian possessives*
Nouns ending with [+lower] (that have plural *-ɒk*) tend to have [–j] (e.g. [ɟaːr-ɒk] 'factories', [ɟaːr-ɒ] 'her factory')

Thus, I propose that speakers learn phonological and morphological dependencies in the same way and apply them together when productively generating unknown forms of new words. This approach captures morphological dependencies without any additional theoretical mechanisms and correctly accounts for the fact that speakers weigh generalizations of different kinds against one another when choosing novel forms of words.

1.2 Road map

In Section 2, I provide a detailed background on the Hungarian plural and possessive, and Section 3 contains a formal analysis. In Section 4, I present a corpus data showing the phonological and morphological factors predicting the distribution of possessive *-V* and *-jV* in the Hungarian lexicon. In Section 5, I present a nonce word study showing that Hungarian speakers productively extend these phonological and morphological generalizations to novel forms. Section 6 describes a theory of phonological and morphological generalization based on Gouskova et al. (2015). Section 7 concludes with considerations of the theoretical import of my proposal.

2 Background

In the introduction, I described lexical variation in two Hungarian suffixes: the possessive and the plural. In this section, I provide more background about this variation and about a related source of suffix alternations, vowel harmony.

As mentioned above, the possessive has two basic allomorphs, *-V* and *-jV*, both of which are very frequent. In the plural, most nouns have *-ok*, while a small class called “lowering stems” instead takes *-ɒk*. Table 1 shows that all four combinations of plural and possessive are possible.

noun	dɒl	tʃont	vaːlː	hold
gloss	‘song’	’bone’	‘shoulder’	‘moon’
plural	dɒl-o̲k	tʃont-o̲k	vaːlː-ɒ̲k	hold-ɒ̲k
possessive	dɒl-ɒ	tʃont-**j**ɒ	vaːlː-ɒ	hold-**j**ɒ

Table 1: Possible combinations of Hungarian plural and possessive suffixes

Nouns are not evenly distributed between the four options, as described in (3): most lowering stems, like [vaːlː] ‘shoulder’, take *-V*, and only a few, like [hold] ‘moon’, take *-jV*.

2.1 Vowel harmony

Hungarian words have either back or front harmony, and suffix vowels alternate accordingly. The mid suffixes also show rounding harmony: front-harmonizing suffixes with mid vowels have rounded and unrounded variants to match the last vowel of the stem. These alternations, for short vowels, are shown in Table 2; see Siptár & Törkenczy (2000: 63–73) for more details. One striking asymmetry is that low and mid vowels are not differentiated for all harmony classes: the same vowel, [ɛ], is the low counterpart of mid [ø] for words with front rounded harmony, and both the low and mid vowel for words with front unrounded harmony. There is no distinct front low vowel like [æ].[1]

height	*front* *unrounded*	*front* *rounded*	*back*	*example suffix*			*example words* kɛrt ‘garden’	føld ‘land’	haːz ‘house’
high	y		u	*-unk*/*-ynk*	1PL	‘our’	kɛrt-ynk	føld-ynk	haːz-unk
mid	ɛ	ø	o	*-hoz*/*-høz*/*-hɛz*	ALL	‘to’	kɛrt-hɛz	føld-høz	haːz-hoz
low			ɒ	*-bɒn*/*-bɛn*	INESS	‘in’	kɛrt-bɛn	føld-bɛn	haːz-bɒn

Table 2: Hungarian suffix vowel harmony alternations (from Siptár & Törkenczy 2000: 65)

Examples in this paper have back harmony. Table 2 can be used to find the front-harmonizing version of each suffix. Thus, the front-harmonizing equivalents of possessive *-ɒ* and *-jɒ* are *-ɛ* and *-jɛ*. Regular-stem plural *-ok* (from the mid vowel set) has two front-harmonizing variants, depending on rounding, *-øk* (e.g. [ʃyl-øk] ‘porcupines’) and *-ɛk*, while the lowering stem plural *-ɒk* (from the low vowel set) only has one front-harmonizing variant, *-ɛk* (e.g. [fyl-ɛk] ‘ears’). Words with front unrounded harmony can

[1] This section describes the standard language; some Hungarian dialects distinguish between phonologically low [ɛ] and mid [e].

only have plural *-ɛk* and thus cannot be distinguished on the surface as lowering stems. Siptár & Törkenczy (2000: 225) mark some nouns with front unrounded harmony as lowering stems on the basis of other properties that correlate (more or less reliably) with lowering stem status. Since this difference is not marked in my corpus and cannot be reliably inferred, I assume that all words with front unrounded harmony are undetermined for stem class. In the nonce word experiment (Section 5), I treat stimuli with front unrounded harmony as fillers.

A stem's harmony class is usually but not always predictable from its vowels (Siptár & Törkenczy 2000; Hayes & Londe 2006; Hayes et al. 2009; Rebrus et al. 2012; 2019)—thus, at least some nouns must be explicitly marked for harmony class. However, I assume that vowel harmony is handled *in the phonology proper*, unlike the distinction between possessive *-V* and *-jV*: *-ɒ* and *-ɛ* are surface variants of a single underlying form, as are *-jɒ* and *-jɛ*.

2.2 The possessive

Table 3 shows the full paradigm of possessives for the four words in Table 1 (see Rounds 2008: 135–137). Hungarian distinguishes between the person and number of possessors, as well as the number of the possessed noun, so [dɒl-om] means 'my song', while [dɒl-ɒ-i-m] means 'my songs', and so on.

noun	dɒl	tʃont	vaːlː	hold
gloss	'song'	'bone'	'shoulder'	'moon'
possessor		*singular noun*		
1SG	dɒlom	tʃontom	vaːlːɒm	holdɒm
2SG	dɒlod	tʃontod	vaːlːɒd	holdɒd
3SG	dɒlɒ	tʃont**j**ɒ	vaːlːɒ	hold**j**ɒ
1PL	dɒlunk	tʃontunk	vaːlːunk	holdunk
2PL	dɒlotok	tʃontotok	vaːlːɒtok	holdɒtok
3PL	dɒluk	tʃont**j**uk	vaːlːuk	hold**j**uk
possessor		*plural noun*		
1SG	dɒlɒim	tʃont**j**ɒim	vaːlːɒim	hold**j**ɒim
2SG	dɒlɒid	tʃont**j**ɒid	vaːlːɒid	hold**j**ɒid
3SG	dɒlɒi	tʃont**j**ɒi	vaːlːɒi	hold**j**ɒi
1PL	dɒlɒink	tʃont**j**ɒink	vaːlːɒink	hold**j**ɒink
2PL	dɒlɒitok	tʃont**j**ɒitok	vaːlːɒitok	hold**j**ɒitok
3PL	dɒlɒik	tʃont**j**ɒik	vaːlːɒik	hold**j**ɒik

Table 3: Hungarian possessive paradigms for some back-harmonizing words

There are two main points of variation among these paradigms. The first is the alternation between [o] and [ɒ] (underlined in Table 3) in the 1SG, 2SG, and 2PL singular. This is the same lowering stem alternation as in the plural, and will be addressed in Section 2.3. The second is the variable presence of [j] (bolded in Table 3) in singular nouns with 3SG and 3PL possessors and plural nouns with all possessors.[2] This is the possessive morpheme, with allomorphs *-V* and *-jV*. Its vowel deletes before 3PL *-uk*.

Under the standard syntactic analysis (cf. Bartos 1999; É. Kiss 2002; Dékány 2018), *-V* and *-jV* are realizations of a Poss head, which has a zero allomorph when adjacent to

[2] Usually, [j] is either present or absent throughout the paradigm. One very rare exception is [bɒraːt] 'friend', which takes *-jV* in the singular ([bɒraːt-jɒ] 'her friend') and *-V* in the plural ([bɒraːt-ɒ-i] 'her friends').

a first- or second-person possessor marker. (The marker for third singular possessors is null.) Thus, I gloss *-V* and *-jV* as POSS (not 3SG), while *-(V)m*, *-(V)d*, etc. mark 1SG, 2SG, and so on.

2.3 Linking vowels

In Section 2.2, I showed that the 1SG, 2SG, and 2PL possessor markers undergo the same lowering stem alternation as the plural. Table 4 shows these forms again for the regular stem [dɒl] 'song' and the lowering stem [vaːlː] 'shoulder', as well as forms of [kɒpu] 'gate'. The non-possessed plural and the three possessive markers show the same pattern: the suffix is a bare consonant (or *-tok*) after a vowel, including the vowel-final noun [kɒpu] and *-i*, which marks plural on possessed nouns. Otherwise, this consonant is preceded by a "linking vowel" (cf. Siptár & Törkenczy 2000: 219), which is mid after [dɒl] and low after [vaːlː].

noun	dɒl		vaːlː		kɒpu	
gloss	'song'		'shoulder'		'gate'	
possessor	*singular*	*plural*	*singular*	*plural*	*singular*	*plural*
none	dɒl	dɒlok	vaːlː	vaːlːɒk	kɒpu	kɒpuk
1SG	dɒlom	dɒlɒim	vaːlːɒm	vaːlːɒim	kɒpum	kɒpuim
2SG	dɒlod	dɒlɒid	vaːlːɒd	vaːlːɒid	kɒpud	kɒpuid
2PL	dɒlotok	dɒlɒitok	vaːlːɒtok	vaːlːɒitok	kɒputok	kɒpuitok

Table 4: Lowering stem alternations in the plural and possessive markers

This suggests an analysis of lowering stems that abstracts over the multiple linking vowel suffixes. I present this analysis, alongside an account of possessive allomorphy, in Section 3.

3 Formal analysis

In this section, I present a formal analysis of Hungarian possessive allomorphy and lowering stems. This analysis is somewhat simplified, but is sufficient for the account of phonological and morphological dependencies in Section 6.

3.1 Possessive allomorphy

In the introduction, I suggested that possessive allomorphy is indexed by a diacritic feature: nouns that take *-jV* are marked with [+j] and, analogously, nouns that take *-V* with [−j]. In (4), we see that these features govern allomorphy selection by appearing in rules of realization that spell out the possessive (ignoring vowel harmony).

(4) *Rules of realization for the Hungarian possessive*
   a. POSS ⟷ jɒ / [+j] __
   b. POSS ⟷ ɒ / [−j] __

In the assumptions of Distributed Morphology (Halle & Marantz 1993; Embick & Marantz 2008) with root-outward vocabulary insertion (Bobaljik 2000), the noun root is inserted first. Its underlying form bears a lexical feature indexing its possessive: for example, the underlying form of [dɒl] 'song' is /dɒl$_{[-\mathrm{j}]}$/.[3] This [–j] feature is then visible during spellout of the possessive, matching the context in (4b). Thus, the possessive is spelled out as *-ɒ*, yielding [dɒl-ɒ].

Following Gouskova et al. (2015), whose framework I adopt in Section 6, I assume that both of the rules in (4) are marked with a diacritic feature; that is, there is no default rule, and every noun root must have a [±j] feature in order to get a possessive form. Rácz & Rebrus (2012), in contrast, that *-jV* is a productive default for most words (meaning that (4a) should be unmarked). Although I adopt the symmetrical, default-free pattern for theoretical reasons, my experimental results do not find evidence of default behavior: participants use *-V* and *-jV* for the same nonce words.

3.2 Lowering stems and the plural

In Section 2.3, I showed that the plural and some possessive suffixes show a three-way distinction between *-C* (after vowel-final stems), *-oC* (after regular stems), and *-ɒC* (after lowering stems). In (5), I assume that these suffixes lack the linking vowel underlyingly, but are marked with a feature, [LV],[4] indicating that they undergo linking vowel alternations:

(5) *Rules of realization for linking vowel suffixes*
  a. PL ⟷ $\mathrm{k}_{[\mathrm{LV}]}$
  b. 1SG ⟷ $\mathrm{m}_{[\mathrm{LV}]}$
  …

Readjustment rules can then insert the appropriate linking vowel after consonants. In particular, lowering stems like [vaːlː] 'shoulder', marked with [+lower] in addition to their possessive feature (/vaːlː$_{[+\mathrm{lower},-\mathrm{j}]}$/), trigger insertion of the low linking vowel [ɒ], while most consonant-final nouns—marked with a complementary [–lower] feature, as in /dal$_{[-\mathrm{lower},-\mathrm{j}]}$/ 'song'—trigger insertion of the linking vowel [o].

(6) *Readjustment rules for linking vowels*
  a. Ø → ɒ / [+lower] __ [LV]
  b. Ø → o / [–lower] __ [LV]

This analysis correctly predicts that a noun that has a low linking vowel in one suffix (e.g. the plural) will have a low linking vowel in all suffixes. It also enables speakers to learn the morphological dependency between lowering stems and the possessive as a correlation between features [+lower] and [–j], as shown in (3) (see Section 6.3).

3.3 The representation of lowering stems

The analysis in the previous section assumes that the lowering alternation is encoded *morphologically*: lowering stems are marked with [+lower], and suffixes with linking vowels have an [LV] feature. Siptár & Törkenczy (2000) instead propose (in Section 8.1.4) an abstract *phonological* analysis: lowering stems have a floating low feature [+open$_1$] and linking vowel suffixes have an underlying vowel unspecified for height. This vowel surfaces as low in the presence of [+open$_1$], otherwise it surfaces as mid after consonants and deletes after vowels.

[3] Like Gouskova & Bobaljik (2022), but unlike many others (e.g. Müller 2004; Embick & Halle 2005; Kramer 2015), I assume that diacritic features are *phonological* properties of *exponents* spelling out syntactic nodes, rather than syntactic properties of the nodes themselves (see Tabachnick 2023: 52–58).

[4] The linking vowel in these suffixes is not predictable from phonotactics, and so must be marked (Siptár & Törkenczy 2000: 219).

These analyses represent two approaches to morphophonologically exceptional morphemes. In my analysis in (4), (5), and (6), exceptional lexical items are marked with a diacritic indexing a morpheme-specific rule or constraint (e.g. Pater 2010; Gouskova 2012; Rysling 2016). Siptár & Törkenczy (2000) instead use defective segments and subsegmental units that cannot surface faithfully and behave differently from full segments (e.g. Lightner 1965; Rubach 2013; Trommer 2021).

The two approaches are not mutually exclusive (for example, Chomsky & Halle (1968) use both), and the choice between them is often one of elegance and coverage. Moreover, both have been criticized on similar grounds: Pater (2006) and Gouskova (2012) argue that accounts with abstract underlying forms can overgenerate and be hard to learn, while Bermúdez-Otero (2012; 2013), Haugen (2016), and Caha (2021) argue that arbitrary lexical marking and readjustment rules are unrestrained and weaken our theory of grammar. In this case, the two analyses are largely equivalent: for Siptár & Törkenczy (2000), the floating feature has no phonological effect beyond producing a low linking vowel, making it akin to what Kiparsky (1982) calls "purely diacritic use of phonological features".[5]

This paper argues that Hungarian speakers learn generalizations over [+lower], a feature marking lowering stems. For Siptár & Törkenczy (2000), the floating [+open$_1$] feature is similarly unique to lowering stems. Both analyses are thus compatible with my main hypothesis that speakers learn generalizations over features that index unpredictable morphophonological behavior.

With this empirical and formal background, we can turn to quantitative study of the distribution of possessive allomorphs and lowering stems and begin to test the correlation between the two.

4 Possessive allomorphy in the Hungarian lexicon

The goal of this paper is to show that Hungarian speakers extend gradient patterns in their lexicon to nonce words, in particular the morphological generalization that lowering stems prefer possessive -*V*. To do this, I must first show what the lexical patterns are. This section presents a corpus study of the Hungarian lexicon that serves as the foundation for the nonce word study in Section 5.

4.1 Representing the Hungarian lexicon

In this section I discuss my representation of the Hungarian lexicon.[6]

My source of data is Papp (1969), a printed morphological dictionary of Hungarian which I digitized manually. I use Papp (1969) for its comprehensive tagging of derivational morphology, but it has potential disadvantages: it is over 50 years old and reflects lexicographic work rather than pure corpus data. In Section 4.2.3, I compare my corpus with that of Rácz & Rebrus (2012), who use a web corpus. The distribution of possessives is quite similar, so I conclude that the dictionary is a relatively accurate representation of contemporary Hungarian.

[5] "Self-lowering" (Siptár & Törkenczy 2000: 228–229) verbal suffixes, which show a vowel–zero alternation whose vowel is always low, could potentially distinguish the two analyses. However, Siptár & Törkenczy (2000) argue that the "self-lowering" alternation is morphological (allomorph selection) rather than phonological (underspecification), converging with my analysis for these cases.

[6] All alternative analyses, comparisons, etc. mentioned in this and the next study are presented in the supplemental materials.

Under standard assumptions in Distributed Morphology, lexical information like allomorph selection is stored for roots and affixes, not complex stems (Embick & Marantz 2008). Thus, if speakers are generalizing over the frequency of types in the lexicon (cf. Bybee 1995; 2001; Pierrehumbert 2001; Albright & Hayes 2003; Hayes & Wilson 2008; Hayes et al. 2009), derived words and compounds with the same head (rightmost affix or root) should not count as separate types. Root-based storage predicts that words ending in the same suffix should take the same possessive, which is largely true in Hungarian (Rácz & Rebrus 2012). I adopt the assumption of root-based storage by limiting my corpus to monomorphemic nouns. In Section 5.7, I suggest that this corpus more accurately reflects the behavior of participants in the nonce word study than a corpus including complex nouns.

Although adjectives can also take possessive suffixes, I limit my corpus to nouns. Unlike nouns, most adjectives are lowering stems (Siptár & Törkenczy 2000: 229–230), so including adjectives would complicate the relationship between lowering stems and possessive allomorphy (see also Rebrus & Szigetvári 2018). Of 35,130 nominals in my corpus, 5,055 are monomorphemic, and 4,443 of those are listed exclusively as nouns. I excluded 1,768 vowel-final words, since these categorically take *-jV* and would be undefined for many of the factors in my regression. I also removed the 27 words ending in orthographic *h*, which is phonologically complicated (Siptár & Törkenczy 2000: 274–276). Finally, I excluded 216 nouns listed in Papp (1969) with variable or unknown possessive to allow for binary coding of the possessive variable (*-V* vs. *-jV*). This leaves 2,432 noun types.

4.2 Corpus study: the distribution of *-V* and *-jV* in the lexicon

In this section, I present the results of a corpus study showing phonological and morphological factors predicting a noun's possessive allomorph in the Hungarian lexicon. Like other cases of lexically specific variation (e.g. Hayes et al. 2009; Becker et al. 2011; Gouskova et al. 2015), the distribution of *-V* and *-jV* shows gradient tendencies; these involve both phonological properties and the morphological dependency of stem class. In Section 5, I test whether speakers productively apply this morphological effect to new forms.

4.2.1 Analysis

In this study, I look at various phonological properties of the stem and one morphological property, stem class. The full model of the lexicon is shown in Section 4.2.2.2. This is a logistic regression with possessive suffix as the dependent variable (*-V* is represented as 0, *-jV* as 1; higher coefficients represent a higher likelihood of *-jV*). The goal of this model is to present an accurate representation of the lexicon, so I initially considered a large number of phonological properties representing the right edge of a stem (local to the suffix): the place (alveolar, labial, palatal, velar) and manner (plosive, non-sibilant fricative, sibilant fricative/affricate, approximant) of its final consonant; the height (mid, high, low), length (short, long), backness (back, front), and roundedness (unrounded, rounded) of its final syllable's vowel; its vowel harmony class (back, front, variable); the complexity of its final coda (singleton, geminate, cluster); and its length (polysyllabic, monosyllabic).[7] All variables were dummy coded with the first listed level (the most

[7] Hungarian has fixed word-initial stress, so this factor also marks whether the suffix is attaching to the stressed syllable.

frequent) as the reference level. Alternating stems were considered in their form when suffixed: for example, [mɒjom] ‘monkey’, which displays a vowel–zero alternation (e.g. [mɒjm-ɒ] ‘her monkey’), is coded as ending in a cluster with [ɒ] as its last vowel. For stem class, nouns were classified as lowering, non-lowering, variable, or indeterminate (nouns with front unrounding harmony, see Section 2.1). This factor was dummy coded with non-lowering as the reference level.

The model was assembled by forward stepwise comparison using the `buildmer` function in R from the package of the same name (R Core Team 2022; Voeten 2022).[8] This function adds factors one at a time such that each additional factor improves (lowers) the model’s Akaike Information Criterion (AIC), which measures how well the model fits the data while penalizing model complexity (number of factors). All included factors lowered the AIC by at least 16.5. Two candidate factors were not added to the model: vowel rounding and vowel backness, the latter of which largely overlaps with harmony class. The model equation is: *possessive* ∼ *C manner* + *C place* + *stem class* + *V height* + *harmony* + *coda* + *V length* + *syllables*.

I also fitted two intermediate regressions: one with only phonological factors, and the other with only the morphological factor of stem class. I present the phonological model in Section 4.2.2.1. I include phonology and stem class as separate factors in my model of the nonce word experiment in Section 5.6.2 (see Section 5.7.1 for discussion), and the phonological model of the lexicon represents speakers’ knowledge of phonological generalizations over possessive allomorphy. Like the full model, vowel rounding did not improve the phonological model’s AIC and it was not added. Unlike in that model, however, vowel backness did slightly improve the AIC (by −.05). I excluded this factor for two reasons: first, the improvement in AIC was very slight, and indeed, adding this factor did not significantly improve the model’s fit ($\chi^2 = 2.05$, $p = .152$).[9] Second, as mentioned previously, this factor is largely conflated with harmony class; including both leads to greater collinearity between the factors and makes the results less interpretable. Thus, the phonological factors are the same in the two models, though added in a slightly different order. The final equation of this model is: *possessive* ∼ *C manner* + *C place* + *harmony* + *V height* + *V length* + *coda* + *syllables*.

I do not present the model predicting possessive from the morphological factor of stem class alone (*possessive* ∼ *stem class*) because I do not use it in the nonce word study: as I discuss in Section 5.5, for my analysis of the experimental results I encode stem class categorically (lowering vs. not).

4.2.2 Results

4.2.2.1 Phonology

Table 5 contains the full model with phonological factors listed in the order in which they were added to the model (roughly corresponding to importance). Most of the factors are significant. The most influential are the place and manner of the final consonant. This effect strength is likely driven by the near-categorical effects of sibilants and palatals, which have the strongest negative effect size (favoring -*V*). Other places and manners have significant effects as well, as do other phonological factors: front-harmonizing words take -*jV* less than back-harmonizing words, and nouns ending in geminates prefer -*jV* relative

[8] Version 2.9 of the package was used in version 4.3.1 of R. Because `buildmer` models are incompatible with several of the ancillary functions I applied to my model, I used the formula generated by `buildmer` to fit a model using the `glm` function in R’s basic *stats* package.

[9] Comparisons were conducted using `lrtest` in version 6.7-0 of R’s *rms* package (Harrell Jr. 2020).

to nouns ending in singleton consonants. The model predicts a word's possessive quite well (Tjur's $R^2 = .68$).[10]

	ΔAIC	β coef	SE	Wald z	p
Intercept		**4.16**	**.30**	**13.85**	**<.0001**
C Manner (default: plosive)	-951.1				
fricative		**-1.44**	**.39**	**-3.73**	**.0002**
sibilant		**-10.69**	**.80**	**-13.38**	**<.0001**
nasal		**-1.95**	**.27**	**-7.16**	**<.0001**
approximant		**-4.08**	**.30**	**-13.47**	**<.0001**
C Place (default: alveolar)	-719.8				
labial		**-2.02**	**.26**	**-7.94**	**<.0001**
palatal		**-8.88**	**1.10**	**-8.06**	**<.0001**
velar		**-3.26**	**.29**	**-11.19**	**<.0001**
Harmony (default: back)	-224.0				
front		**-2.03**	**.18**	**-10.96**	**<.0001**
variable		**2.27**	**.97**	**2.33**	**.0197**
V Height (default: mid)	-68.5				
high		**1.73**	**.22**	**7.89**	**<.0001**
low		0.28	.19	1.50	.1342
V Length (default: short)	-44.4				
long		**1.40**	**.17**	**7.98**	**<.0001**
Coda (default: singleton)	-44.2				
geminate		**2.47**	**.40**	**6.25**	**<.0001**
cluster		0.04	.21	0.17	.8617
Syllables (default: polysyllabic)	-44.4				
monosyllabic		**-1.15**	**.17**	**-6.67**	**<.0001**

Table 5: Regression model with phonological predictors of possessive *-jV*, with significant effects bolded

For a given word, this regression calculates a coefficient x which measures the predicted probability P that that word takes *-jV*, $P = \frac{e^x}{1+e^x}$. This coefficient is the sum of the β coefficients of the intercept and a word's value for each factor when it differs from the default. The model can predict the possessive of nonce words as well. For example, the nonce word [lufɒn] has a coefficient of $\beta_{\text{Intercept}} + \beta_{\text{C place: nasal}} + \beta_{\text{V height: low}} = 4.16 - 1.95 + 0.28 = 2.49$, corresponding to a probability of $\frac{e^{2.49}}{1+e^{2.49}} = .923 = 92.3\%$: if this were a real word, its possessive would likely be [lufɒn-jɒ]. I refer to these coefficients as *phon_odds* and use them as predictors of the nonce word experiment in Section 5.6.

4.2.2.2 Phonology and morphology

Adding stem class to the model significantly improves it ($\chi^2 = 112.9$, $p < .0001$), raising the correlation to Tjur's $R^2 = .71$. Stem class is significant and the most important factor after final C manner and place. Otherwise, the new model, shown in Table 6, is very similar to the phonological model in Table 5: the same phonological factors are added to the model and the effect sizes are quite similar. The effect of lowering stems is strongly

[10] This and other correlation coefficients were calculated using the `r2` function of R's *performance* package (Lüdecke et al. 2021).

negative: independent of their phonology, lowering stems are more likely to take *-V* than non-lowering stems. The difference between non-lowering stems and nouns with undetermined stem class is smaller and not significant. As discussed in Section 2.1, this class comprises nouns with front unrounded harmony, so its effect should be masked by the factor of harmony.

	ΔAIC	*β coef*	*SE*	*Wald z*	*p*
Intercept		**4.32**	**.31**	**13.99**	**<.0001**
C Manner (default: plosive)	−951.1				
fricative		**−1.03**	**.44**	**−2.37**	**.0179**
sibilant		**−11.07**	**.80**	**−13.87**	**<.0001**
nasal		**−2.07**	**.28**	**−7.39**	**<.0001**
approximant		**−4.06**	**.31**	**−13.10**	**<.0001**
C Place (default: alveolar)	−719.8				
labial		**−2.22**	**.27**	**−8.35**	**<.0001**
palatal		**−9.25**	**1.13**	**−8.22**	**<.0001**
velar		**−3.54**	**.31**	**−11.55**	**<.0001**
Stem class (default: non-lowering)	−241.3				
lowering		**−3.71**	**.44**	**−8.44**	**<.0001**
undetermined		−0.25	.25	−0.98	.3278
variable		**−2.76**	**.69**	**−4.00**	**<.0001**
V Height (default: mid)	−114.8				
high		**1.85**	**.23**	**8.09**	**<.0001**
low		**0.77**	**.21**	**3.66**	**.0003**
Harmony (default: back)	−81.6				
front		**−1.98**	**.27**	**−7.41**	**<.0001**
variable		**2.25**	**1.04**	**2.17**	**.0297**
Coda (default: singleton)	−39.2				
geminate		**2.43**	**.41**	**5.97**	**<.0001**
cluster		−0.08	.22	−0.36	.7147
V Length (default: short)	−38.7				
long		**1.30**	**.19**	**6.97**	**<.0001**
Syllables (default: polysyllabic)	−16.8				
monosyllabic		**−0.79**	**.18**	**−4.31**	**<.0001**

Table 6: Regression model with phonological predictors of possessive *-jV* and stem class, with significant effects bolded

I confirmed the independence of stem class by testing its variance inflation factor (VIF) using the `check_collinearity` function from R's *performance* package (Lüdecke et al. 2021). This measures whether different factors are describing the same effect. Stem class had a low correlation (a VIF of 2.96) with the other factors (see James et al. 2013), meaning that its effect cannot be reduced to some combination of phonological factors.

4.2.3 Discussion

The corpus study shows that a number of phonological factors and stem class are good predictors of a noun's possessive. Some of these results differ from those of Rácz & Rebrus (2012). In this section, I compare their results with mine, concluding that the differences are due to analytical choices.

There are two key differences between my study and that of Rácz & Rebrus (2012). First, they use counts rather than statistical modelling. As a consequence, they must be careful about how they count: they generally exclude nouns ending in palatals and sibilants, which nearly categorically take *-V*. Statistical models, on the other hand, balance numerous factors together without the need for exclusion. To facilitate comparison, I present counts from my corpus below using the same calculations as they use.

The second difference is in the choice of corpus. Rácz & Rebrus (2012) tabulate type and token counts from an unlemmatized web corpus that includes all words appearing with possessive suffixes: monomorphemic nouns; derived nouns; and occasional adjectives, numerals, and similar. Their results, shown in Table 7, include three phonological factors for consonant-final nouns not ending in sibilants or palatals: final consonant place (limited to stops), final coda complexity, and vowel harmony class.

	tokens (thousands)			*types (thousands)*		
	-V	*-jV*	*% -jV*	*-V*	*-jV*	*% -jV*
labial plosive	126	186	60%	0.3	0.3	50%
alveolar plosive	1039	350	25%	1.7	1.4	45%
velar plosive	1706	150	8%	3.2	0.5	14%
singleton	4158	619	13%	7.9	2.4	20%
geminate/cluster	395	287	42%	0.9	1.2	57%
back harmony	1817	789	30%	4.2	2.6	38%
front harmony	2737	117	4%	4.6	1.0	18%
total	4554	906	17%	8.8	3.6	29%

Table 7: Phonological distribution of possessive allomorphs from Rácz & Rebrus (2012: 57–59) for nominals ending in consonants that are not palatal or sibilant

The corpus in this study differs from theirs in two ways: first, mine is limited to monomorphemic nouns. Second, my corpus is derived from Papp (1969), an older dictionary (see Section 4.1). To distinguish these two factors, Table 8 includes counts from my corpus and two supersets: one also including complex nouns, and one including all words with listed possessive including adjectives and others.

	monomorphemic nouns			*all nouns*			*all nominals*		
	-V	*-jV*	*% -jV*	*-V*	*-jV*	*% -jV*	*-V*	*-jV*	*% -jV*
labial plosive	23	76	76.8%	288	380	56.9%	291	406	58.2%
alveolar plosive	13	359	96.5%	1594	1269	44.3%	1633	1502	47.9%
velar plosive	126	224	64.0%	2796	667	19.3%	2851	743	20.7%
singleton	337	846	71.5%	7749	2323	23.1%	7944	2652	25.0%
geminate/cluster	348	122	74.0%	983	1121	53.3%	996	1343	57.4%
back harmony	195	921	82.5%	4024	2700	40.2%	4072	3062	42.9%
front harmony	264	235	47.1%	4698	685	12.7%	4858	858	15.0%
total	459	1194	72.2%	8732	3444	28.3%	8940	3995	30.9%

Table 8: Phonological distribution of possessive allomorphs from Papp (1969) for subgroups of nominals ending in consonants that are not palatal or sibilant

Table 8 shows that including complex nouns makes a drastic difference: most saliently, the overall rate of *-jV* is much lower across the board in complex nouns than monomorphemic nouns, in part because most derivational suffixes categorically take *-V* (Rácz &

Rebrus 2012). Monomorphemic nouns ending in [t d], like [hold] 'moon', overwhelmingly take *-jV* (e.g. [hold-jɒ]); according to Table 9, monomorphemic nouns ending in alveolars take *-jV* significantly more often than those ending in other consonants. However, when complex nouns are taken into account, nouns ending in [t d] actually take *-jV less* than those ending in labial stops. This larger corpus includes many nouns with suffixes like the nominalizer *-ɒt*, which takes *-V*, as in [viʒgaːl-ɒt-ɒ] 'her examination' (from [viʒgaːl] 'examine'). Likewise, monomorphemic nouns ending in singletons take *-jV* slightly less than others (in Table 9, geminates significantly prefer *-jV* relative to singletons), but the gap is much wider among complex nouns. Including adjectives and other nominals has little effect: the rates of *-jV* between all nouns and all nominals differ by at most 4.1% in each category.

The percentages of Rácz & Rebrus (2012) in Table 7 are very similar to those for all nominals in Table 8, differing by at most 8.2% in any category. The relative rates of *-jV* between different phonological properties are quite consistent. Differences between my results and those of Rácz & Rebrus (2012) are thus due overwhelmingly to my use of monomorphemic nouns only. This is confirmation that Papp (1969) is sufficiently similar to contemporary sources to be an adequate representation of Hungarian. The comparison also casts doubt on explanations Rácz & Rebrus (2012) offer for their effects. They suggest that the relatively high rates of *-jV* for nouns ending in coronals, labials, and complex codas are likely due (at least in part) to the nearly categorical preference for *-jV* among two geminate-final suffixes: the past participle *-tː* (e.g. [ɒlkɒlmɒz-otː-jɒ] 'her employee') and the comparative *-bː*.[11] However, these effects are equally strong among words listed in Papp (1969) as nouns—which excludes essentially all comparatives and most past participles, which are usually listed as "adjective and noun". Thus, these suffixes are not the primary drivers of asymmetries in the distribution of *-jV*. In fact, inspection of Papp (1969) suggests that the complex coda effect is driven primarily by the large number of compounds ending in clusters (e.g. [uːj-hold-jɒ] 'her new moon') and, conversely, the large number of derived forms—which, again, mostly take *-V*—ending in singletons.

This section shows that corpus choice can have a substantial effect on the distribution of possessive allomorphs, especially on the baseline rate of *-jV*. The choice between corpora is theoretically driven—for example, my use of monomorphemic nouns is grounded in a theory of morphology that stores roots rather than stems—and an empirical question: which corpus more closely reflects Hungarian speakers' behavior? In Section 5.7.3, I suggest that the monomorphemic corpus is a better match for my experimental results. However, the primary results of this paper (Hungarian speakers apply learned generalizations about the possessive according to both the phonology and the stem class of nouns) are robust no matter the corpus.

5 A nonce word study of the Hungarian morphological dependency

In this section, I present a nonce word study testing whether Hungarian speakers productively apply gradient phonological and morphological effects on possessive allomorphy from the lexicon. While previous nonce word studies have focused on phonological generalizations (e.g. Hayes et al. 2009; Becker et al. 2011; Gouskova et al. 2015), I show that speakers apply a morphological dependency as well: nonce words are assigned *-V* more often when presented as lowering stems (with plural *-ɒk*). To test this, I use a novel

[11] These suffixes both take linking vowels (see Section 2.3).

extension of the nonce word paradigm: in most nonce word studies, speakers are presented with a single form and make inferences based on its phonological properties. In this study, I change both the base form and a second, inflected, form, which provides information to participants about the nonce word's inflectional patterns. This novel experimental condition enables me to test the psychological reality of correlations between morphological patterns.

5.1 Predictions

I hypothesize that speakers form the possessives of novel words by matching the distribution of *-V* and *-jV* in the lexicon (described in Section 4) according to both phonological factors and stem class, following the "law of frequency matching" of Hayes et al. (2009) (see Section 1.1). Since my primary concern is the morphological dependency, I focus on phonological frequency matching in the aggregate rather than individual phonological effects.

Rácz & Rebrus (2012) argue that *-jV* is the productive default for most words (see Section 2.2). If this is true, speakers should instead categorically assign *-jV* to most words.

5.2 Participants

Subjects were recruited through Prolific (https://app.prolific.co/) and were born in Hungary and raised as monolingual Hungarian speakers. I recruited 30 participants for the stimulus norming study and 91 for the stimulus testing study. One participant was rejected for poor quality (see Section 5.4.2), and an additional 48 subjects were recruited for earlier versions of the stimulus testing study; because the experimental task was substantially different (fill-in-the-blank rather than forced-choice, or missing the attention check for the plural), their data are not presented here.

5.3 Stimuli

I trained the UCLA Phonotactic Learner (Hayes & Wilson 2008) on the corpus of Hungarian nouns used in Section 4.1. The program produces a "sample salad" of 1,968 nonce words randomly generated from the probability distribution defined by the phonotactic grammar it learned. Many of the generated words included long strings of consonants like [ʒuːgkfkb]. These ridiculous words reflect the limitations of the learned grammar: in particular, its constraints were limited to bigram sequences of segments or natural classes (that is, two segments/classes long) plus a word boundary, so the learner could not learn restrictions against long strings of consonants. The learned grammar was much less restrictive than actual Hungarian, so words like [ʒuːgkfkb] were filtered out manually. However, the grammar was good enough that many of its generated words looked like reasonable Hungarian words (like [tuːs]), and some were actually existing Hungarian words, like [iːɲ] 'gum'.

I selected the 494 generated nonce words with the shape (C)VC(C) or (CV)CVC(C) and removed 162 disyllabic disharmonic words (which had one front vowel and one back vowel). Finally, I removed 19 generated stimuli that were headwords of any part of speech in Papp (1969).[12] This left a final set of 317 nonce word stimuli, some of which were still, impressionistically, phonologically questionable (e.g. [ɲɒsm]). Each word was presented in the singular and plural, in the latter case with either a regular or lowering plural suffix.

5.4 Procedure

This experiment was split into two studies. First, subjects rated all stimuli for plausibility as Hungarian words. The ratings obtained in this study were used to select a smaller set of stimuli for the main experiment, in which subjects selected possessive forms.

5.4.1 Stimulus norming and selection

The goal of this study was to confirm the intuitions of the author (not a native Hungarian speaker) that the stimuli were broadly plausible as Hungarian words, including when presented as lowering stems. Participants completed 50 trials, each with a different stimulus. Stimuli were chosen randomly: five words never appeared at all, and the remaining 312 words were used in 1–10 trials each (mean 4.8 trials, standard deviation 2.1). An example trial is shown in Figure 1. Each trial had a frame sentence containing the target stimulus twice, presented in written form with regular Hungarian orthography (*lufan*, *lufanok*). In its first occurrence, the stimulus appeared in bare nominative form; the second time, the stimulus had a plural suffix (and sometimes subsequent suffixes). Most stimuli were shown with regular plurals (e.g. *-ok*), but 8 randomly chosen trials instead showed stimuli as lowering stems (e.g. with plural *-ɒk*).[13] Participants rated each stimulus on a five-point scale according to the question *Could the underlined words be Hungarian?*, with the ends of the scale labelled *no* (1) and *yes* (5). Frame sentences, stimuli, and the sample trial in the original Hungarian are included in the supplemental materials.

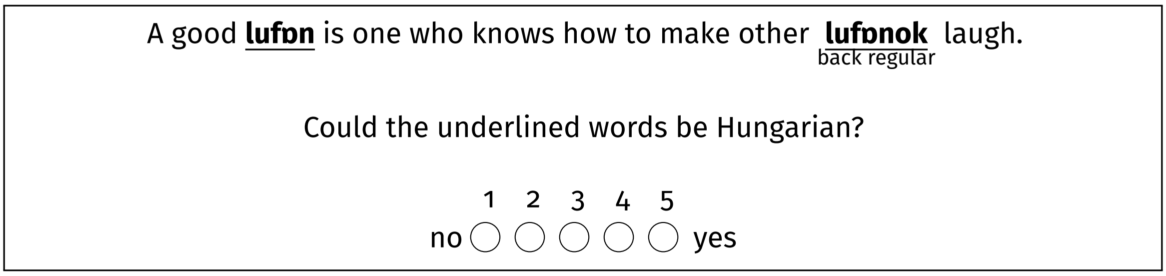

Figure 1: Trial for Hungarian stimulus norming study, with forms annotated for harmony and stem class

Three participants were discarded because their ratings disagreed substantially from the average ratings for a given word (an average deviation of 1.5 from the mean for stimuli that were rated by at least 3 participants). The ratings of the remaining 27 participants were used as inputs to a Python script (available in the supplemental materials) that generated potential subsets of stimuli to be used in the stimulus testing study. These sets were evaluated on three criteria: distance from a desired phonological distribution

[12] A further six stimuli appeared in the norming study but were removed from consideration for the main experiment because they matched inflected forms of existing words found in MTA Nyelvtudományi Intézet (2006), an online resource with Hungarian morphological paradigms.

[13] This proportion was chosen as a balance between the relative rarity of lowering stems (about 5% of the noun types in my corpus) and a need to collect sufficient data in the lowering stem condition.

similar to that of the base corpus, a high average overall rating, and a consistent average rating for stimuli falling into each phonological category. The distribution was chosen to allow for a sizable number of trials for stimuli with each phonological property while still ensuring that the aggregate distribution of all trials can be compared to the aggregate lexicon. The goal of maximizing the average rating was to make the experiment as naturalistic as possible: while I would not expect any individual low-rated stimuli to be treated differently from high-rated stimuli, participants are supposed to pretend that the stimuli are real words of Hungarian, and too great a presence of phonotactically deviant words might make the experience more obviously artificial. I examined high-ranked sets manually and selected a set with 81 stimuli to use for the main testing phase.

5.4.2 Morphological dependency testing

Participants each completed 35–50 trials,[14] which had the format shown in Figure 2. First, the stimuli were presented in written form (*lufan*, *lufanok*, etc.) in the same frame sentences as in the stimulus norming experiment. As before, stimuli were chosen randomly: each stimulus appeared in 44–62 trials (mean 53.1, standard deviation 4.2). In Figure 2, the nonce word [lufɒn] has a regular plural -*ok*, but in 8–12 trials, the stimuli were presented as lowering stems, e.g. plural [lufɒn-ɒk]. As an attention check, participants had to correctly select the plural form appearing in the first sentence. Next, a second frame sentence appeared, in which participants had to select 1SG and possessive forms. The 1SG suffix has the same regular and lowering stem variants as the plural (see Section 2.3), so the linking vowel should match that of the plural: in this case, [lufɒn-om].[15] The choices included both back and front variants; the possessive should have the same vowel harmony as the plural (in this case, [lufɒn-ɒ] or [lufɒn-jɒ]).

A good <u>**lufɒn**</u> is one who knows how to make other <u>**lufɒnok**</u> (back regular) laugh.

Please select the word's plural form: [lufɒnøk (front regular) / lufɒnɒk (back lowering) / lufɒnɛk (front lowering) / **lufɒnok** (back regular)]

That's correct! Now select the word in the appropriately inflected form according to you:

My [lufɒnɒm (back lowering) / lufɒnɛm (front lowering) / lufɒnøm (front regular) / **lufɒnom** (back regular)] couldn't sing well, however my husband's [lufɒnɛ (front -*V*) / lufɒnjɛ (front -*jV*) / **lufɒnɒ** (back -*V*) / **lufɒnjɒ** (back -*jV*)] sang brilliantly.

Figure 2: Trial for Hungarian stimulus testing study, with forms annotated for harmony and stem class and acceptable answers bolded

Trials in which participants chose a discordant 1SG or antiharmonic possessive were discarded. Of 4,305 total trials, 141 had a harmony mismatch in one or both selected forms and another 565 had a stem class mismatch. Setting aside the trials with a harmony

[14] I initially presented participants with 35 trials, as I was worried about the length of the test. Early participants completed the study more quickly than I expected, so I gradually increased the number of trials to 50 and the number of lowering stem trials accordingly.

[15] The possessive morpheme -*V*/-*jV* does not appear in first singular possessive forms, only the possessor marker (see Section 2.2).

mismatch, most stem class mismatches came in lowering stem trials: in 1,003 trials in which the stimulus was presented as a lowering stem, participants only selected a lowering stem suffix for the 1SG form in 525 (52.3%). By contrast, participants matched non-lowering stems 96.2% of the time (1,890 of 1,964 trials). These results indicate that participants often resisted treating nonce words as lowering stems. Individuals varied in their matching of stem class: 10 successfully matched all lowering stem trials, while two matched none and another 11 matched 20% or fewer (the standard deviation for individual match rate was 28.7%; I compare rate rather than number of matched trials because participants saw different numbers of trials). The results presented here exclude one participant who was particularly bad (matching 44% of trials, and only matching the harmony in 62% of trials, suggesting overall lack of attention); an analysis excluding more poor performers (14 additional participants who matched in 60% or less of all trials or 20% or less of lowering stem trials) yields similar results.

5.5 Analysis

After removing discordant trials and trials of the one excluded speaker, 3,577 trials remained. Of these, 1,179 had filler stimuli with front unrounded harmony, which do not exhibit lowering stem alternations (see Section 2.3); removing these leaves 2,398 trials with 57 stimuli.

This experiment tests the hypothesis that speakers apply patterns from their lexicon in producing novel possessive forms—in particular, the morphological correlation between stem class and possessive. The clearest confirmation of this hypothesis would be to show that the odds assigned to nonce words by the model of the lexicon in Table 6, which includes phonological factors as well as stem class (which I call *phon_morph_odds*), closely predicts the experimental results and does a better job of doing so than the odds assigned to nonce words by the lexicon model in Table 5, which includes phonological factors but not stem class (*phon_odds*). However, this model performs *worse* than the model with *phon_odds* (presented in Section 5.6.1). In Section 5.7.1, I show that this mismatch is, in fact, expected and that it is methodologically justified to separate stem class out from the phonological factors, as done in Section 5.6.2. This second model shows that taking the morphological factor of stem class into account leads to a better fit: participants' choice of possessive for nonce words is influenced by whether it is presented as a lowering stem. I still represent the effect of a word's phonology as a single aggregate factor, reflecting the assumption that speakers apply all of the phonological effects from the lexicon together. While this may paper over some differences between the effects of individual phonological factors, it is sufficient for this study, whose main purpose is to test the effect of stem class. (A model using individual phonological factors yields the same effect size for stem class.)

Parallel to the corpus study, I present two models of the experimental results, differing in the presence of stem class, whose dependent variable is the possessive suffix selected in a trial (*-V*, coded as 0, vs. *-jV*, coded as 1). Both are mixed logistic regressions with random intercepts for participant and item. The first regression includes a fixed factor of *phon_odds* (the log odds that a nonce word takes *-jV* according to the phonological model of the lexicon in Table 5; see Section 4.2.2) and a by-participant random slope for *phon_odds*, yielding the model equation *possessive* ∼ *phon_odds* + *(phon_odds* | *participant)* + *(1* | *item)*. The second regression also includes the factor of stem class. I represent this as a dummy-coded categorical variable, rather than as a *morph_odds* factor derived from a lexicon model trained on stem class alone, for ease of presentation: given that this is a factor with only two levels (regular and lowering), all observations would have one of two values for *morph_odds* anyway, meaning that the two representations behave similarly. (The equivalent model using *morph_odds* yields identical results.) This second model includes the same random effects as the previous one: adding a by-participant random intercept for stem class does not significantly improve the model ($\chi^2 = 3.05$, $p =$ .383) and is worse according to the Akaike Information Criterion (AIC), which penalizes model complexity. The final model is *possessive* ~ *phon_odds* + *stem class* + *(phon_odds | participant)* + *(1 | item)*. Models were fitted with the `glmer` function in version 1.1-34 of R's *lme4* package (Bates et al. 2015) using the `bobyqa` optimizer, which was better at finding converging models than the default `Nelder_Mead` optimizer.

5.6 Results

5.6.1 Phonology

In Table 9, we see the effects of the mixed logistic regression predicting participant responses by phonology alone, using the *phon_odds* calculated from the phonological model of the lexicon in Table 5 as described in Section 4.2.2.1.

Random effects	*variance*	*SD*
Participant		
Intercept	0.61	.78
Phon_odds	0.01	.11
Item	0.50	.70

Fixed effects	*β coef*	*SE*	*Wald z*	*p*
Intercept	**0.68**	**.14**	**4.76**	**<.0001**
Phon_odds	**0.38**	**.03**	**12.49**	**<.0001**

Table 9: Effects of mixed logistic model with predictions of the phonological model of the lexicon (Table 5) for experimental use of possessive *-jV*, with significant effects bolded

This model shows an overall bias towards *-jV* (since the intercept is positive): there were 1,031 responses of *-V* and 1,367 of *-jV*. The results also show a correspondence between predicted rates and actual rates: the β coefficient for *phon_odds* is positive. However, this coefficient is smaller than 1. This means that the overall range of likelihood predicted by the experimental model is narrower than that predicted by the model of the lexicon: a change of 1 in a word's predicted log odds of taking *-jV* according to its phonology corresponds to a change of only .38 in its predicted log odds in the experiment.

The random intercept for item shows that different words were given fairly divergent rates of *-jV* even once phonology (*phon_odds*) is taken into account. The random intercept for participant shows that different participants had different baseline rates of *-jV*, but the by-participant random slope for *phon_odds* has a low variance, suggesting that participants treated a nonce word's phonology in similar ways. Including this random slope significantly improves the model ($\chi^2 = 15.60$, $p < .001$).

We see the difference between the lexical and experimental models in Table 10, which shows the two words predicted to be most ([olu:nt]) and least ([jøs]) likely to take *-jV* and the word with a predicted rate closest to 50% ([ʃok:ol]). These predictions follow from the words' phonology: words ending in sibilants like [jøs] take *-V* nearly categorically in the lexicon, and other properties of this word, like front harmony, push its prediction further towards *-V*. By contrast, [olu:nt] has numerous properties that strongly prefer *-jV*, especially its final cluster ending in an alveolar stop.

The two extremes, [jøs] and [olu:nt], are predicted by Table 5 to be essentially categorical in the lexicon, but showed mixed responses in the experiment. Correspondingly, the model trained on the experimental results (Table 9) predicts that one variant should be dominant but not categorical. Table 10 also shows the effect of the random intercept for stimulus, which is to account for variance that the fixed effects alone cannot by bringing observations closer to the line of best fit (see Figure 3). In the case of [jøʃ], the fixed coefficients alone substantially underestimate the likelihood of *-jV*: 4.9% instead of the observed 17.4%. This word thus has a substantial positive random intercept, which adjusts the predicted rate up to 9.3%—closer to the observed rate. For the other two words, the model with fixed intercepts for item already does a very good job at matching the experimental rate: for example, [ʃok:ol] is predicted to have a rate of 67.0% and has an actual rate of 67.5%. For this word, the random intercept overcompensates slightly and brings the predicted rate up to 69.5%; for [olu:nt], the random intercept yields a slight improvement, reaching 96.3%, close to the true rate of 94.7%.

nonce word	*predicted likelihood of -jV in lexicon model*	*experimental rate of -jV*		*predicted likelihood of -jV in experimental model*	
				fixed intercept	*random intercept*
jøʃ	0.006%	17.4%	(8/46)	4.9%	9.3%
ʃok:ol	52.095%	67.5%	(27/40)	67.0%	69.5%
olu:nt	99.934%	94.7%	(36/38)	96.9%	96.3%

Table 10: Predicted likelihood of *-jV* for nonce words according to models trained on lexicon (Table 5) and experimental results (Table 9), including the adjustment of the random intercept for item

Figure 3 shows the relationship between the predicted likelihood (according to the *phon_odds*) of each nonce word taking *-jV* and its experimental rate of *-jV*. Both axes are shown in terms of log odds (that is, coefficients), making the relationship linear. A rate of 0 corresponds to a log odds of negative infinity, so the nonce word *fátyúsz* [fa:cu:s], which speakers assigned *-V* in every trial, should be at negative infinity. It is included at the bottom edge of the graph in Figure 3. The graph shows each nonce word twice: in black, the word's position on the x-axis assumes a fixed intercept, so each word's position is solely a function of its *phon_odds*. The lighter gray includes the adjustment of the random intercept for word, bringing them closer towards the line of best fit: stimuli above the line have a random intercept shifting them to the right, while those below the line move left with the random intercept. Figure 4 shows the same data plotted on scales of raw probability. The graphs also include a line corresponding to the fit of the model in Table 9.

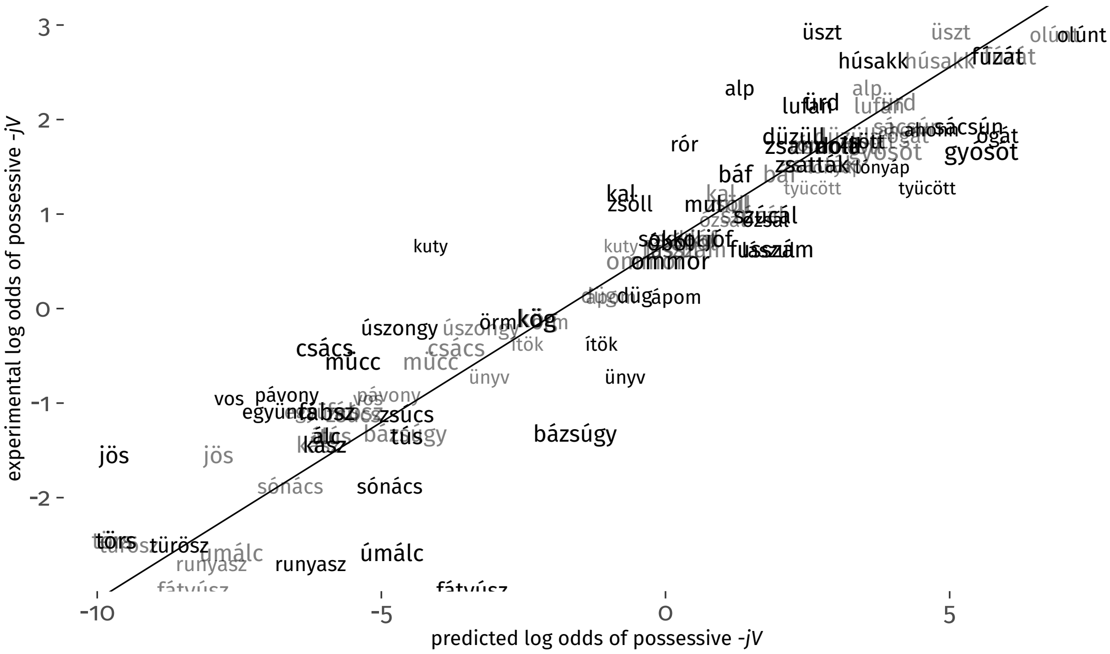

Figure 3: The relationship between predicted and experimental log odds of possessive *-jV* for individual nonce words with (gray) and without (black) the random intercept, sized according to number of trials, with a line showing the fit of the experimental model in Table 9

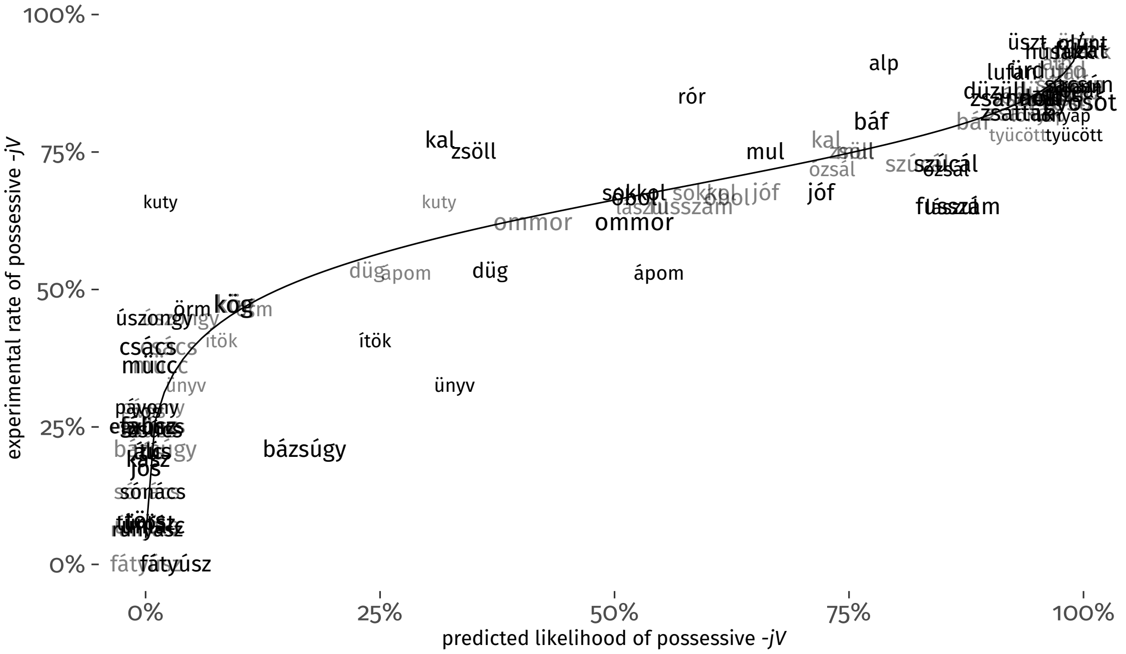

Figure 4: The relationship between predicted likelihood and experimental rate of possessive *-jV* for individual nonce words with (gray) and without (black) the random intercept, sized according to number of trials, with a line showing the fit of the experimental model in Table 9

Figure 3 shows that the log odds of *-jV* predicted from the lexicon map well onto the experimental log odds: the relationship is a tight linear fit. Figure 4 shows that the experimental results are less extreme than the predicted likelihood, especially on the low end: nouns ending in palatals and sibilants, which nearly categorically take *-V* in the lexicon and thus had a near-zero predicted likelihood of *-jV*, were assigned *-jV* in the experiment up to nearly 50% of the time, or even higher in the case of *kuty* [kuc].

5.6.2 Phonology and stem class

Table 11 shows the effects of the regression predicting participant responses by both phonology and stem class (that is, the plural shown on the nonce word in a given trial).

Random effect	*variance*	*SD*		
Participant				
Intercept	0.61	.78		
Phon_odds	0.01	.12		
Item	0.51	.72		
Fixed effects	*β coef*	*SE*	*Wald z*	*p*
Intercept	**0.77**	**.15**	**5.27**	**<.0001**
Phon_odds	**0.38**	**.03**	**12.43**	**<.0001**
Stem class (default: non-lowering)				
lowering	**−0.43**	**.13**	**−3.21**	**.0013**

Table 11: Effects of mixed logistic model with predictions of the phonological model of the lexicon (Table 5) and stem class for experimental use of possessive *-jV*, with significant effects bolded

Stimuli (e.g. [huːʃɒkː]) presented as regular stems (with plural [huːʃɒkː-ok]) were assigned *-jV* ([huːʃɒkː-jɒ]) 58.1% of the time (1,090 out of 1,876 trials), while participants assigned *-jV* to stimuli slightly less often when they were presented as lowering stems (with plural [huːʃɒkː-ɒk]), 53.1% of the time (277 of 522 trials). This may seem like a rather small difference, but it is significant in the model in Table 11. This model performs significantly better than the model without the morphological factor shown in Table 9 ($\chi^2 = 9.96$, $p = .002$). As before, the random slope for *phon_odds* has a small variance of .01.

Figure 5 and Figure 6 show the same data as Figure 3 and Figure 4, but each nonce word is now split between trials when it was presented as a regular stem (in black), and a lowering stem (in gray). (Both versions show the fixed intercept only.) The lowering stem words are always smaller than the regular words because each word appeared in fewer lowering stem trials. A line connects the two conditions for each word, going leftward from the regular word to the lowering stem word because the model predicts a lower likelihood of *-jV* for lowering stems. Probabilities of 0 and 1 correspond to log odds of (negative) infinity, so words with categorical behavior in one condition are shown at the bottom and top edges of Figure 5. The graphs show lines corresponding to the fit of the model in Table 11 for regular (black) and lowering stem (gray) conditions.

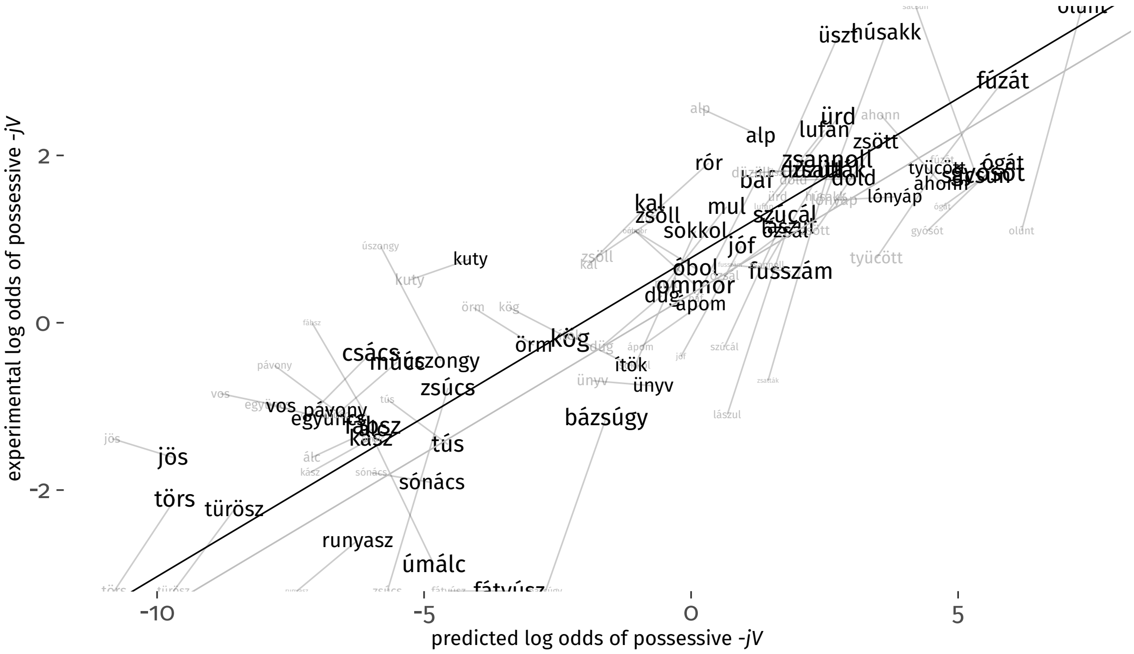

Figure 5: The relationship between predicted and experimental log odds of possessive *-jV* for individual nonce words presented with regular (black) and lowering (gray) plurals, sized according to number of trials, with a line showing the fit of the experimental model in Table 9

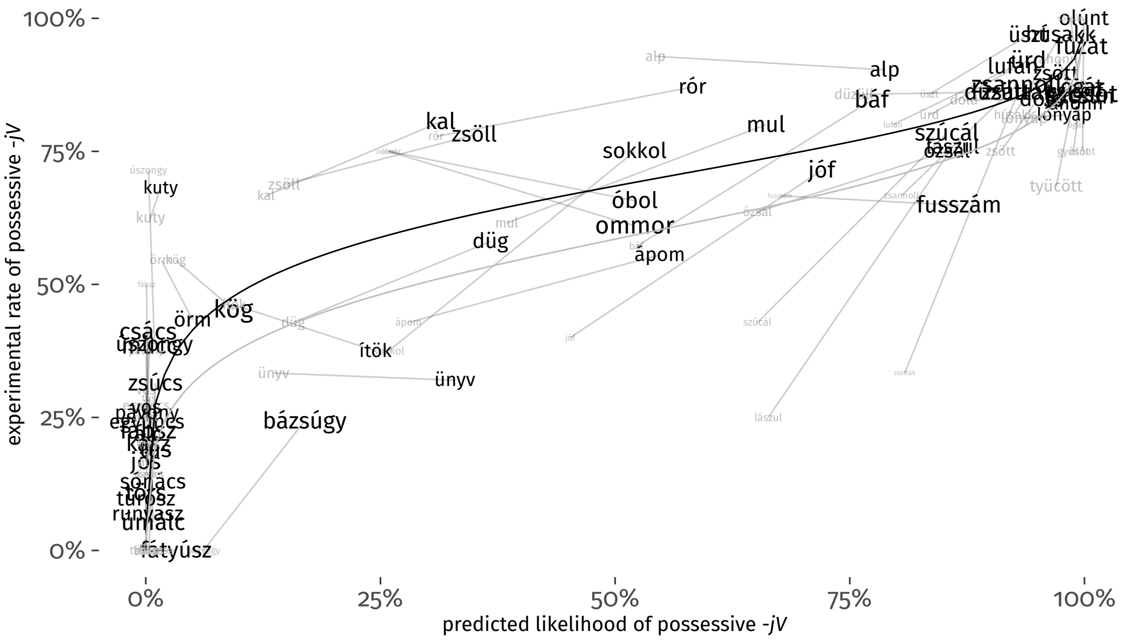

Figure 6: The relationship between predicted likelihood and experimental rate of possessive *-jV* for individual nonce words presented with regular (black) and lowering (gray) plurals, sized according to number of trials, with a line showing the fit of the experimental model in Table 9

Most of the lines in these figures slope downward and to the left, indicating that most nonce words had a lower experimental rate of *-jV* when presented as a lowering stem. This is especially true in the top right section of the graph—that is, nonce words with a higher rate of *-jV*. On the other hand, nouns with a lower expected rate of *-jV* are more likely to have a *higher* rate of *-jV* when presented as a lowering stem. However, as mentioned above, these words generally behaved unexpectedly: they almost always take *-V* in the lexicon but showed moderate rates of *-jV* in the experiment.

5.7 Discussion

5.7.1 Model construction

The operating assumption in my models of the experimental results is that speakers are applying the distribution of the lexicon to nonce words. Accordingly, as described in Section 5.5, the complex model of the lexicon (with both phonological factors and stem class as predictors of whether a noun takes *-jV*) should be the best predictor of the experimental results (whether participants assign *-jV* to nonce words). However, the full lexicon model in Table 6 (assigning *phon_morph_odds* to nonce words) does *worse* at predicting experimental results than the lexicon model in Table 5, which only includes phonological factors and not stem class (assigning *phon_odds*).

Table 11 shows that stem class does have an effect. The model with *phon_morph_odds* fails because the experimental effect of stem class is *weaker* than the phonological effects relative to the lexicon. We can see this from the fixed effects in Table 11. The effect size of *phon_odds* in this model is .38, meaning that the effect of the phonological factors was

38% as strong experimentally as in the lexicon (see Section 5.6.1). The effect size of stem class in the model of the experimental results is −.43, while in the model of the lexicon (Table 6), it is −3.71. That is, the effect of stem class was $\frac{-.43}{-3.71} = .116 = 11.6\%$ as strong experimentally as in the lexicon. Accordingly, a model with *phon_morph_odds* predicting experimental results must split the difference between accurately capturing the relative strength of the phonological effects (thus overweighting stem class) and accurately capturing the relative strength of stem class (thus underweighting the phonological factors). Including *phon_odds* and stem class as separate factors gives the model a degree of freedom enabling it to capture the difference in strength between the two types of factors. In the next section, I propose interpretations for these strength differences, arguing that this degree of freedom is methodologically justified.

5.7.2 How did speakers match the frequency of the lexicon?

In both phonology and stem class, participants applied patterns from the lexicon to the possessive of nonce words, but in less extreme form. The difference is particularly stark for nouns ending in sibilants and palatals, which nearly categorically take *-V* in the lexicon but were assigned *-jV* in the experiment in a substantial minority of trials. One likely source of this discrepancy is a task effect: the forced choice task makes alternatives more salient, leading participants to select "unlikely" possessives more often than they would in normal production.[16] There is also likely some noise. The rate of discarded trials where participants chose suffixes with the wrong vowel harmony for the stem (141 out of 4,305; see Section 5.4.2), while low, suggests occasional inattention or random guessing, and a similar number of trials presumably include a randomly chosen answer that was not filtered out.

One relatively minor factor is that of individual participant differences. Although participants differed in their propensity to assign *-jV* (the random intercept for participant has substantial variance), this difference was only weakly mediated by phonology (the by-participant random slope for *phon_odds* has a low variance).

The low rate of harmony mismatches can be contrasted with the much higher rate of stem class mismatches. Despite my study design, which forced participants to pay attention to the presented plural suffix, they were not very accepting of nonce lowering stems, in particular: nearly half (478 of 1,003) of all trials in which the stimulus was presented with a lowering stem plural (e.g. *-ɒk*) were discarded because the participant assigned it a non-lowering 1SG suffix (e.g. *-om*), although individuals did so at different rates (see Section 5.4.2). This means that lack of consideration for (or rejection of) the experimental manipulation plays a larger role specifically with the factor of stem class, which would make its effect on the experimental results weaker. Thus, looking at quantitative evidence from discordant trials, we would expect the effect of stem class to be weaker than that of the phonological factors. This is what we find.

There is another reason to treat stem class separately from the phonological factors and to expect its effect to be attenuated. Participants had to choose between suffixed forms, meaning that they were directly confronted with the phonological form of each stimulus when selecting a possessive. In contrast, information about stem class is only available from other forms, so speakers are not explicitly reminded of it when selecting the possessive. These are two different types of information, and the former should have a stronger effect, because it is more immediate.

[16] I thank Volya Kapatsinski for raising this point.

The treatment of stem class as separate from phonological factors pursued in this study is thus the most accurate way to represent this factor. This discussion supports the validity of the experimental results confirming my hypotheses: subjects applied gradient patterns from the lexicon, counter to the claim by Rácz & Rebrus (2012), discussed in Section 2.2, that novel words categorically take one possessive suffix (*-V* for nouns ending in palatals and sibilants, *-jV* otherwise). In particular, subjects learned the generalization that lowering stems (which take plural suffixes like *-ɒk*) are more likely to take *-V* in the possessive. The relatively weak experimental correlation is expected, since there is good reason to believe that the experiment was only able to detect this correlation quite weakly—that is, the difference in strength is a result of the experiment rather than a fact about Hungarian speakers' grammars.

5.7.3 Are speakers generalizing over roots or stems?

In Section 4.2.3, I compared my corpus of monomorphemic words, which captures the assumption that lexical items are stored as roots, with a stem-based corpus that counts derived forms and compounds separately. I conclude the nonce word study discussion by briefly discussing which corpus is a better fit for the experimental results. I suggest that participants are roughly matching the frequency of the lexicon counting roots, not stems.

As discussed in Section 4.2.3, the corpora are quite similar in the relative distribution of *-V* and *-jV*, and indeed, *phon_odds* produced by models trained on the two are quite closely correlated ($R^2 = .86$), meaning that stimuli predicted to be more likely to take *-jV* in the model of monomorphemic nouns are also predicted to be more likely to take *-jV* in the model of all nouns. The corpora differed on three points. Most saliently, the stem-based corpus had a much higher overall rate of *-V*. Second, the root-based corpus had a higher rate of *-jV* among alveolar stops than labial stops, while the stem-based corpus showed the opposite pattern. Finally, the stem-based corpus had a much higher rate of *-jV* among words ending in complex codas (geminates and clusters) than those ending in singletons, while the difference in the root-based corpus was much smaller. In Table 12, I compare the proportions of *-jV* in the root- and stem-based corpora from Table 8 with the experimental results. The manner and coda complexity counts are inconclusive: the experimental proportions lie in between those of the two corpora. The baseline rate of *-jV* in the experiment, however, is much more in line with the predictions of the root-based corpus: participants assigned *-jV* 75.2% of the time to stimuli not ending in palatals or sibilants, much closer to the rate of 72.2% in the corpus of monomorphemic nouns than the 28.3% rate across all nouns.[17]

[17] The rates of *-jV* are quite similar if we include trials with the filler stimuli.

	experiment (responses)			lexicon (types) monomorphemic	all
	-V	-jV	% -jV	% -jV	% -jV
labial plosive	11	72	86.7%	76.8%	56.9%
alveolar plosive	47	339	87.8%	96.5%	44.3%
velar plosive	81	137	62.8%	64.0%	19.3%
singleton	270	691	71.9%	71.5%	23.1%
geminate/cluster	118	488	80.5%	74.0%	53.3%
total	388	1179	75.2%	72.2%	28.3%

Table 12: Experimental frequency of *-jV* responses compared with type frequency of *-jV* in Papp (1969) for nouns ending in consonants that are not palatal or sibilant (see Table 8)

The baseline rate of *-jV* suggests that participants, on the whole, were mirroring the frequencies of the root-based corpus. More sophisticated statistical analysis is less conclusive. On the one hand, the *phon_odds* produced by the stem-based corpus are *better* predictors of the experimental results than the root-based *phon_odds*: the equivalent of Table 11 using the stem-based corpus yields a better model ($\chi^2 = 10.27$) of the experimental results than that in Table 11. However, there are two reasons to question this as support for the stem-based model. First, these models have a free intercept parameter that sets the baseline, so they do not penalize the stem-based *phon_odds* for the large difference in baseline rate of *-jV* shown in Table 12. Second, the better performance of the stem-based *phon_odds* seems to be an artifact of the surprising behavior of nouns ending in palatals and sibilants: when these are removed, the root-based *phon_odds* yield a better-fitting model ($\chi^2 = 6.20$). The issue requires further study, but the comparisons in this section tentatively suggest that Hungarian speakers are counting over roots and affixes, not complex stems. This root-based storage, used in Distributed Morphology (e.g. Halle & Marantz 1993; Embick & Marantz 2008), fits with the theory of morphological dependencies I present in Section 6.

6 Sublexicon models and morphological dependencies

To account for the results in Section 5.6, we need a theory that can apply patterns of allomorphy from the lexicon productively to new words. In particular, this theory must be able to learn *morphological* patterns like the correlation between lowering stems and possessive *-V*. In Section 1.1, I argued that using diacritic features on lexical items provides an easy symbolic representation for lexically specific morphological behavior that can be used in generalizations. Moreover, as discussed in Section 5.7, these generalizations should be made over roots, not stems. In this section, I sketch out such a theory, based on the sublexicon model of phonological analogy (Gouskova et al. 2015). I extend this basic model with a novel set of morphological constraints that allow for generalizations over cases of lexically specific allomorphy. As before, I assume Distributed Morphology (Halle & Marantz 1993), which uses root-based storage and diacritic features.

6.1 The basic sublexicon model

The sublexicon model (Allen & Becker 2015; Gouskova et al. 2015; Becker & Gouskova 2016) encodes phonological generalizations in lexically specific variation. This allows learners to pick up on the partial phonological predictability determining a given lexical item's choice of allomorph. As such, it follows in the path of previous models, like the Minimal Generalization Learner (Albright & Hayes 2003), that use phonological analogy to determine the set of lexical items to which a given morphophonological rule applies (see Guzmán Naranjo (2019) for an overview).

In the sublexicon model, the learner divides the lexicon into *sublexicons* that pattern together. These sublexicons correspond with the morphological features described in Section 3: [+lower] for lowering stems, [−lower] for other consonant-final nouns, [−j] for nouns that take possessive *-V*, and [+j] for nouns that take *-jV*. These groupings are shown below:

(7) *Lexical entries for Hungarian nouns*

a. [+lower]: /vaːlː$_{[\textbf{+lower},-j]}$/ 'shoulder', /hold$_{[\textbf{+lower},+j]}$/ 'moon', …
b. [−lower]: /dɒl$_{[\textbf{−lower},-j]}$/ 'song', /tʃont$_{[\textbf{−lower},+j]}$/ 'bone', …
c. [+j]: /tʃont$_{[-lower,\ \textbf{+j}]}$/ 'bone', /hold$_{[+lower,\textbf{+j}]}$/ 'moon', …
d. [−j]: /dɒl$_{[-lower,\ \textbf{−j}]}$/ 'song', /vaːlː$_{[+lower,\textbf{−j}]}$/ 'shoulder', …

6.2 Sublexical phonotactic grammars

Hayes & Wilson (2008) present a model of phonotactic learning in which a learner captures generalizations over a language's surface forms through a constraint-based phonotactic grammar. In their proposal, the learner keeps track of sounds or sequences of sounds (defined in terms of features) that are rare or absent in the lexicon and proposes constraints against them, weighting them in accordance with the strength of the generalization. For example, geminate consonants in Hungarian generally do not appear in clusters, especially within a morpheme, so the phonotactic learner trained on my corpus of monomorphemic nouns generates strong phonotactic constraints penalizing geminate consonants adjacent to other consonants: *[−syllabic,+long][−syllabic] and *[−syllabic][−syllabic,+long].[18] This is how the speaker knows that clusters with geminates are unlikely in Hungarian.

The sublexicon model extends the notion of phonotactic learning to capture generalizations over subsets of the lexicon that pattern together—that is, sublexicons. The learner induces a phonotactic grammar for each sublexicon, capturing patterns specific to that sublexicon. As discussed in Section 4.2.3, Hungarian nouns ending in sibilants and palatals categorically take possessive *-V*, while nouns ending in vowels always take *-jV*. The sublexical grammar for the [+j] sublexicon should include heavily weighted constraints penalizing final sibilants and palatals, and the [−j] sublexicon should penalize word-final vowels.

These sublexical grammars are then reflected in speakers' behavior. When a speaker wishes to form the possessive of a novel word, they evaluate the stem against each sublexicon's grammar, where each sublexical grammar yields a score for that word. The better a word fares on the [+j] sublexicon relative to the [−j] sublexicon, the more likely it is to be placed into this sublexicon, and thus take *-jV*.

[18] Hayes & Wilson (2008) released an implementation of their learning model, the UCLA Phonotactic Learner. In practice, it does capture many strong phonotactic tendencies, but also learns many constraints that strike linguists as phonologically unnatural and do not correspond to the phonotactic knowledge of real speakers (Hayes & White 2013). When applied to the Hungarian data, the Phonotactic Learner also failed to learn many moderate tendencies that speakers displayed sensitivity to. In this section, I focus on the conceptual framework of sublexical phonotactic grammars rather than any particular model of how they are learned.

In Figure 7 and Figure 8, we see two nonce words from the experiment in Section 5, *runyasz* [ruɲɒs] and *fúzát* [fu:za:t], tested on toy sublexical grammars with the constraints described above. Here, [ruɲɒs] is penalized by *[+strident]#, penalizing word-final sibilants, in the [+j] sublexicon, but not by the constraint against word-final vowels in the [–j] sublexicon; [fu:za:t] accrues no penalities. I assume that all three constraints have a weight of 5.

constraint *weight*	*[+strident]# 5	*[+palatal]# 5	total
ruɲɒs	−5	0	−5
fu:za:t	0	0	0

Figure 7: Evaluation of nonce words *runyasz* and *fúzát* on the [+j] sublexical grammar

constraint *weight*	*[+syllabic]# 5	total
ruɲɒs	0	0
fu:za:t	0	0

Figure 8: Evaluation of nonce words *runyasz* and *fúzát* on the [–j] sublexical grammar

Since [ruɲɒs] has a better score on the [–j] sublexical grammar than the [+j] sublexical grammar, it is much more likely to be placed into the former and form its possessive with *-V*. Specifically, this is a maximum entropy model (Hayes & Wilson 2008): a word's likelihood of being placed into a sublexicon is proportional to its (negative) score raised to the power of e. Here, the probability of [ruɲɒs] being assigned to the [+j] sublexicon is $\frac{e^{-5}}{e^{0}+e^{-5}} = .0067 = .67\%$. On the other hand, since [fu:za:t] has the same score on both sublexicons, it has a 50% chance of being assigned to each.

The sublexicon model is designed to capture generalizations over the phonological shape of each sublexicon's members. In Section 5, I showed that speakers also observe a morphological generalization: lowering stems are more likely to have possessive *-V*. In the feature-based analysis of Section 3, this means that [+lower] and [–j] are likely to cooccur on lexical items, as stated in (3). In the next section, I extend the sublexicon model to accommodate these relations.

6.3 A sublexicon model with morphology

In my proposal, each sublexicon's grammar has constraints penalizing diacritic features alongside those penalizing phonological features. For example, every member of the [+j] sublexicon has [+j] (by definition), but very few have [+lower], since lowering stems rarely take *-jV*. Since [+lower] is underrepresented in the [+j] sublexicon, the [+j] sublexical grammar should contain a heavily weighted constraint *[+lower] penalizing nouns with both [+lower] and [+j]. The [–j] sublexicon, comprising words that take *-V*, will also have a *[+lower] constraint, but it will not be as strong, since lowering stems are better represented among *-V* words (though still uncommon).

Figure 9 and Figure 10 show the evaluation of our two nonce words on the toy grammars, now containing *[+lower]. This constraint has a heavier weight in the [+j] grammar than in the [–j] grammar (2 and 1, respectively). Here, the speaker knows that the plurals of these words are [ruɲɒs-ɒk] and [fu:za:t-ɒk], so she has marked both with [+lower].

constraint weight	*[+strident]# 5	*[+palatal]# 5	*[+lower] 2	total
$\text{ruɲɒs}_{[+\text{lower}]}$	−5	0	−2	−7
$\text{fu:za:t}_{[+\text{lower}]}$	0	0	−2	−2

Figure 9: Evaluation of nonce lowering stems *runyasz* and *fúzát* on the [+j] sublexical grammar with *[+lower]

constraint weight	*[+syllabic]# 5	*[+lower] 1	total
$\text{ruɲɒs}_{[+\text{lower}]}$	0	−1	−1
$\text{fu:za:t}_{[+\text{lower}]}$	0	−1	−1

Figure 10: Evaluation of nonce lowering stems *runyasz* and *fúzát* on the [−j] sublexical grammar with *[+lower]

The *[+lower] constraint brings the likelihood of [+j] (and possessive *-jV*) being assigned to /$\text{ruɲɒs}_{[+\text{lower}]}$/ slightly further, from .67% to .25%. For /$\text{fu:za:t}_{[+\text{lower}]}$/, the effect is more visible: the likelihood of [+j] goes from 50% to $\frac{e^{-2}}{e^{-1}+e^{-2}} = .269 = 26.9\%$. This shows how the sublexicon model can accommodate the effects found in the nonce word experiment, both phonological and morphological: nonce words ending in sibilants are less likely to be assigned *-jV* (that is, be placed in the [+j] sublexicon), as are words shown as lowering stems. These effects can all be assessed in a single calculation, correctly allowing them to compound or cancel out.

The sublexicon model presented in this section has two major theoretical benefits as a tool for capturing morphological dependencies. First of all, it does not require any additional theoretical mechanisms beyond those already proposed for well-established experimental effects (phonotactic constraints penalizing underrepresented feature combinations). Second, it relies on diacritic features, a common theoretical construct (though not universally accepted; see Section 3.3), making the sublexicon model compatible with theories of morphology that commonly use diacritic features—in my case, Distributed Morphology.

7 Conclusion

The study presented in this paper is intended as a beachhead for experimental study of morphological dependencies. The results are a proof of concept for a new experimental paradigm: a nonce word study in which stimuli are presented with varying morphological behavior. Hungarian speakers were shown to be sensitive to this manipulation: they assigned the possessive suffix *-V* more frequently to stimuli presented as lowering stems, with plural *-ɒk*, than to stimuli presented with the more common plural suffix *-ok*. This finding supports the claims of Ackerman et al. (2009), Ackerman & Malouf (2013), and others: speakers learn lexical correlations between a word's morphological patterns and use these correlations to infer unknown forms of words.

This paper is also a response to the theoretical concerns of Ackerman & Malouf (2013), who note that piece-based theories of morphology like Distributed Morphology typically ignore "the patterns of organization among related words" as "the epiphenomenal result of representations and operations designed to produce individual words". It is true that, with very limited exceptions, in particular Halle & Marantz (2008), linguists working in such theories have addressed correspondences between related forms only in the case of syncretism, i.e. *identity*. Thus, the goal of many formal analyses is to explain why two inflected forms differing in morphosyntactic featural content have the same realization. As I have shown in this paper, however, theories like Distributed Morphology are compatible with learning of morphological dependencies—even complementary to it. In my proposal, morphological knowledge is split between a piece-based generative grammar and vocabulary handling derivations of familiar words (including the rules of realization in Section 3 and the lexical entries in (7)) and a pattern-learning mechanism used productively to fill in the gaps in stored knowledge (the sublexical grammars of Section 6). On the one hand, this approach enables us to account for morphological dependencies without discarding the insights of generative morphological accounts. On the other, it sharpens our study of these dependencies by forcing us to be explicit about what, exactly, is being generalized over (lexical features indexing rules of realization and readjustment rules). This study thus offers not just novel empirical results, but also a proposed synthesis of two approaches to morphology that have often been set against one another.

Abbreviations

ALL = allative, DAT = dative, INESS = inessive, INS = instrumental, NOM = nominative, PL = plural, POSS = possessive, SG = singular, 1SG = first person singular possessor marker, etc.

Data availability

All data, analysis code, experimental materials, calculations, and alternate models referenced in this study can be found in the supplemental materials at: https://osf.io/mdra8/?view_only=3794e1a584f34fc8a26a7f7b38e8f251.

Ethics and consent

The studies conducted in this paper were deemed exempt by the New York University Institutional Review Board, IRB-FY2022-5933.

Competing interests

The author has no competing interests to declare.

References

Ackerman, Farrell & Blevins, James P. & Malouf, Robert. 2009. Parts and wholes: Implicative patterns in inflectional paradigms. In Blevins, James P. & Blevins, Juliette (eds.), *Analogy in grammar: Form and acquisition*, chap. 3, 54–82. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199547548.003.0003

Ackerman, Farrell & Malouf, Robert. 2013. Morphological organization: The low conditional entropy conjecture. *Language* 89(3). 429–464. https://doi.org/10.1353/lan.2013.0054

Albright, Adam & Hayes, Bruce. 2003. Rules vs. analogy in English past tenses: a computational/experimental study. *Cognition* 90(2). 119–161. https://doi.org/10.1016/s0010-0277(03)00146-x

Allen, Blake & Becker, Michael. 2015. Learning alternations from surface forms with sublexical phonology. Lingbuzz/002503.

Bartos, Huba. 1999. *Morfoszintaxis és interpretáció. a magyar inflexiós jelenségek szintaktikai háttere*. Budapest: Eötvös Loránd University dissertation.

Bates, Douglas & Mächler, Martin & Bolker, Ben & Walker, Steve. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48.

Becker, Michael & Gouskova, Maria. 2016. Source-oriented generalizations as grammar inference in Russian vowel deletion. *Linguistic Inquiry* 47(3). 391–425. https://doi.org/10.1162/ling_a_00217

Becker, Michael & Ketrez, Nihan & Nevins, Andrew. 2011. The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish laryngeal alternations. *Language* 87(1). 84–125. https://doi.org/10.1353/lan.2011.0016

Bermúdez-Otero, Ricardo. 2012. The architecture of grammar and the division of labor in exponence. In Trommer, Jochen (ed.), *The morphology and phonology of exponence* (Oxford Studies in Theoretical Linguistics 41), chap. 2, 8–83. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199573721.003.0002

Bermúdez-Otero, Ricardo. 2013. The Spanish lexicon stores stems with theme vowels, not roots with inflectional class features. *Probus: International Journal of Romance Linguistics* 25(1). 3–103. https://doi.org/10.1515/probus-2013-0009

Bobaljik, Jonathan David. 2000. The ins and outs of contextual allomorphy. In Grohmann, Kleanthes & Struijke, Caro (eds.), *University of Maryland working papers in linguistics*. 35–71. College Park: University of Maryland.

Bonami, Olivier & Beniamine, Sacha. 2016. Joint predictiveness in inflectional paradigms. *Word Structure* 9(2). 156–182. https://doi.org/doi.org/10.3366/word.2016.0092

Bybee, Joan. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10(5). 425–455. https://doi.org/10.1080/01690969508407111

Bybee, Joan. 2001. *Phonology and language use* (Cambridge Studies in Linguistics 94). Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511612886

Caha, Pavel. 2021. Modeling declensions without declension features: The case of Russian. *Acta Linguistica Academica* 68(4). 385–425. https://doi.org/10.1556/2062.2021.00433

Chomsky, Noam & Halle, Morris. 1968. *The sound pattern of English* (Studies in Language). New York, Evanston, and London: Harper & Row.

Dékány, Éva. 2018. The position of case markers relative to possessive agreement: Variation within Hungarian. *Natural Language & Linguistic Theory* 36. 365–400. https://doi.org/10.1007/s11049-017-9379-7

É. Kiss, Katalin. 2002. *The syntax of Hungarian* (Cambridge Syntax Guides). Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511755088

Embick, David & Halle, Morris. 2005. On the status of stems in morphological theory. In Geerts, Twan & van Ginneken, Ivo & Jacobs, Heike (eds.), *Romance languages and linguistic theory 2003: Selected papers from 'Going Romance' 2003, Nijmegen, 20–22*

November (Current Issues in Linguistic Theory 270), 37–62. Amsterdam, Philadelphia: John Benjamins. https://doi.org/10.1075/cilt.270.03emb

Embick, David & Marantz, Alec. 2008. Architecture and blocking. *Linguistic Inquiry* 39(1). 1–53. https://doi.org/10.1162/ling.2008.39.1.1

Ernestus, Mirjam & Baayen, R. Harald. 2003. Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language* 79(1). 5–38. https://doi.org/10.1353/lan.2003.0076

Finkel, Raphael & Stump, Gregory. 2007. Principal parts and morphological typology. *Morphology* 17. 39–75. https://doi.org/10.1007/s11525-007-9115-9

Gouskova, Maria. 2012. Unexceptional segments. *Natural Language and Linguistic Theory* 30(1). 79–133. https://doi.org/10.1007/s11049-011-9142-4

Gouskova, Maria & Bobaljik, Jonathan David. 2022. The lexical core of a complex functional affix: Russian baby diminutive *-onok*. *Natural Language & Linguistic Theory* 40. 1075–1115. https://doi.org/10.1007/s11049-021-09530-1

Gouskova, Maria & Newlin-Łukowicz, Luiza & Kasyanenko, Sofya. 2015. Selectional restrictions as phonotactics over sublexicons. *Lingua* 167. 41–81. https://doi.org/10.1016/j.lingua.2015.08.014

Guzmán Naranjo, Matías. 2019. *Analogical classification in formal grammar* (Empirically Oriented Theoretical Morphology and Syntax 5). Berlin: Language Science Press. https://doi.org/10.5281/zenodo.3191825

Halle, Morris & Marantz, Alec. 1993. Distributed morphology and the pieces of inflection. In Hale, Ken & Keyser, Samuel Jay (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger* (Current Studies in Linguistics 24), chap. 3, 111–176. Cambridge: MIT Press.

Halle, Morris & Marantz, Alec. 2008. Clarifying "Blur": Paradigms, defaults, and inflectional classes. In Bachrach, Asaf & Nevins, Andrew (eds.), *Inflectional identity* (Oxford Studies in Theoretical Linguistics 18), chap. 3, 55–72. Oxford: Oxford University Press.

Harrell Jr., Frank E. 2020. *rms: Regression modeling strategies*. https://CRAN.R-project.org/package=rms. R package version 6.1-0.

Haugen, Jason D. 2016. Readjustment: Rejected? In Siddiqi, Daniel & Harley, Heidi (eds.), *Morphological metatheory* (Linguistic Aktuell/Linguistics Today 229), 303–340. Amsterdam, Philadelphia: John Benjamins. https://doi.org/10.1075/la.229.11hau

Hayes, Bruce & Londe, Zsuzsa Cziráky. 2006. Stochastic phonological knowledge: the case of Hungarian vowel harmony. *Phonology* 23. 59–104. https://doi.org/10.1017/s0952675706000765

Hayes, Bruce & White, James. 2013. Phonological naturalness and phonotactic learning. *Linguistic Inquiry* 44(1). 45–75. https://doi.org/10.1162/ling_a_00119

Hayes, Bruce & Wilson, Colin. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39(3). 379–440. https://doi.org/10.1162/ling.2008.39.3.379

Hayes, Bruce & Zuraw, Kie & Siptár, Péter & Londe, Zsuzsa. 2009. Natural and unnatural constraints in Hungarian vowel harmony. *Language* 85(4). 822–863. https://doi.org/10.1353/lan.0.0169

Hudson Kam, Carla L. & Newport, Elissa L. 2005. Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development* 1(2). 151–195. https://doi.org/10.1207/s15473341lld0102_3

James, Gareth & Witten, Daniela & Hastie, Trevor & Tibshirani, Robert. 2013. *An introduction to statistical learning with applications in R* (Springer Texts in Statistics 103). New York: Springer. https://doi.org/10.1007/978-1-4614-7138-7

Kiparsky, Paul. 1982. *Explanation in phonology* chap. 6, 119–164 (Publications in Language Science) Dordrecht, Cinnaminson: Foris Publications. Written in 1968 and originally published in 1973.

Kramer, Ruth. 2015. *The morphosyntax of gender* (Oxford Studies in Theoretical Linguistics). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199679935.001.0001

Lightner, Theodore. 1965. *Segmental phonology of Modern Standard Russian*. Cambridge: Massachusetts Institute of Technology dissertation.

Lüdecke, Daniel & Ben-Shachar, Mattan S. & Patil, Indrajeet & Waggoner, Philip & Makowski, Dominique. 2021. performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software* 6(60). 3139. https://doi.org/10.21105/joss.03139

MTA Nyelvtudományi Intézet. 2006. Alaktani táblázatok. http://corpus.nytud.hu/cgi-bin/e-szokincs/alaktan.

Müller, Gereon. 2004. On decomposing inflection class features: Syncretism in Russian noun inflection. In Müller, Gereon & Gunkel, Lutz & Zifonun, Gisela (eds.), *Explorations in nominal inflection* (Interface Explorations 10), 189–227. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110197501.189

Papp, Ferenc. 1969. *A magyar nyelv szóvégmutató szótára*. Budapest: Akadémiai Kiadó.

Parker, Jeff & Sims, Andrea D. 2020. Irregularity, paradigmatic layers, and the complexity of inflection class systems: A study of Russian nouns. In Arkadiev, Peter & Gardani, Francesco (eds.), *The complexities of morphology*, chap. 2, 23–51. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198861287.003.0002

Pater, Joe. 2006. The locus of exceptionality: Morpheme-specific phonology. In Bateman, Leah & O'Keefe, Michael & Reilly, Ehren & Werley, Adam (eds.), *Papers in optimality theory III* (University of Massachusetts Occasional Papers in Linguistics 32). 259–296. Amherst: Graduate Linguistics Student Association.

Pater, Joe. 2010. Morpheme-specific phonology: Constraint indexation and inconsistency resolution. In Parker, Steve (ed.), *Phonological argumentation: Essays on evidence and motivation*, 123–154. London: Equinox.

Pertsova, Katya. 2004. *Distribution of genitive plural allomorphs in the Russian lexicon and in the internal grammar of native speakers*: University of California, Los Angeles MA thesis.

Pierrehumbert, Janet. 2001. Stochastic phonology. *Glot International* 5(6). 195–207.

R Core Team. 2022. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna. https://www.R-project.org/.

Rácz, Péter & Rebrus, Péter. 2012. Variation in the possessive allomorphy of Hungarian. In Kiefer, Ferenc & Ladányi, Mária & Siptár, Péter (eds.), *Current issues in morphological theory: (ir)regularity, analogy and frequency: Selected papers from the 14th International Morphology Meeting, budapest, 13–16 may 2010* (Current Issues in Linguistic Theory 322). 51–64. Amsterdam, Philadelphia: John Benjamins. https://doi.org/10.1075/cilt.322.03rac

Rebrus, Péter & Szigetvári, Péter. 2018. Morphosyntactically conditioned morphophonological variation. Presentation at PhonVar 1, Barcelona.

Rebrus, Péter & Szigetvári, Péter & Törkenczy, Miklós. 2019. Vacillation and lexical variation in Hungarian front/back harmony. Poster presented at the 7th Annual Meeting on Phonology, Stony Brook University.

Rebrus, Péter & Szigetvári, Péter & Tőrkenczy, Miklós. 2012. Dark secrets of Hungarian vowel harmony. In Cyran, Eugeniusz & Kardela, Henryk & Szymanek, Bogdan (eds.), *Sound, structure and sense: Studies in memory of Edmund Gussmann*, 491–508. Lublin: Wydawnictwo KUL.

Rounds, Carol. 2008. *Hungarian: An essential grammar*. London: Routledge 2nd edn. https://doi.org/10.4324/9780203886199

Rubach, Jerzy. 2013. Exceptional segments in Polish. *Natural Language & Linguistic Theory* 31(4). 1139–1162. https://doi.org/10.1007/s11049-013-9204-x

Rysling, Amanda. 2016. Polish yers revisited. *Catalan Journal of Linguistics* 15. 121–143. https://doi.org/10.5565/rev/catjl.187

Siptár, Péter & Törkenczy, Miklós. 2000. *The phonology of Hungarian* (The Phonology of the World's Languages). Oxford, New York: Oxford University Press.

Tabachnick, Guy. 2023. *Morphological dependencies*. New York: New York University dissertation.

Trommer, Jochen. 2021. The subsegmental structure of German plural allomorphy. *Natural Language & Linguistic Theory* 39. 601–656. https://doi.org/10.1007/s11049-020-09479-7

Voeten, Cesko C. 2022. *buildmer: Stepwise elimination and term reordering for mixed-effects regression*. https://CRAN.R-project.org/package=buildmer. R package version 2.6.

Wurzel, Wolfgang Ullrich. 1989. *Inflectional morphology and naturalness*. Dordrecht: Kluwer.