Speakers apply morphological dependencies in the inflection of novel forms

Guy Tabachnick

New York University

University of Connecticut LingLunch,
April 18, 2023

Introduction

Theories of morphology must account for stems inflecting in *different*, often *arbitrary* ways

- irregulars: English plural *oxen*, *sheep*, *syllabi*
- inflection classes: Russian nouns in class I–IV

Introduction

Theories of morphology must account for stems inflecting in *different*, often *arbitrary* ways

- irregulars: English plural *oxen*, *sheep*, *syllabi*
- inflection classes: Russian nouns in class I–IV

Arbitrary inflection of lexical items must be somehow *grammatically marked*

Today's study

Hungarian wug test explores one aspect of arbitrary inflection: correlations *between* inflected forms of a word (e.g. Ackerman et al., 2009; Ackerman and Malouf, 2013)

Today's study

Hungarian wug test explores one aspect of arbitrary inflection: correlations *between* inflected forms of a word (e.g. Ackerman et al., 2009; Ackerman and Malouf, 2013)

- If a Hungarian noun has -ɒk in the plural, it is likely to have -ɒ in the possessive

Today's study

Hungarian wug test explores one aspect of arbitrary inflection: correlations *between* inflected forms of a word (e.g. Ackerman et al., 2009; Ackerman and Malouf, 2013)

- If a Hungarian noun has -ɒk in the plural, it is likely to have -ɒ in the possessive
- Participants observed this gradient lexical tendency in the selection of novel forms

Today's study

Hungarian wug test explores one aspect of arbitrary inflection: correlations *between* inflected forms of a word (e.g. Ackerman et al., 2009; Ackerman and Malouf, 2013)

- If a Hungarian noun has **-ɒk in the plural**, it is likely to have **-ɒ in the possessive**
- Participants observed this gradient lexical tendency in the selection of novel forms

Gradient patterns and productivity

Productive extension of gradient lexical patterns

- Previous work: speakers extend phonological generalizations in wug tests, for example:

Gradient patterns and productivity

Productive extension of gradient lexical patterns

- Previous work: speakers extend phonological generalizations in wug tests, for example:
  - Hayes and Londe (2006); Hayes et al. (2009): certain vowels in Hungarian more likely to trigger back vs. front harmony for ambiguous words
  - Gouskova et al. (2015): Russian masculine nouns ending in consonant clusters form diminutives with -ik, not -ok or tʃʲik

Gradient patterns and productivity

Productive extension of gradient lexical patterns

- Previous work: speakers extend phonological generalizations in wug tests, for example:
  - Hayes and Londe (2006); Hayes et al. (2009): certain vowels in Hungarian more likely to trigger back vs. front harmony for ambiguous words
  - Gouskova et al. (2015): Russian masculine nouns ending in consonant clusters form diminutives with -ik, not -ok or tʃʲik
- Correlations between inflected forms can be handled using the same grammatical tools

Inflectional structure in morphological theory

Generative morphology has focused on *syncretism* and *inflection classes*

Inflectional structure in morphological theory

Generative morphology has focused on *syncretism* and *inflection classes*

- Syncretism: identity between realizations of different morphosyntactic features (e.g. Müller, 2004; Kramer, 2016; Caha, 2021)
  - Grammar induces identity, e.g. through shared structure, underspecification, impoverishment
  - Ex: Russian agreement doesn't show gender distinctions in the plural → rule deletes gender features in the context of PL

Inflectional structure in morphological theory

Generative morphology has focused on *syncretism* and *inflection classes*

- Syncretism: identity between realizations of different morphosyntactic features (e.g. Müller, 2004; Kramer, 2016; Caha, 2021)
  - Grammar induces identity, e.g. through shared structure, underspecification, impoverishment
  - Ex: Russian agreement doesn't show gender distinctions in the plural → rule deletes gender features in the context of PL
- Inflection class: lexical items "whose members each select for the same set of inflectional realizations" (Aronoff, 1994: 64)
  - Often assumed as discrete units of analysis, e.g. Russian "class I" (Corbett and Fraser, 1993; Müller, 2004; Caha, 2021)
  - These "macroclasses" often hide overlaps and complexities in inflectional patterns (Cameron-Faulkner and Carstairs-McCarthy, 2000; Finkel and Stump, 2007; Ackerman et al., 2009; Ackerman and Malouf, 2013; Bonami and Beniamine, 2016; Parker and Sims, 2020)

Main takeaway

Correlations between inflected forms exist outside of syncretism and inflectional macroclasses, and speakers apply them productively

Main takeaway

Correlations between inflected forms exist outside of syncretism and inflectional macroclasses, and speakers apply them productively

- Not a new insight (see Wurzel, 1989), but rarely discussed in generative work (rare exception: Halle and Marantz (2008))
- Well-established in lexicon (e.g. Ackerman et al., 2009; Ackerman and Malouf, 2013), but rarely if ever tested experimentally (rare exception: Bybee and Moder (1983))
- Theoretical work done by “inflection classes” can be shifted from hard-coded grammar to gradient pattern generalization

Main takeaway

Correlations between inflected forms exist outside of syncretism and inflectional macroclasses, and speakers apply them productively

- Not a new insight (see Wurzel, 1989), but rarely discussed in generative work (rare exception: Halle and Marantz (2008))
- Well-established in lexicon (e.g. Ackerman et al., 2009; Ackerman and Malouf, 2013), but rarely if ever tested experimentally (rare exception: Bybee and Moder (1983))
- Theoretical work done by “inflection classes” can be shifted from hard-coded grammar to gradient pattern generalization

These patterns can be learned as *phonotactics*

Outline

Morphological arbitrariness

- Arbitrary inflection of exceptional lexical items must be *grammatically marked*
- One common approach: *morphological features* (e.g. Lieber, 1980; Corbett and Baerman, 2006) that are attached as *diacritics* to lexical entries

Morphological arbitrariness

- Arbitrary inflection of exceptional lexical items must be *grammatically marked*
- One common approach: *morphological features* (e.g. Lieber, 1980; Corbett and Baerman, 2006) that are attached as *diacritics* to lexical entries
- Common subtype: *inflection class features*, which group together lexical items "whose members each select for the same set of inflectional realizations" (Aronoff, 1994: 64)

Inflection class features: the case of Russian

Russian feminine nouns: class II and III (Corbett and Baerman, 2006)

class	II	III
example	'newspaper'	'bone'
nominative	gazet-a	$kost^{j}$
dative	gazet-e	$kost^{j}$-i
instrumental	gazet-oj	$kost^{j}$-ju

Inflection class features: the case of Russian

Russian feminine nouns: class II and III (Corbett and Baerman, 2006)

class	II	III
example	'newspaper'	'bone'
nominative	gazet-a	kost^j-Ø
dative	gazet-e	kost^j-i
instrumental	gazet-oj	kost^j-ju

Feature-based analysis of Russian

The features II and III are each referenced in *multiple* (DM-style) *vocabulary insertion rules* (see Halle and Marantz, 1993; Müller, 2004; Embick and Marantz, 2008)

(1) *Vocabulary insertion rules for Russian cases*

a. NOM ↔ a / II ___
b. DAT ↔ e / II ___
c. INS ↔ oj / II ___
d. NOM ↔ Ø / III ___
e. DAT ↔ i / III ___
f. INS ↔ ju / III ___

Feature-based analysis of Russian

The features II and III are each referenced in *multiple* (DM-style) *vocabulary insertion rules* (see Halle and Marantz, 1993; Müller, 2004; Embick and Marantz, 2008)

(1) *Vocabulary insertion rules for Russian cases*

a.	NOM ↔ a / II ___	d.	NOM ↔ Ø / III ___
b.	DAT ↔ e / II ___	e.	DAT ↔ i / III ___
c.	INS ↔ oj / II ___	f.	INS ↔ ju / III ___

Feature-based analysis of Russian

The features II and III are each referenced in *multiple* (DM-style) *vocabulary insertion rules* (see Halle and Marantz, 1993; Müller, 2004; Embick and Marantz, 2008)

(1) *Vocabulary insertion rules for Russian cases*

a. NOM $\leftrightarrow$ a / II ___
b. DAT $\leftrightarrow$ e / II ___
c. INS $\leftrightarrow$ oj / II ___
d. NOM $\leftrightarrow$ Ø / III ___
e. DAT $\leftrightarrow$ i / III ___
f. INS $\leftrightarrow$ ju / III ___

(2) *Lexical entries for Russian nouns*

a. II: /gazet_{II}/ 'newspaper', /tʃert_{II}/ 'characteristic', /$\text{dol}^{\text{j}}_{\text{II}}$/ 'portion', …

b. III: /$\text{kost}^{\text{j}}_{\text{III}}$/ 'bone', /$\text{tetrad}^{\text{j}}_{\text{III}}$/ 'notebook', /$\text{ploɕːad}^{\text{j}}_{\text{III}}$/ 'square', …

Russian diacritic features facilitate inference

Novel dative [grid^j-i] ‘princely retinue’, **what’s the instrumental**?

Russian diacritic features facilitate inference

Novel dative [grid^j-i] 'princely retinue', **what's the instrumental**?

- Dative [i] → rule (7e)

(7e) DAT ↔ i / III ___

Russian diacritic features facilitate inference

Novel dative [grid^j-i] ‘princely retinue’, **what’s the instrumental**?

- Dative [i] → rule (7e)

(7e) DAT ↔ i / III ___

- Rule (7e) → III in noun’s lexical entry

(3) /grid^j$_{\text{III}}$/

Russian diacritic features facilitate inference

Novel dative [grid^j-i] ‘princely retinue’, **what’s the instrumental**?

- Dative [i] → rule (7e)

  (7e) DAT ↔ i / III ___

- Rule (7e) → III in noun’s lexical entry

  (3) /grid$^j_{III}$/

- III in noun’s lexical entry → rule (7f)

  (7f) INS ↔ ju / III ___

Russian diacritic features facilitate inference

Novel dative [grid^j-i] 'princely retinue', **what's the instrumental?**

- Dative [i] → rule (7e)

(7e) DAT ↔ i / III ___

- Rule (7e) → III in noun's lexical entry

(3) /grid$^j_{III}$/

- III in noun's lexical entry → rule (7f)

(7f) INS ↔ ju / III ___

- Rule (7f) → instrumental [ju]

(4) [grid^j-ju]

Russian diacritic features facilitate inference

Novel dative [grid^j-i] ‘princely retinue’, **what’s the instrumental**?

- Dative [i] → rule (7e)

(7e) DAT ↔ i / III ___

- Rule (7e) → III in noun’s lexical entry

(3) /grid$^j_{III}$/

- III in noun’s lexical entry → rule (7f)

(7f) INS ↔ ju / III ___

- Rule (7f) → instrumental [ju]

(4) [grid^j-ju]

The structure of the grammar, with features used in multiple rules, facilitates inference of new forms!

Narrowly tailored features: the case of Hungarian

Russian feminine nouns: class II and III (Corbett and Baerman, 2006)

class / *example*	II / 'newspaper'	III / 'bone'
nominative	gazet-a	kost^j-Ø
dative	gazet-e	kost^j-i
instrumental	gazet-oj	kost^j-ju

Narrowly tailored features: the case of Hungarian

Russian feminine nouns: class II and III (Corbett and Baerman, 2006)

class	II	III
example	'newspaper'	'bone'
nominative	gazet-a	kostʲ-Ø
dative	gazet-e	kostʲ-i
instrumental	gazet-oj	kostʲ-ju

Hungarian plural (-ok/-ɒk) and possessive (-ɒ/-jɒ): all four possible combinations (Rácz and Rebrus, 2012)

			"lowering stems"	
noun	dɒl	tʃont	vaːlː	hold
gloss	'song'	'bone'	'shoulder'	'moon'
plural	dɒl-ok	tʃont-ok	vaːlː-ɒk	hold-ɒk
possessive	dɒl-ɒ	tʃont-jɒ	vaːlː-ɒ	hold-jɒ

Narrowly tailored features: the case of Hungarian

Russian feminine nouns: class II and III (Corbett and Baerman, 2006)

class	II	III
example	'newspaper'	'bone'
nominative	gazet-a	$kost^j$-Ø
dative	gazet-e	$kost^j$-i
instrumental	gazet-oj	$kost^j$-ju

Hungarian plural (-ok/-ɒk) and possessive (-ɒ/-jɒ): all four possible combinations (Rácz and Rebrus, 2012)

			"lowering stems"	
noun	dɒl	tʃont	vaːlː	hold
gloss	'song'	'bone'	'shoulder'	'moon'
plural	dɒl-ok	tʃont-ok	vaːlː-ɒk	hold-ɒk
possessive	dɒl-ɒ	tʃont-jɒ	vaːlː-ɒ	hold-jɒ

Feature-based analysis of Hungarian

Features for the possessive ([±j]) and plural ([lower]) are each referenced in *one rule* (see Siptár and Törkenczy (2000) for an alternate analysis)

(5) *Vocabulary insertion rules for Hungarian plural and possessive*

a. PL ↔ ɒk / [lower] ___

b. PL ↔ ok

c. POSS ↔ jɒ / [+j] ___

d. POSS ↔ ɒ / [−j] ___

Feature-based analysis of Hungarian

Features for the possessive ([±j]) and plural ([lower]) are each referenced in *one rule* (see Siptár and Törkenczy (2000) for an alternate analysis)

(5) *Vocabulary insertion rules for Hungarian plural and possessive*

a. PL ↔ ɒk / [lower] ___

b. PL ↔ ok

c. POSS ↔ jɒ / [+j] ___

d. POSS ↔ ɒ / [−j] ___

Feature-based analysis of Hungarian

Features for the possessive ([±j]) and plural ([lower]) are each referenced in *one rule* (see Siptár and Törkenczy (2000) for an alternate analysis)

(5) *Vocabulary insertion rules for Hungarian plural and possessive*

a. PL ↔ ɒk / [lower] ___
b. PL ↔ ok
c. POSS ↔ jɒ / [+j] ___
d. POSS ↔ ɒ / [−j] ___

(6) *Lexical entries for Hungarian nouns*

a. [lower]: /vaːlː$_{[\text{lower},-\text{j}]}$/ 'shoulder', /hold$_{[\text{lower},+\text{j}]}$/ 'moon', /ɟaːr$_{[\text{lower},-\text{j}]}$/ 'factory', /ɲaːr$_{[\text{lower},+\text{j}]}$/ 'poplar', …

b. [+j]: /tʃont$_{[+\text{j}]}$/ 'bone', /hold$_{[\text{lower},+\text{j}]}$/ 'moon', /paːr$_{[+\text{j}]}$/ 'pair', /ɲaːr$_{[\text{lower},+\text{j}]}$/ 'poplar', …

c. [−j]: /dɒl$_{[-\text{j}]}$/ 'song', /vaːlː$_{[\text{lower},-\text{j}]}$/ 'shoulder', /kaːr$_{[-\text{j}]}$/ 'damage', /ɟaːr$_{[\text{lower},-\text{j}]}$/ 'factory', …

Feature-based analysis of Hungarian

Features for the possessive ([±j]) and plural ([lower]) are each referenced in *one rule* (see Siptár and Törkenczy (2000) for an alternate analysis)

(5) *Vocabulary insertion rules for Hungarian plural and possessive*

a. PL ↔ ɒk / [lower] ___

b. PL ↔ ok

c. POSS ↔ jɒ / [+j] ___

d. POSS ↔ ɒ / [−j] ___

(6) *Lexical entries for Hungarian nouns*

a. [lower]: /vaːlː$_{[\text{lower},-\text{j}]}$/ ‘shoulder’, /hold$_{[\text{lower},+\text{j}]}$/ ‘moon’, /ɟaːr$_{[\text{lower},-\text{j}]}$/ ‘factory’, /ɲaːr$_{[\text{lower},+\text{j}]}$/ ‘poplar’, …

b. [+j]: /tʃont$_{[+\text{j}]}$/ ‘bone’, /hold$_{[\text{lower},+\text{j}]}$/ ‘moon’, /paːr$_{[+\text{j}]}$/ ‘pair’, /ɲaːr$_{[\text{lower},+\text{j}]}$/ ‘poplar’, …

c. [−j]: /dɒl$_{[-\text{j}]}$/ ‘song’, /vaːlː$_{[\text{lower},-\text{j}]}$/ ‘shoulder’, /kaːr$_{[-\text{j}]}$/ ‘damage’, /ɟaːr$_{[\text{lower},-\text{j}]}$/ ‘factory’, …

Feature-based analysis of Hungarian

Features for the possessive ([±j]) and plural ([lower]) are each referenced in *one rule* (see Siptár and Törkenczy (2000) for an alternate analysis)

(5) *Vocabulary insertion rules for Hungarian plural and possessive*

a. PL ↔ ɒk / [lower] ___
b. PL ↔ ok
c. POSS ↔ jɒ / [+j] ___
d. POSS ↔ ɒ / [−j] ___

(6) *Lexical entries for Hungarian nouns*

a. [lower]: /vaːlː$_{[lower,-j]}$/ 'shoulder', /hold$_{[lower,+j]}$/ 'moon', /ɟaːr$_{[lower,-j]}$/ 'factory', /ɲaːr$_{[lower,+j]}$/ 'poplar', …

b. [+j]: /tʃont$_{[+j]}$/ 'bone', /hold$_{[lower,+j]}$/ 'moon', /paːr$_{[+j]}$/ 'pair', /ɲaːr$_{[lower,+j]}$/ 'poplar', …

c. [−j]: /dɒl$_{[-j]}$/ 'song', /vaːlː$_{[lower,-j]}$/ 'shoulder', /kaːr$_{[-j]}$/ 'damage', /ɟaːr$_{[lower,-j]}$/ 'factory', …

Feature-based analysis of Hungarian

Features for the possessive ([±j]) and plural ([lower]) are each referenced in *one rule* (see Siptár and Törkenczy (2000) for an alternate analysis)

(5) *Vocabulary insertion rules for Hungarian plural and possessive*

a. PL ↔ ɒk / [lower] ___
b. PL ↔ ok
c. POSS ↔ jɒ / [+j] ___
d. POSS ↔ ɒ / [−j] ___

(6) *Lexical entries for Hungarian nouns*

a. [lower]: /vaːlː$_{[\text{lower},-\text{j}]}$/ 'shoulder', /hold$_{[\text{lower},+\text{j}]}$/ 'moon', /ɟaːr$_{[\text{lower},-\text{j}]}$/ 'factory', /ɲaːr$_{[\text{lower},+\text{j}]}$/ 'poplar', …

b. [+j]: /tʃont$_{[+\text{j}]}$/ 'bone', /hold$_{[\text{lower},+\text{j}]}$/ 'moon', /paːr$_{[+\text{j}]}$/ 'pair', /ɲaːr$_{[\text{lower},+\text{j}]}$/ 'poplar', …

c. [−j]: /dɒl$_{[-\text{j}]}$/ 'song', /vaːlː$_{[\text{lower},-\text{j}]}$/ 'shoulder', /kaːr$_{[-\text{j}]}$/ 'damage', /ɟaːr$_{[\text{lower},-\text{j}]}$/ 'factory', …

Feature-based analysis of Hungarian

Features for the possessive ([±j]) and plural ([lower]) are each referenced in *one rule* (see Siptár and Törkenczy (2000) for an alternate analysis)

(5) *Vocabulary insertion rules for Hungarian plural and possessive*

a. PL ↔ ɒk / [lower] ___

b. PL ↔ ok

c. POSS ↔ jɒ / [+j] ___

d. POSS ↔ ɒ / [−j] ___

(6) *Lexical entries for Hungarian nouns*

a. [lower]: /vaːlː$_{[\text{lower},-\text{j}]}$/ 'shoulder', /hold$_{[\text{lower},+\text{j}]}$/ 'moon', /ɟaːr$_{[\text{lower},-\text{j}]}$/ 'factory', /ɲaːr$_{[\text{lower},+\text{j}]}$/ 'poplar', …

b. [+j]: /tʃont$_{[+\text{j}]}$/ 'bone', /hold$_{[\text{lower},+\text{j}]}$/ 'moon', /paːr$_{[+\text{j}]}$/ 'pair', /ɲaːr$_{[\text{lower},+\text{j}]}$/ 'poplar', …

c. [−j]: /dɒl$_{[-\text{j}]}$/ 'song', /vaːlː$_{[\text{lower},-\text{j}]}$/ 'shoulder', /kaːr$_{[-\text{j}]}$/ 'damage', /ɟaːr$_{[\text{lower},-\text{j}]}$/ 'factory', …

Hungarian diacritic features do not facilitate inference

Novel plural [maːl-ɒk] ‘belly furs’, **what’s the possessive**?

Hungarian diacritic features do not facilitate inference

Novel plural [maːl-ɒk] ‘belly furs’, **what’s the possessive**?

- Plural [ɒk] → rule (5a)

(5a) PL ↔ ɒk / [lower] ___

Hungarian diacritic features do not facilitate inference

Novel plural [maːl-ɒk] ‘belly furs’, **what’s the possessive**?

- Plural [ɒk] → rule (5a)

(5a) PL ↔ ɒk / [lower] ___

- Rule (5a) → [lower] in noun’s lexical entry

(7) $/\text{maːl}_{\text{[lower]}}/$

Hungarian diacritic features do not facilitate inference

Novel plural [maːl-ɒk] 'belly furs', **what's the possessive**?

- Plural [ɒk] → rule (5a)

(5a) PL ↔ ɒk / [lower] ___

- Rule (5a) → [lower] in noun's lexical entry

(7) /maːl$_{\text{[lower]}}$/

- [lower] in noun's lexical entry ↛ …

(5c) POSS ↔ jɒ / [+j] ___

(5d) POSS ↔ ɒ / [−j] ___

Hungarian diacritic features do not facilitate inference

Novel plural [maːl-ɒk] ‘belly furs’, **what’s the possessive**?

- Plural [ɒk] → rule (5a)

(5a) PL ↔ ɒk / [lower] ___

- Rule (5a) → [lower] in noun’s lexical entry

(7) /maːl$_{\text{[lower]}}$/

- [lower] in noun’s lexical entry ↛ …

(5c) POSS ↔ jɒ / [+j] ___

(5d) POSS ↔ ɒ / [−j] ___

Unlike in Russian, the structure of the grammar, with each feature used in a single rule, **does not** facilitate inference of new forms.

But *something must* facilitate inference

In the lexicon (Rácz and Rebrus, 2012):

But *something must* facilitate inference

In the lexicon (Rácz and Rebrus, 2012):

- Irregular nouns (including lowering stems) usually take -ɒ
- Some alveolar-final lowering stems do take -jɒ

But *something must* facilitate inference

In the lexicon (Rácz and Rebrus, 2012):

- Irregular nouns (including lowering stems) usually take -ɒ
- Some alveolar-final lowering stems do take -jɒ

Results:

But *something must* facilitate inference

In the lexicon (Rácz and Rebrus, 2012):

- Irregular nouns (including lowering stems) usually take -ɒ
- Some alveolar-final lowering stems do take -jɒ

Results:

- Speakers are more likely to assign possessive -ɒ to nonce lowering stems

But *something must* facilitate inference

In the lexicon (Rácz and Rebrus, 2012):

- **Irregular nouns (including lowering stems) usually take -ɒ**
- Some alveolar-final lowering stems do take -jɒ

Results:

- **Speakers are more likely to assign possessive -ɒ to nonce lowering stems**

Outline

Task

Stimulus presented twice in frame sentence

Task

Stimulus presented twice in frame sentence

- bare: lufɒn
- plural: lufɒnok (regular stem)

Task

Stimulus presented twice in frame sentence

- bare: lufɒn
- plural: lufɒnok (regular stem)

Participants see another frame sentence, select possessive from drop-down menu

Task

Stimulus presented twice in frame sentence

- bare: lufɒn
- plural: lufɒnok (regular stem)

Participants see another frame sentence, select possessive from drop-down menu

- [lufɒnɒ / lufɒnjɒ]

Task

Stimulus presented twice in frame sentence

- bare: lufɒn
- plural: lufɒnok **(regular stem)**

Participants see another frame sentence, select possessive from drop-down menu

- [lufɒnɒ / lufɒnjɒ]

Task

Stimulus presented twice in frame sentence

- bare: lufɒn
- plural: lufɒnɒk **(lowering stem)**

Participants see another frame sentence, select possessive from drop-down menu

- [lufɒnɒ / lufɒnjɒ]

Stats

- 90 participants
- 35–50 trials per participant
- ...of which 8–12 lowering stem trials
- 81 stimuli (57 target, 24 filler)
- 2,398 total target trials

Phonological frequency matching

Baseline: phonological model trained on corpus of monomorphemic words from Papp (1969)

Phonological frequency matching

Baseline: phonological model trained on corpus of monomorphemic words from Papp (1969)

- Given *word*, predicts odds of -jɒ as coefficient *phon_odds*
- final C manner + final C place + harmony class + final V height + final V length + final coda complexity + word length

Phonological frequency matching

Baseline: phonological model trained on corpus of monomorphemic words from Papp (1969)

- Given *word*, predicts odds of -jɒ as coefficient *phon_odds*
- final C manner + final C place + harmony class + final V height + final V length + final coda complexity + word length

Then: predict experimental results from phonological model

Phonological frequency matching

Baseline: phonological model trained on corpus of monomorphemic words from Papp (1969)

- Given *word*, predicts odds of -jɒ as coefficient *phon_odds*
- final C manner + final C place + harmony class + final V height + final V length + final coda complexity + word length

Then: predict experimental results from phonological model

- Given nonce word phonology and participant, predicts odds of -jɒ
- (1 | participant) + phon_odds

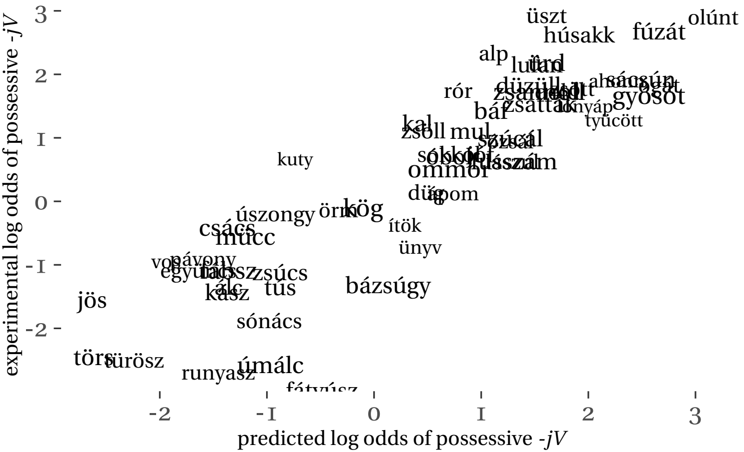
Phonological frequency matching

Baseline: phonological model trained on corpus of monomorphemic words from Papp (1969)

- Given *word*, predicts odds of -jɒ as coefficient ***phon_odds***
- final C manner + final C place + harmony class + final V height + final V length + final coda complexity + word length

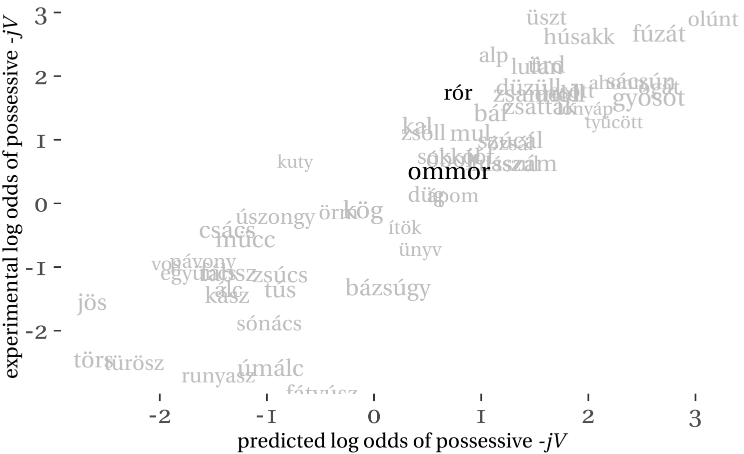
Then: predict experimental results from phonological model

- Given nonce word phonology and participant, predicts odds of -jɒ
- (1 | participant) + **phon_odds**



Baseline: the phonological model predicts experimental rate of possessives for *individual nonce words* quite well

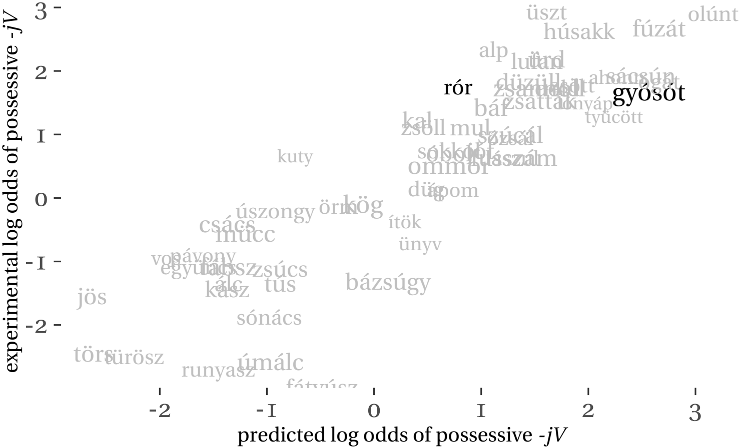
Results: phonological frequency matching



predicted: roːrjɒ = omːorjɒ

actual: roːrjɒ > omːorjɒ

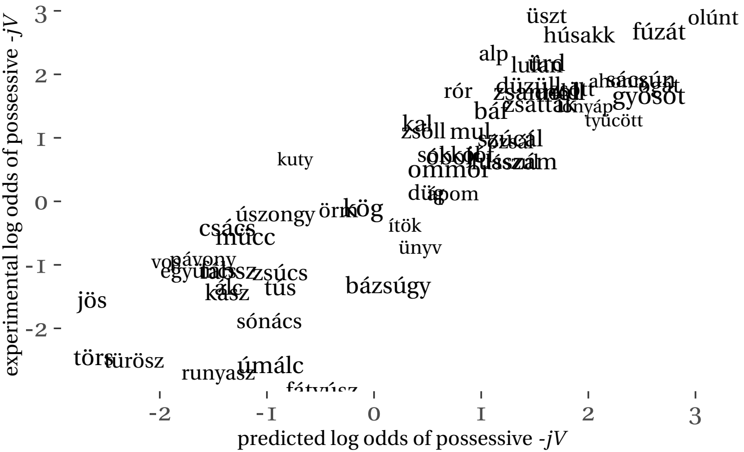
Results: phonological frequency matching



predicted: roːrjɒ < ɟoːʃoːtjɒ

actual: roːrjɒ = ɟoːʃoːtjɒ

Results: phonological frequency matching



Baseline: the phonological model predicts experimental rate of possessives for *individual nonce words* quite well

Sensitivity to morphology

Baseline: phonological model trained on corpus of monomorphemic words from Papp (1969)

- Given *word*, predicts odds of -jɒ as coefficient *phon_odds*
- final C manner + final C place + harmony class + final V height + final V length + final coda complexity + word length

Then: predict experimental results from phonological model

- Given nonce word phonology and participant, predicts odds of -jɒ
- (1 | participant) + phon_odds

Sensitivity to morphology

Baseline: phonological model trained on corpus of monomorphemic words from Papp (1969)

- Given *word*, predicts odds of -jɒ as coefficient *phon_odds*
- final C manner + final C place + harmony class + final V height + final V length + final coda complexity + word length

Then: predict experimental results from phonological model **and stem class**

- Given nonce word phonology and participant, predicts odds of -jɒ
- (1 | participant) + phon_odds

Sensitivity to morphology

Baseline: phonological model trained on corpus of monomorphemic words from Papp (1969)

- Given *word*, predicts odds of -jɒ as coefficient *phon_odds*
- final C manner + final C place + harmony class + final V height + final V length + final coda complexity + word length

Then: predict experimental results from phonological model **and stem class**

- Given nonce word phonology **and plural** and participant, predicts odds of -jɒ
- (1 | participant) + phon_odds

Sensitivity to morphology

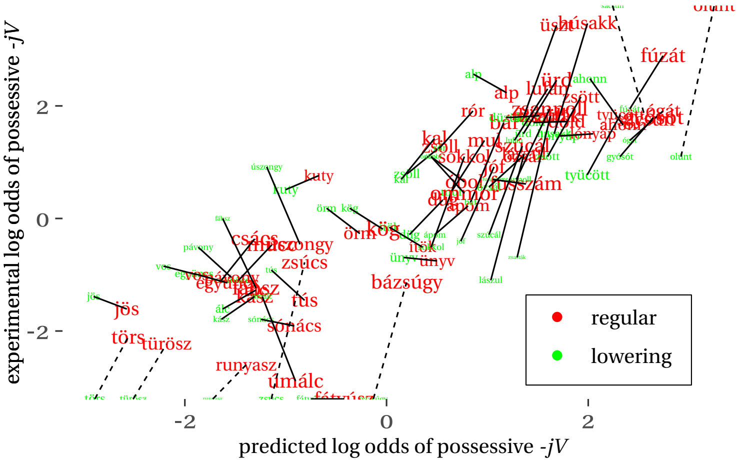
Baseline: phonological model trained on corpus of monomorphemic words from Papp (1969)

- Given *word*, predicts odds of -jɒ as coefficient *phon_odds*
- final C manner + final C place + harmony class + final V height + final V length + final coda complexity + word length

Then: predict experimental results from phonological model **and stem class**

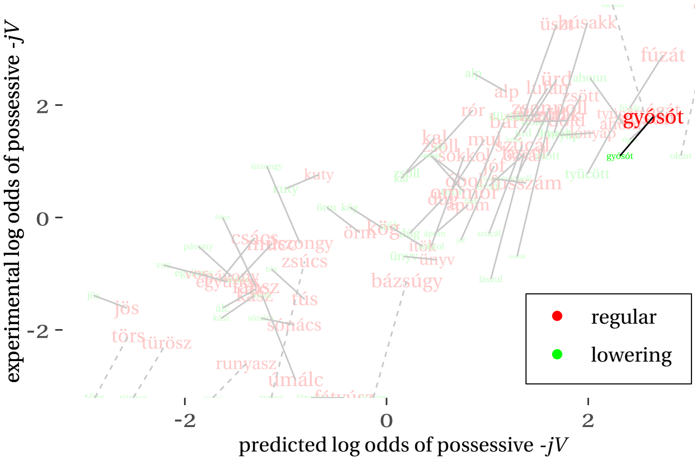
- Given nonce word phonology **and plural** and participant, predicts odds of -jɒ
- (1 | participant) + phon_odds **+ plural**

Results: sensitivity to morphology



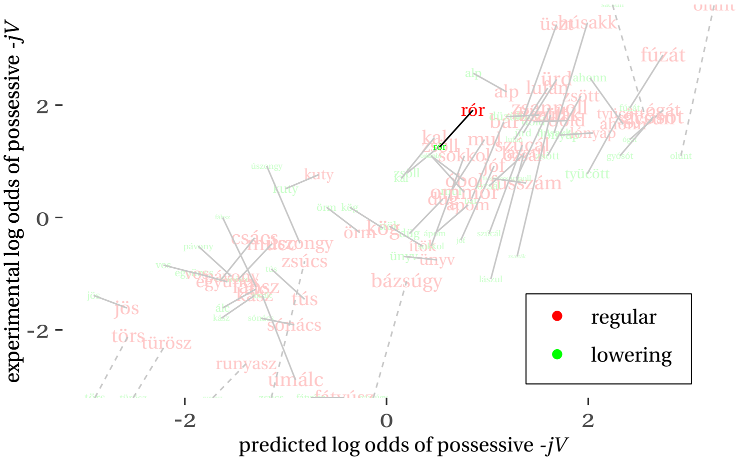
Target condition: most nonce words had a *lower* rate of -jɒ when presented as lowering stems

Results: sensitivity to morphology



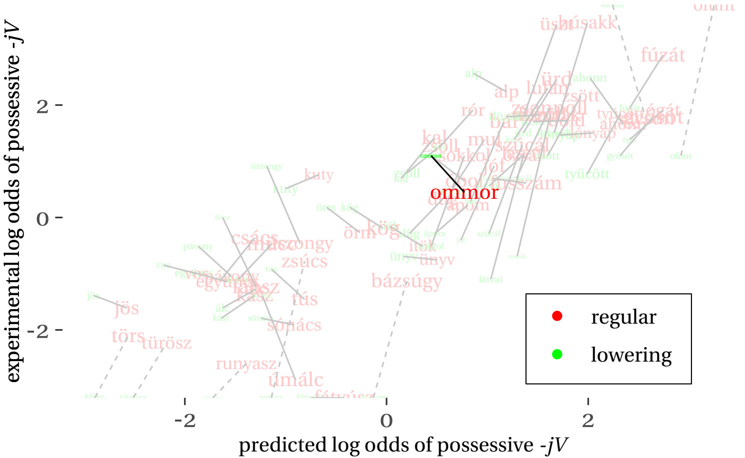
predicted: ɟoːsoːtok, ɟoːsoːtjɒ > ɟoːsoːtɒk, ɟoːsoːtjɒ
actual: ɟoːsoːtok, ɟoːsoːtjɒ > ɟoːsoːtɒk, ɟoːsoːtjɒ

Results: sensitivity to morphology



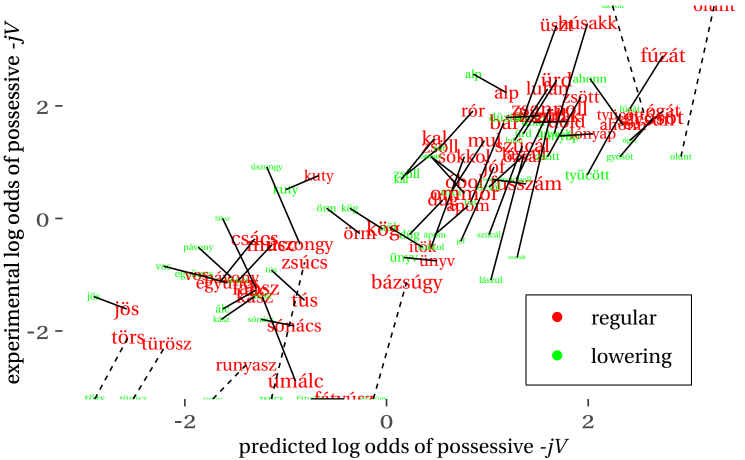
predicted: roːrok, roːrjɒ > roːrɒk, roːrjɒ
actual: roːrok, roːrjɒ > roːrɒk, roːrjɒ

Results: sensitivity to morphology



predicted: om:orok, om:orjɒ > om:orɒk, om:orjɒ
actual: om:orok, om:orjɒ < om:orɒk, om:orjɒ

Results: sensitivity to morphology



Target condition: most nonce words had a *lower* rate of -jɒ when presented as lowering stems

Results: summary

- Participants matched the phonological distribution of -jɒ and -ɒ in the lexicon

Results: summary

- Participants matched the phonological distribution of -jɒ and -ɒ in the lexicon
- …Taking this into account, they also assigned -ɒ more to nonce words with plural -ɒk

Outline

1. Background: morphological features and inflection class
2. Experiment: Hungarian possessive and plural
3. Discussion

Interpretation of results

Rácz and Rebrus (2012) and others: -jɒ is the productive default for most words

- recent loans and neologisms take -jɒ
- … unless they end in palatals and sibilants, in which case, they take -ɒ

Interpretation of results

Rácz and Rebrus (2012) and others: -jɒ is the productive default for most words

- recent loans and neologisms take -jɒ
- … unless they end in palatals and sibilants, in which case, they take -ɒ

My results: speakers used -jɒ and -ɒ on the same words

- gradient patterns extended from the lexicon

Interpretation of results

Rácz and Rebrus (2012) and others: -jɒ is the productive default for most words

- recent loans and neologisms take -jɒ
- … unless they end in palatals and sibilants, in which case, they take -ɒ

My results: speakers used -jɒ and -ɒ on the same words

- gradient patterns extended from the lexicon

No obvious explanation for difference, but …

Interpretation of results

Rácz and Rebrus (2012) and others: -jɒ is the productive default for most words

- recent loans and neologisms take -jɒ
- … unless they end in palatals and sibilants, in which case, they take -ɒ

My results: speakers used -jɒ and -ɒ on the same words

- gradient patterns extended from the lexicon

No obvious explanation for difference, but …

- clear that speakers have and can apply generalizations over the distribution of -jɒ and -ɒ in the lexicon
- these generalizations are both *phonological* and *morphological*

Generalizations and productivity

Existing formal models for productively learning phonological generalizations (e.g. Albright and Hayes, 2003; Hayes et al., 2009; Gouskova et al., 2015)

Generalizations and productivity

Existing formal models for productively learning phonological generalizations (e.g. Albright and Hayes, 2003; Hayes et al., 2009; Gouskova et al., 2015)

- speakers learn gradient phonological patterns as *weighted constraints* over portions of the lexicon (Hayes and Wilson, 2008; Gouskova et al., 2015)

Generalizations and productivity

Existing formal models for productively learning phonological generalizations (e.g. Albright and Hayes, 2003; Hayes et al., 2009; Gouskova et al., 2015)

- speakers learn gradient phonological patterns as *weighted constraints* over portions of the lexicon (Hayes and Wilson, 2008; Gouskova et al., 2015)
  - Hungarian nouns ending in sibilants always take -ɒ ([−j])
  - Hungarian nouns ending in vowels always take -jɒ ([+j])
  - for [+j] words: *[+strident]# (weight: 5)
  - for [−j] words: *[+syllabic]# (weight: 5)

Generalizations and productivity

Existing formal models for productively learning phonological generalizations (e.g. Albright and Hayes, 2003; Hayes et al., 2009; Gouskova et al., 2015)

- speakers learn gradient phonological patterns as *weighted constraints* over portions of the lexicon (Hayes and Wilson, 2008; Gouskova et al., 2015)
  - Hungarian nouns ending in sibilants always take -ɒ ([−j])
  - Hungarian nouns ending in vowels always take -jɒ ([+j])
  - for [+j] words: *[+strident]# (weight: 5)
  - for [−j] words: *[+syllabic]# (weight: 5)
- if inflectional patterns are marked with lexical diacritic features, weighted constraints can handle morphological dependencies as well

Generalizations and productivity

Existing formal models for productively learning phonological generalizations (e.g. Albright and Hayes, 2003; Hayes et al., 2009; Gouskova et al., 2015)

- speakers learn gradient phonological patterns as *weighted constraints* over portions of the lexicon (Hayes and Wilson, 2008; Gouskova et al., 2015)
  - Hungarian nouns ending in sibilants always take -ɒ ([−j])
  - Hungarian nouns ending in vowels always take -jɒ ([+j])
  - for [+j] words: *[+strident]# (weight: 5)
  - for [−j] words: *[+syllabic]# (weight: 5)
- if inflectional patterns are marked with lexical diacritic features, weighted constraints can handle morphological dependencies as well
  - Hungarian nouns with plural -ɒk ([lower]) usually take -ɒ ([−j])
  - for [+j] words: *[lower] (weight: 1)

Extending patterns to new words

To generate the possessive of a novel word, speakers evaluate it on the constraints for [+j] and [−j] words

Extending patterns to new words

To generate the possessive of a novel word, speakers evaluate it on the constraints for [+j] and [−j] words

Ex: speaker has seen plurals [ruɲɒs-ɒk] and [fuːzaːt-ɒk]

Extending patterns to new words

To generate the possessive of a novel word, speakers evaluate it on the constraints for [+j] and [−j] words

Ex: speaker has seen plurals [ruɲɒs-ɒk] and [fuːzaːt-ɒk]

1. evaluate /ruɲɒs$_{\text{[lower]}}$/ and /fuːzaːt$_{\text{[lower]}}$/ as [+j] words

constraint weight	*[+strident]# 5	*[lower] 1	total
ruɲɒs$_{\text{[lower]}}$	−5	−1	−6
fuːzaːt$_{\text{[lower]}}$	0	−1	−1

Extending patterns to new words

To generate the possessive of a novel word, speakers evaluate it on the constraints for [+j] and [−j] words

Ex: speaker has seen plurals [ruɲɒs-ɒk] and [fuːzaːt-ɒk]

1. evaluate /ruɲɒs$_{\text{[lower]}}$/ and /fuːzaːt$_{\text{[lower]}}$/ as [+j] words

constraint *weight*	*[+strident]# 5	*[lower] 1	total
ruɲɒs$_{\text{[lower]}}$	−5	−1	−6
fuːzaːt$_{\text{[lower]}}$	0	−1	−1

2. evaluate /ruɲɒs$_{\text{[lower]}}$/ and /fuːzaːt$_{\text{[lower]}}$/ as [−j] words

constraint *weight*	*[+syllabic]# 5	total
ruɲɒs$_{\text{[lower]}}$	0	0
fuːzaːt$_{\text{[lower]}}$	0	0

Extending patterns to new words

To generate the possessive of a novel word, speakers evaluate it on the constraints for [+j] and [−j] words

Ex: speaker has seen plurals [ruɲɒs-ɒk] and [fuːzaːt-ɒk]

1. evaluate /ruɲɒs$_{\text{[lower]}}$/ and /fuːzaːt$_{\text{[lower]}}$/ as [+j] words

constraint weight	*[+strident]# 5	*[lower] 1	total
ruɲɒs$_{\text{[lower]}}$	−5	−1	−6
fuːzaːt$_{\text{[lower]}}$	0	−1	−1

2. evaluate /ruɲɒs$_{\text{[lower]}}$/ and /fuːzaːt$_{\text{[lower]}}$/ as [−j] words

constraint weight	*[+syllabic]# 5	total
ruɲɒs$_{\text{[lower]}}$	0	0
fuːzaːt$_{\text{[lower]}}$	0	0

3. stochastically assign [+j] or [−j] (maximum entropy (Hayes and Wilson, 2008): for outcome x with score $H(x)$, $P(x) \propto e^{H(x)}$)

Extending patterns to new words

Ex: speaker has seen plurals [ruɲɒs-ɒk] and [fuːzaːt-ɒk]

1. evaluate /ruɲɒs$_{\text{[lower]}}$/ and /fuːzaːt$_{\text{[lower]}}$/ as [+j] words

constraint weight	*[+strident]# 5	*[lower] 1	total
ruɲɒs$_{\text{[lower]}}$	−5	−1	−6
fuːzaːt$_{\text{[lower]}}$	0	−1	−1

2. evaluate /ruɲɒs$_{\text{[lower]}}$/ and /fuːzaːt$_{\text{[lower]}}$/ as [−j] words

constraint weight	*[+syllabic]# 5	total
ruɲɒs$_{\text{[lower]}}$	0	0
fuːzaːt$_{\text{[lower]}}$	0	0

3. stochastically assign [+j] or [−j] (maximum entropy (Hayes and Wilson, 2008): for outcome x with score $H(x)$, $P(x) \propto e^{H(x)}$)
   - $P(/\text{ruɲɒs}_{\text{[lower,[+j]]}}/) = \frac{e^{H([+j])}}{e^{H([+j])}+e^{H([-j])}} = \frac{e^{-6}}{e^{-6}+e^{0}} = .002 = 0.2\%$
   - $P(/\text{fuːzaːt}_{\text{[lower,[+j]]}}/) = \frac{e^{H([+j])}}{e^{H([+j])}+e^{H([-j])}} = \frac{e^{-1}}{e^{-1}+e^{0}} = .269 = 26.9\%$

Weighted constraint models and their uses

Maximum entropy grammars with weighted constraints are useful across phonological and morphological domains

Weighted constraint models and their uses

Maximum entropy grammars with weighted constraints are useful across phonological and morphological domains

- *phonotactics*: grammar for the entire lexicon defines phonotactically good and bad words (Hayes and Wilson, 2008)
- *generalizations over lexically conditioned allomorphy*: grammar for morphologically defined sublexicons defines good and bad words for a lexically exceptional class (Gouskova et al., 2015)
- *morphological dependencies*: since morphological features like [+j] and [lower] are present in underlying forms, they can also define good and bad words for a different lexical class

Weighted constraint models and their uses

Maximum entropy grammars with weighted constraints are useful across phonological and morphological domains

- *phonotactics*: grammar for the entire lexicon defines phonotactically good and bad words (Hayes and Wilson, 2008)
- *generalizations over lexically conditioned allomorphy*: grammar for morphologically defined sublexicons defines good and bad words for a lexically exceptional class (Gouskova et al., 2015)
- *morphological dependencies*: since morphological features like [+j] and [lower] are present in underlying forms, they can also define good and bad words for a different lexical class

Phonological and morphological effects are evaluated together, in a single analysis

We can handle morphological dependencies using independently necessary general phonological mechanisms

Do we need inflection class features at all?

Previously: Russian and Hungarian are categorically distinct

Do we need inflection class features at all?

Previously: Russian and Hungarian are categorically distinct

(5) *Vocabulary insertion rules for Hungarian plural and possessive*

a. PL ↔ ɒk / [lower] ___

b. PL ↔ ok

c. POSS ↔ jɒ / [+j] ___

d. POSS ↔ ɒ / [−j] ___

(1) *Vocabulary insertion rules for Russian cases*

a. NOM ↔ a / II ___

b. DAT ↔ e / II ___

c. INS ↔ oj / II ___

d. NOM ↔ Ø / III ___

e. DAT ↔ i / III ___

f. INS ↔ ju / III ___

Do we need inflection class features at all?

Previously: Russian and Hungarian are categorically distinct

(5) *Vocabulary insertion rules for Hungarian plural and possessive*

a. PL ↔ ɒk / [lower] ___

b. PL ↔ ok

c. POSS ↔ jɒ / [+j] ___

d. POSS ↔ ɒ / [−j] ___

(1) *Vocabulary insertion rules for Russian cases*

a. NOM ↔ a / II ___

b. DAT ↔ e / II ___

c. INS ↔ oj / II ___

d. NOM ↔ Ø / III ___

e. DAT ↔ i / III ___

f. INS ↔ ju / III ___

Ackerman et al. (2009); Baerman et al. (2017) and others: Russian and Hungarian differ in *degree* of cohesion, not kind
(indeed, actual Russian inflection is messier than the oversimplified four-class analysis (Parker and Sims, 2020))

Do we need inflection class features at all?

Previously: Russian and Hungarian are categorically distinct

(5) *Vocabulary insertion rules for Hungarian plural and possessive*

a. PL ↔ ɒk / [lower] ___
b. PL ↔ ok
c. POSS ↔ jɒ / [+j] ___
d. POSS ↔ ɒ / [−j] ___

(1) *Vocabulary insertion rules for Russian cases*

a. NOM ↔ a / II ___
b. DAT ↔ e / II ___
c. INS ↔ oj / II ___
d. NOM ↔ Ø / III ___
e. DAT ↔ i / III ___
f. INS ↔ ju / III ___

Ackerman et al. (2009); Baerman et al. (2017) and others: Russian and Hungarian differ in *degree* of cohesion, not kind
(indeed, actual Russian inflection is messier than the oversimplified four-class analysis (Parker and Sims, 2020))

- We need separate generalizations to capture Hungarian morphological dependency between -ɒk and -ɒ
- Maybe Russian-style "inflection classes" are just very strong morphological generalizations

Do we need inflection class features at all?

Alternate Russian analysis

Do we need inflection class features at all?

Alternate Russian analysis

(1′) *Vocabulary insertion rules for Russian cases*

a.	NOM ↔ a	/ [N:a] ___		d.	NOM ↔ Ø	/ [N:Ø] ___
b.	DAT ↔ e	/ [D:e] ___		e.	DAT ↔ i	/ [D:i] ___
c.	INS ↔ oj	/ [I:oj] ___		f.	INS ↔ ju	/ [I:ju] ___

Do we need inflection class features at all?

Alternate Russian analysis

(1′) *Vocabulary insertion rules for Russian cases*

a.	NOM $\leftrightarrow$ a / [N:a] ___		d.	NOM $\leftrightarrow$ Ø / [N:Ø] ___	
b.	DAT $\leftrightarrow$ e / [D:e] ___		e.	DAT $\leftrightarrow$ i / [D:i] ___	
c.	INS $\leftrightarrow$ oj / [I:oj] ___		f.	INS $\leftrightarrow$ ju / [I:ju] ___	

(2′) *Lexical entries for Russian nouns*

a. II: /$\text{gazet}_{[\text{N:a,D:e,I:oj}]}$/ 'newspaper', /$\text{tʃert}_{[\text{N:a,D:e,I:oj}]}$/ 'characteristic', /$\text{dol}^{\text{j}}{}_{[\text{N:a,D:e,I:oj}]}$/ 'portion', …

b. III: /$\text{kost}^{\text{j}}{}_{[\text{N:Ø,D:i,I:ju}]}$/ 'bone', /$\text{tetrad}^{\text{j}}{}_{[\text{N:Ø,D:i,I:ju}]}$/ 'notebook', /$\text{ploɕːad}^{\text{j}}{}_{[\text{N:Ø,D:i,I:ju}]}$/ 'square', …

Do we need inflection class features at all?

Alternate Russian analysis

(1′) *Vocabulary insertion rules for Russian cases*

a.	NOM ↔ a / [N:a] ___		d.	NOM ↔ Ø / [N:Ø] ___
b.	DAT ↔ e / [D:e] ___		e.	DAT ↔ i / [D:i] ___
c.	INS ↔ oj / [I:oj] ___		f.	INS ↔ ju / [I:ju] ___

(2′) *Lexical entries for Russian nouns*

a. II: /$\text{gazet}_{[\text{N:a,D:e,I:oj}]}$/ ‘newspaper’, /$\text{tʃert}_{[\text{N:a,D:e,I:oj}]}$/ ‘characteristic’, /$\text{dol}^{\text{j}}{}_{[\text{N:a,D:e,I:oj}]}$/ ‘portion’, …

b. III: /$\text{kost}^{\text{j}}{}_{[\text{N:Ø,D:i,I:ju}]}$/ ‘bone’, /$\text{tetrad}^{\text{j}}{}_{[\text{N:Ø,D:i,I:ju}]}$/ ‘notebook’, /$\text{ploɕ:ad}^{\text{j}}{}_{[\text{N:Ø,D:i,I:ju}]}$/ ‘square’, …

(8) *Heavily weighted constraints for sublexicons*

a.	for [N:a] nouns: *[D:i]		c.	for [N:Ø] nouns: *[D:e]
b.	for [N:a] nouns: *[I:ju]		d.	for [N:Ø] nouns: *[I:oj]
	…			…

Summary

- Hungarian speakers productively apply correlations between inflected forms in the lexicon
- These cases are not well-suited for an “inflection class” analysis
- We need a way to account for gradient correlations between narrowly targeted inflectional features
- Gradient constraint-based phonotactic models can be easily extended to do this
- Inflection classes can be recast as *emergent* clusters of strong correlations between narrowly targeted features

References I

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

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

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

Aronoff, M. (1994). *Morphology by Itself: Stems and Inflectional Classes.* Number 22 in Linguistic Inquiry Monographs. MIT Press, Cambridge.

Baerman, M., Brown, D., and Corbett, G. G. (2017). *Morphological Complexity.* Number 153 in Cambridge Studies in Linguistics. Cambridge University Press, Cambridge.

Bermel, N. and Knittl, L. (2012). Morphosyntactic variation and syntactic constructions in Czech nominal declension: corpus frequency and native-speaker judgements. *Russian Linguistics*, 36(1):91–119.

Bonami, O. and Beniamine, S. (2016). Joint predictiveness in inflectional paradigms. *Word Structure*, 9(2):156–182.

Bybee, J. L. and Moder, C. L. (1983). Morphological classes as natural categories. *Language*, 59(2):251–270.

Caha, P. (2021). Modeling declensions without declension features: The case of Russian. *Acta Linguistica Academica*, 68(4):385–425.

Cameron-Faulkner, T. and Carstairs-McCarthy, A. (2000). Stem alternants as morphological stigmata: Evidence from blur avoidance in Polish nouns. *Natural Language and Linguistic Theory*, 18:813–835.

Corbett, G. G. and Baerman, M. (2006). Prolegomena to a typology of morphological features. *Morphology*, 16:231–246.

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

Embick, D. and Marantz, A. (2008). Architecture and blocking. *Linguistic Inquiry*, 39(1):1–53.

Finkel, R. and Stump, G. (2007). Principal parts and morphological typology. *Morphology*, 17:39–75.

References II

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

Guzmán Naranjo, M. and Bonami, O. (2021). Overabundance and inflectional classification: Quantitative evidence from Czech. *Glossa: a journal of general linguistics*, 6(1):88. 1–31.

Halle, M. and Marantz, A. (1993). Distributed morphology and the pieces of inflection. In Hale, K. and Keyser, S. J., editors, *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, number 24 in Current Studies in Linguistics, chapter 3, pages 111–176. MIT Press, Cambridge.

Halle, M. and Marantz, A. (2008). Clarifying "Blur": Paradigms, defaults, and inflectional classes. In Bachrach, A. and Nevins, A., editors, *Inflectional Identity*, number 18 in Oxford Studies in Theoretical Linguistics, chapter 3, pages 55–72. Oxford University Press, Oxford.

Hayes, B. and Londe, Z. C. (2006). Stochastic phonological knowledge: the case of Hungarian vowel harmony. *Phonology*, 23:59–104.

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

Hayes, B., Zuraw, K., Siptár, P., and Londe, Z. (2009). Natural and unnatural constraints in Hungarian vowel harmony. *Language*, 85(4):822–863.

Kramer, R. (2016). Syncretism in paradigm function morphology and distributed morphology. In Siddiqi, D. and Harley, H., editors, *Morphological Metatheory*, number 229 in Linguistic Aktuell/Linguistics Today, pages 95–120. John Benjamins, Amsterdam, Philadelphia.

Křen, M., Cvrček, V., Hnátková, M., Jelínek, T., Kocek, J., Kovářiková, D., Křivan, J., Milička, J., Petkevič, V., Procházka, P., Skoumalová, H., Šindlerová, J., and Škrabal, M. (2022). Korpus SYN, verze 11 z 14.12.2022. www.korpus.cz.

Lieber, R. (1980). *On the Organization of the Lexicon*. PhD thesis, Massachusetts Institute of Technology, Cambridge.

References III

Müller, G. (2004). On decomposing inflection class features: Syncretism in Russian noun inflection. In Müller, G., Gunkel, L., and Zifonun, G., editors, *Explorations in Nominal Inflection*, number 10 in Interface Explorations, pages 189–227. Mouton de Gruyter, Berlin.

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

Parker, J. and Sims, A. D. (2020). Irregularity, paradigmatic layers, and the complexity of inflection class systems: A study of Russian nouns. In Arkadiev, P. and Gardani, F., editors, *The Complexities of Morphology*, chapter 2, pages 23–51. Oxford University Press, Oxford.

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

Siptár, P. and Törkenczy, M. (2000). *The Phonology of Hungarian.* The Phonology of the World's Languages. Oxford University Press, Oxford, New York.

Wurzel, W. U. (1989). *Inflectional Morphology and Naturalness.* Kluwer, Dordrecht.

Full procedure

Sample trial (regular plural)

Full procedure

Sample trial (regular plural)

In 1997, the **lufɒn** entered into the competition for flowery **lufɒnok** for the first time.

Please select the word's plural form: [lufɒnøk / lufɒnɒk / lufɒnɛk / lufɒnok]

Full procedure

Sample trial (regular plural)

In 1997, the **lufɒn** entered into the competition for flowery **lufɒnok** for the first time.

Please select the word's plural form: [lufɒnøk / lufɒnɒk / lufɒnɛk / **lufɒnok**]

Full procedure

Sample trial (regular plural)

In 1997, the **lufɒn** entered into the competition for flowery **lufɒnok** for the first time.

Please select the word's plural form: [lufɒnøk / lufɒnɒk / lufɒnɛk / **lufɒnok**]

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

My [lufɒnɒm / lufɒnɛm / lufɒnøm / lufɒnom] couldn't sing well, however my husband's [lufɒnɛ / lufɒnjɛ / lufɒnɒ / lufɒnjɒ] sang brilliantly.

Full procedure

Sample trial (regular plural)

In 1997, the **lufɒn** entered into the competition for flowery **lufɒnok** for the first time.

Please select the word's plural form: [lufɒnøk / lufɒnɒk / lufɒnɛk / **lufɒnok**]

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

My [lufɒnɒm / lufɒnɛm / lufɒnøm / **lufɒnom**] couldn't sing well, however my husband's [lufɒnɛ / lufɒnjɛ / lufɒnɒ / lufɒnjɒ] sang brilliantly.

Full procedure

Sample trial (regular plural)

In 1997, the **lufɒn** entered into the competition for flowery **lufɒnok** for the first time.

Please select the word's plural form: [lufɒnøk / lufɒnɒk / lufɒnɛk / **lufɒnok**]

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

My [lufɒnɒm / lufɒnɛm / lufɒnøm / **lufɒnom**] couldn't sing well, however my husband's [lufɒnɛ / lufɒnjɛ / **lufɒnɒ** / **lufɒnjɒ**] sang brilliantly.

Full procedure

Sample trial (lowering stem)

In 1997, the **lufɒn** entered into the competition for flowery **lufɒnɒk** for the first time.

Please select the word's plural form: [lufɒnøk / **lufɒnɒk** / lufɒnɛk / lufɒnok]

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

My [**lufɒnɒm** / lufɒnɛm / lufɒnøm / lufɒnom] couldn't sing well, however my husband's [lufɒnɛ / lufɒnjɛ / **lufɒnɒ** / **lufɒnjɒ**] sang brilliantly.

Phonological model of lexicon

	β coef	*SE*	*Wald z*	*p*
Intercept	**3.02**	**.32**	**9.55**	**<.0001**
C Manner (default: plosive)				
fricative	**−1.44**	**.39**	**−3.73**	**.0002**
sibilant	**−10.69**	**.80**	**−13.36**	**<.0001**
nasal	**−1.95**	**.27**	**−7.16**	**<.0001**
approximant	**−4.08**	**.30**	**−13.47**	**<.0001**
C Place (default: alveolar)				
labial	**−2.02**	**.26**	**−7.94**	**<.0001**
palatal	**−8.88**	**1.10**	**−8.06**	**<.0001**
velar	**−3.26**	**.29**	**−10.96**	**<.0001**
Harmony (default: back)				
front	**−2.03**	**.18**	**−10.96**	**<.0001**
variable	**2.26**	**.97**	**2.33**	**.0197**
V Height (default: mid)				
high	**1.73**	**.22**	**7.89**	**<.0001**
low	.28	.19	1.50	.1342
V Length (default: short)				
long	**1.40**	**.17**	**7.98**	**<.0001**
Coda (default: singleton)				
geminate	**2.47**	**.40**	**6.25**	**<.0001**
cluster	.04	.21	0.18	.8602
Syllables (default: monosyllabic)				
polysyllabic	**1.15**	**.17**	**6.67**	**<.0001**

Phonological model of experimental results

Random effect	*variance*	*SD*		
Participant	.55	.74		

Fixed effects	β *coef*	*SE*	*Wald z*	*p*
Intercept	**.67**	**.10**	**7.03**	**$<$.0001**
Phon_odds	**.34**	**.01**	**22.76**	**$<$.0001**

Phonological and morphological model of experimental results

Random effect	*variance*	*SD*		
Participant	.54	.74		
Fixed effects	*β coef*	*SE*	*Wald z*	*p*
Intercept	**.74**	**.10**	**7.48**	**<.0001**
Phon_odds	**.34**	**.02**	**22.77**	**<.0001**
Plural (default: -ok)				
-ɒk	**−.33**	**.13**	**−2.62**	**.0086**

Experiment 2: Czech locative and genitive

Czech genitive (-u/-a) and locative (-u/-ɛ): all four possible combinations (for *masculine inanimate hard-stem* nouns)

noun	problɛːm	zaːpas	vɛtʃɛr	kostɛl
gloss	‘problem’	’match’	‘evening’	’church’
genitive	problɛːm-u	zaːpas-u	vɛtʃɛr-a	kostɛl-a
locative	problɛːm-u	zaːpas-ɛ	vɛtʃɛr-u	kostɛl-ɛ

Experiment 2: Czech locative and genitive

Czech genitive (-u/-a) and locative (-u/-ɛ): all four possible combinations (for *masculine inanimate hard-stem* nouns)

noun *gloss*	problɛːm ‘problem’	zaːpas ’match’	vɛtʃɛr ‘evening’	kostɛl ’church’
genitive	problɛːm-u	zaːpas-u	vɛtʃɛr-a	kostɛl-a
locative	problɛːm-u	zaːpas-ɛ	vɛtʃɛr-u	kostɛl-ɛ

Experiment 2: Czech locative and genitive

Czech genitive (-u/-a) and locative (-u/-ɛ): all four possible combinations (for *masculine inanimate hard-stem* nouns)

noun *gloss*	problɛːm ‘problem’	zaːpas ’match’	vɛtʃɛr ‘evening’	kostɛl ’church’
genitive	problɛːm-u	zaːpas-u	vɛtʃɛr-a	kostɛl-a
locative	problɛːm-u	zaːpas-ɛ	vɛtʃɛr-u	kostɛl-ɛ

Historically: innovative -u has pushed out original -a and -ɛ in both cases

Experiment 2: Czech locative and genitive

Czech genitive (-u/-a) and locative (-u/-ɛ): all four possible combinations (for *masculine inanimate hard-stem* nouns)

noun	problɛːm	zaːpas	vɛtʃɛr	kostɛl
gloss	‘problem’	’match’	‘evening’	’church’
genitive	problɛːm-u	zaːpas-u	vɛtʃɛr-a	kostɛl-a
locative	problɛːm-u	zaːpas-ɛ	vɛtʃɛr-u	kostɛl-ɛ

Historically: innovative -u has pushed out original -a and -ɛ in both cases

- Today -u is much more common
- Morphological dependency: nouns that take genitive -a also tend to take locative -ɛ

Background: variation

Most nouns that take genitive -a or locative -ɛ do so *variably* (Bermel and Knittl, 2012; Guzmán Naranjo and Bonami, 2021)

Background: variation

Most nouns that take genitive -a or locative -ɛ do so *variably* (Bermel and Knittl, 2012; Guzmán Naranjo and Bonami, 2021)

- *lexically conditioned*: variable nouns have different rates of -ɛ: [most] 'bridge' prefers -ɛ, while [uːr̝ad] 'office' prefers -u

Background: variation

Most nouns that take genitive -a or locative -ɛ do so *variably* (Bermel and Knittl, 2012; Guzmán Naranjo and Bonami, 2021)

- *lexically conditioned*: variable nouns have different rates of -ɛ: [most] 'bridge' prefers -ɛ, while [uːr̝ad] 'office' prefers -u
- *syntactically conditioned*: "canonical" locative prepositions like [v] 'in' prefer -ɛ relative to less "canonical" prepositions like [o] 'about'

Background: variation

Most nouns that take genitive -a or locative -ɛ do so *variably* (Bermel and Knittl, 2012; Guzmán Naranjo and Bonami, 2021)

- *lexically conditioned*: variable nouns have different rates of -ɛ: [most] ‘bridge’ prefers -ɛ, while [uːɻad] ‘office’ prefers -u
- *syntactically conditioned*: “canonical” locative prepositions like [v] ‘in’ prefer -ɛ relative to less “canonical” prepositions like [o] ‘about’

	in		*about*
bridge	v mosc-ɛ	>	o mosc-ɛ
	∨		∨
office	v uːɻaɟ-ɛ	>	o uːɻaɟ-ɛ

Background: variation

Most nouns that take genitive -a or locative -ɛ do so *variably* (Bermel and Knittl, 2012; Guzmán Naranjo and Bonami, 2021)

- *lexically conditioned*: variable nouns have different rates of -ɛ: [most] 'bridge' prefers -ɛ, while [uːɹ̝ad] 'office' prefers -u
- *syntactically conditioned*: "canonical" locative prepositions like [v] 'in' prefer -ɛ relative to less "canonical" prepositions like [o] 'about'

	in		*about*
bridge	v mosc-ɛ	>	o mosc-ɛ
	∨		∨
office	v uːɹ̝aɟ-ɛ	>	o uːɹ̝aɟ-ɛ

For variable nouns, a higher rate of genitive -a corresponds to a higher rate of locative -ɛ

Task

Stimulus presented twice in frame sentence

Task

Stimulus presented twice in frame sentence

- bare: cɪs
- prep + genitive: z cɪsu ([z] ‘out of’)

Task

Stimulus presented twice in frame sentence

- bare: cɪs
- prep + genitive: z cɪsu ([z] ‘out of’)

Participants see another frame sentence, select genitive and locative from drop-down menus

Task

Stimulus presented twice in frame sentence

- bare: cɪs
- prep + genitive: z cɪsu ([z] 'out of')

Participants see another frame sentence, select genitive and locative from drop-down menus

- prep + genitive: [do cɪsu / do cɪsa] ([do] 'into')
- prep + locative: [na cɪsu / na cɪsɛ] ([na] 'on')

Task

Stimulus presented twice in frame sentence

- bare: cɪs
- prep + genitive: z cɪsa ([z] 'out of')

Participants see another frame sentence, select genitive and locative from drop-down menus

- prep + genitive: [do cɪsu / do cɪsa] ([do] 'into')
- prep + locative: [na cɪsu / na cɪsɛ] ([na] 'on')

Stats

- 88 participants
- 50 trials per participant
- …of which 12 shown with genitive -a
- 82 stimuli
- 4,397 total target trials

Phonological frequency matching

Baseline: phonological model trained on noun tokens from the Czech National Corpus (Křen et al., 2022)

Phonological frequency matching

Baseline: phonological model trained on noun tokens from the Czech National Corpus (Křen et al., 2022)

- Nouns binned as categorical ($>$ 99% of tokens) locative -u vs. variable/categorical -ɛ

Phonological frequency matching

Baseline: phonological model trained on noun tokens from the Czech National Corpus (Křen et al., 2022)

- Nouns binned as categorical ($>$ 99% of tokens) locative -u vs. variable/categorical -ɛ
- Given *word*, predicts odds of -ɛ as coefficient *phon_odds*
- final C place + final C manner + final coda complexity + final V length + final V front + final V height + word length

Phonological frequency matching

Baseline: phonological model trained on noun tokens from the Czech National Corpus (Křen et al., 2022)

- Nouns binned as categorical ($> 99\%$ of tokens) locative -u vs. variable/categorical -ɛ
- Given *word*, predicts odds of -ɛ as coefficient *phon_odds*
- final C place + final C manner + final coda complexity + final V length + final V front + final V height + word length
- Phonology has a *much smaller effect* than in Hungarian ($R^2 = .09$)
- In part because the data are so skewed (94% of nouns take -u)

Phonological frequency matching

Baseline: phonological model trained on noun tokens from the Czech National Corpus (Křen et al., 2022)

- Nouns binned as categorical ($>$ 99% of tokens) locative -u vs. variable/categorical -ɛ
- Given *word*, predicts odds of -ɛ as coefficient *phon_odds*
- final C place + final C manner + final coda complexity + final V length + final V front + final V height + word length
- Phonology has a *much smaller effect* than in Hungarian ($R^2 = .09$)
- In part because the data are so skewed (94% of nouns take -u)

Then: predict experimental results from phonological model

Phonological frequency matching

Baseline: phonological model trained on noun tokens from the Czech National Corpus (Křen et al., 2022)

- Nouns binned as categorical ($> 99\%$ of tokens) locative -u vs. variable/categorical -ɛ
- Given *word*, predicts odds of -ɛ as coefficient *phon_odds*
- final C place + final C manner + final coda complexity + final V length + final V front + final V height + word length
- Phonology has a *much smaller effect* than in Hungarian ($R^2 = .09$)
- In part because the data are so skewed (94% of nouns take -u)

Then: predict experimental results from phonological model

- Given nonce word phonology, syntactic context, and participant, predicts odds of -ɛ
- (1 | participant) + phon_odds + preposition

Phonological frequency matching

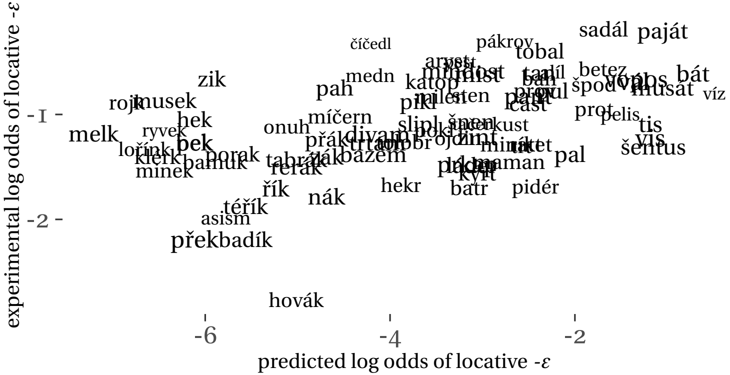
Baseline: phonological model trained on noun tokens from the Czech National Corpus (Křen et al., 2022)

- Nouns binned as categorical (> 99% of tokens) locative -u vs. variable/categorical -ɛ
- Given *word*, predicts odds of -ɛ as coefficient ***phon_odds***
- final C place + final C manner + final coda complexity + final V length + final V front + final V height + word length
- Phonology has a *much smaller effect* than in Hungarian ($R^2 = .09$)
- In part because the data are so skewed (94% of nouns take -u)

Then: predict experimental results from phonological model

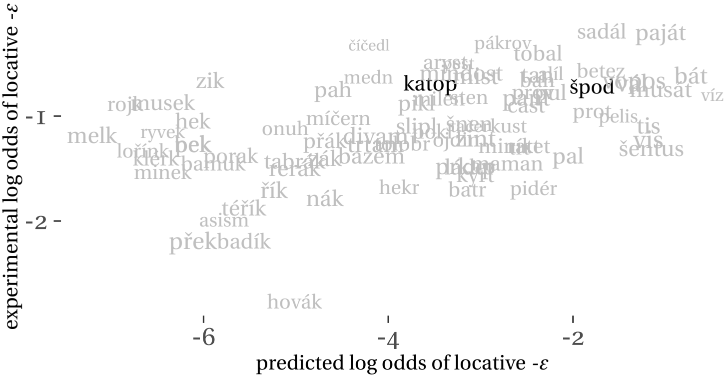
- Given nonce word phonology, syntactic context, and participant, predicts odds of -ɛ
- (1 | participant) + **phon_odds** + preposition

Results: phonological frequency matching



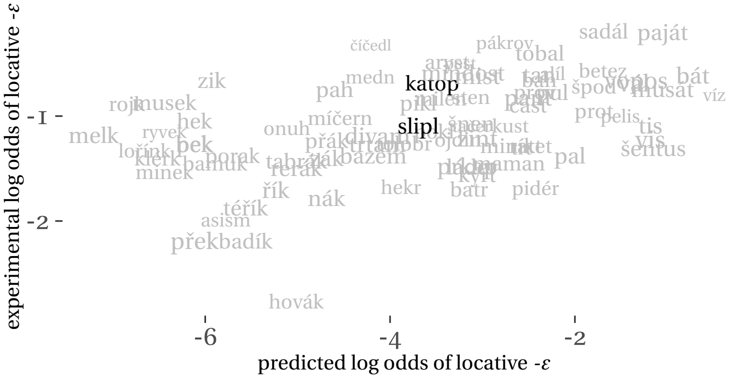
Baseline: the phonological model is slightly predictive of experimental rate of locatives for *individual nonce words*

Results: phonological frequency matching



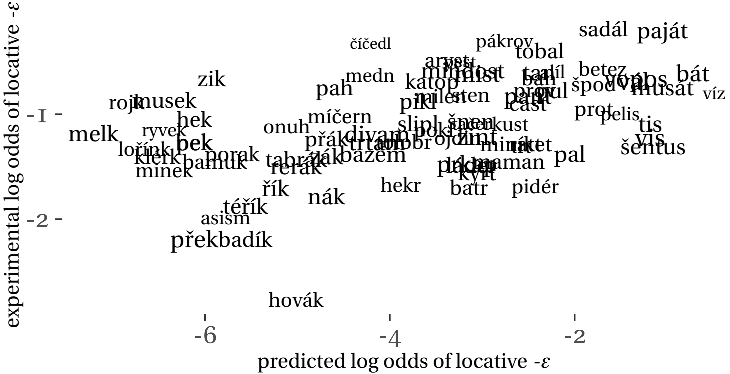
predicted: katopjɛ < ʃpoɟɛ

actual: katopjɛ = ʃpoɟɛ



predicted: katopjɛ = slɪplɛ
actual: katopjɛ > slɪplɛ

Results: phonological frequency matching



Baseline: the phonological model is slightly predictive of experimental rate of locatives for *individual nonce words*

Sensitivity to morphology

Baseline: phonological model trained on noun tokens from the Czech National Corpus (Křen et al., 2022)

- Nouns binned as categorical ($>$ 99% of tokens) locative -u vs. variable/categorical -ɛ
- Given *word*, predicts odds of -ɛ as coefficient *phon_odds*
- final C place + final C manner + final coda complexity + final V length + final V front + final V height + word length
- Phonology has a *much smaller effect* than in Hungarian ($R^2 = .09$)
- In part because the data are so skewed (94% of nouns take -u)

Then: predict experimental results from phonological model

- Given nonce word phonology, syntactic context, and participant, predicts odds of -ɛ
- (1 | participant) + phon_odds + preposition

Sensitivity to morphology

Baseline: phonological model trained on noun tokens from the Czech National Corpus (Křen et al., 2022)

- Nouns binned as categorical ($>$ 99% of tokens) locative -u vs. variable/categorical -ɛ
- Given *word*, predicts odds of -ɛ as coefficient *phon_odds*
- final C place + final C manner + final coda complexity + final V length + final V front + final V height + word length
- Phonology has a *much smaller effect* than in Hungarian ($R^2 = .09$)
- In part because the data are so skewed (94% of nouns take -u)

Then: predict experimental results from phonological model **and genitive**

- Given nonce word phonology, syntactic context, and participant, predicts odds of -ɛ
- (1 | participant) + phon_odds + preposition

Sensitivity to morphology

Baseline: phonological model trained on noun tokens from the Czech National Corpus (Křen et al., 2022)

- Nouns binned as categorical ($> 99\%$ of tokens) locative -u vs. variable/categorical -ɛ
- Given *word*, predicts odds of -ɛ as coefficient *phon_odds*
- final C place + final C manner + final coda complexity + final V length + final V front + final V height + word length
- Phonology has a *much smaller effect* than in Hungarian ($R^2 = .09$)
- In part because the data are so skewed (94% of nouns take -u)

Then: predict experimental results from phonological model **and genitive**

- Given nonce word phonology, syntactic context, **genitive**, and participant, predicts odds of -ɛ
- (1 | participant) + phon_odds + preposition

Sensitivity to morphology

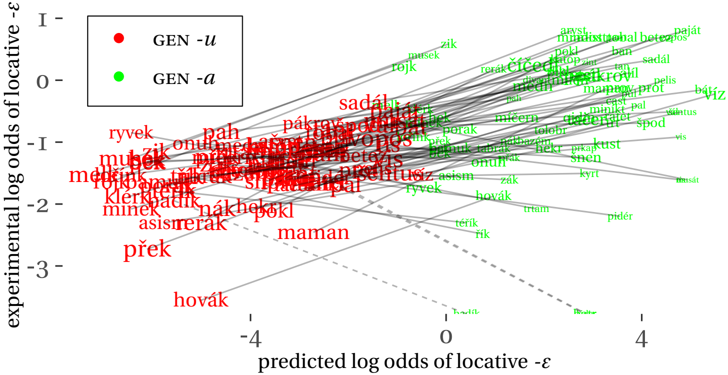
Baseline: phonological model trained on noun tokens from the Czech National Corpus (Křen et al., 2022)

- Nouns binned as categorical ($>$ 99% of tokens) locative -u vs. variable/categorical -ε
- Given *word*, predicts odds of -ε as coefficient *phon_odds*
- final C place + final C manner + final coda complexity + final V length + final V front + final V height + word length
- Phonology has a *much smaller effect* than in Hungarian ($R^2 = .09$)
- In part because the data are so skewed (94% of nouns take -u)

Then: predict experimental results from phonological model **and genitive**

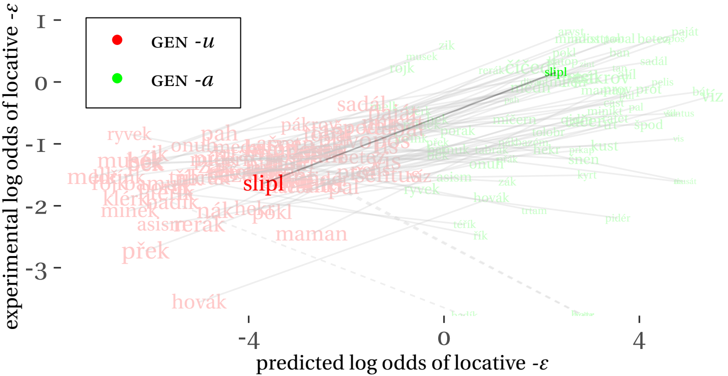
- Given nonce word phonology, syntactic context, **genitive**, and participant, predicts odds of -ε
- (1 | participant) + phon_odds + preposition **+ genitive**

Results: sensitivity to morphology



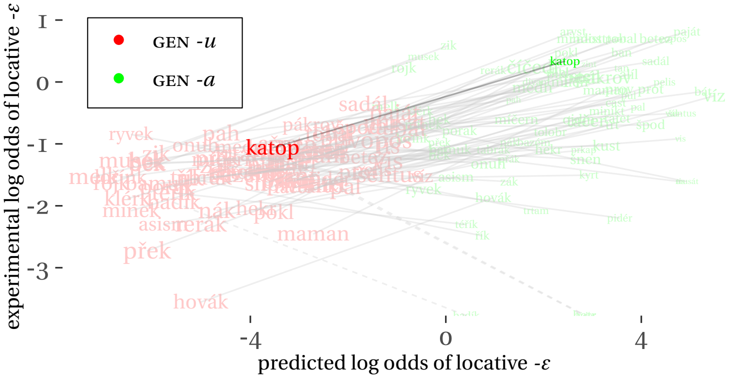
Target condition: most nonce words had a *much higher* rate of -ε when also assigned genitive as -a

Results: sensitivity to morphology



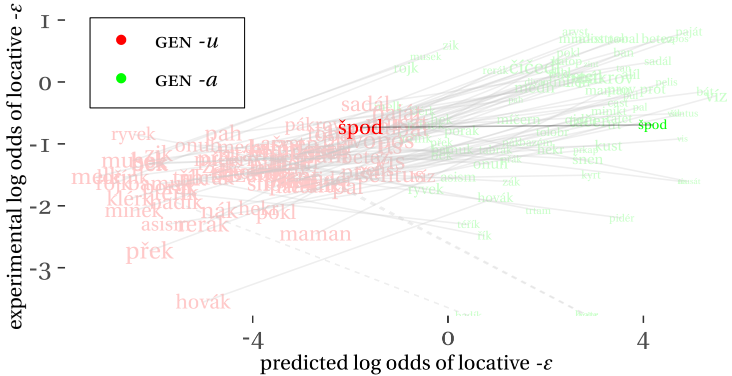
predicted: slɪplu, slɪplɛ < slɪpla, slɪplɛ
actual: slɪplu, slɪplɛ < slɪpla, slɪplɛ

Results: sensitivity to morphology



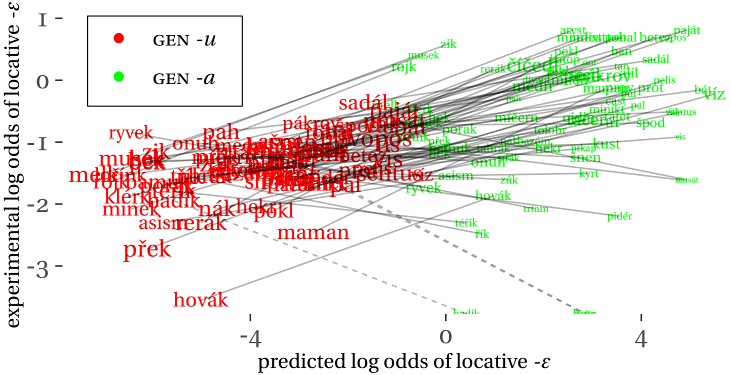
predicted: katopu, katopjɛ < katopa, katopjɛ
actual: katopu, katopjɛ < katopa, katopjɛ

Results: sensitivity to morphology



predicted: ʃpodu, ʃpoɟɛ < ʃpoda, ʃpoɟɛ
actual: ʃpodu, ʃpoɟɛ = ʃpoda, ʃpoɟɛ

Results: sensitivity to morphology



Target condition: most nonce words had a *much higher* rate of -ɛ when also assigned genitive as -a

Results: summary

- Participants (very loosely) matched the phonological distribution of -u and -ɛ in the lexicon

Results: summary

- Participants (very loosely) matched the phonological distribution of -u and -ɛ in the lexicon
- Not shown: syntactic context (preposition) also closely mirrored the lexicon

Results: summary

- Participants (very loosely) matched the phonological distribution of -u and -ɛ in the lexicon
- Not shown: syntactic context (preposition) also closely mirrored the lexicon
- They assigned -ɛ much more to nonce words with genitive -a