Czech speakers learn and apply morphological dependencies

Guy Tabachnick

University of Nova Gorica

November 23, 2023

Introduction

Two things that speakers know about their language:

Introduction

Two things that speakers know about their language:

- the behavior of *individual words*

Introduction

Two things that speakers know about their language:

- the behavior of *individual words*

present	*past*
love	loved
sing	sang
grow	grew

Introduction

Two things that speakers know about their language:

- the behavior of *individual words*
- *patterns* of word behavior

present	*past*
love	loved
sing	sang
grow	grew

Introduction

Two things that speakers know about their language:

- the behavior of *individual words*
- *patterns* of word behavior

present	*past*
love	loved
sing	sang
grow	grew

Introduction

Two things that speakers know about their language:

- the behavior of *individual words*
- *patterns* of word behavior – *how do we know they know this?*

present	*past*
love	loved
sing	sang
grow	grew

Introduction

Two things that speakers know about their language:

- the behavior of *individual words*
- *patterns* of word behavior – *how do we know they know this?*

present	*past*
love	loved
sing	sang
grow	grew
spling	

Introduction

Two things that speakers know about their language:

- ~~the behavior of *individual words*~~
- ***patterns* of word behavior** – *how do we know they know this?*

present	*past*
love	loved
sing	sang
grow	grew
spling	

Introduction

Two things that speakers know about their language:

- ~~the behavior of *individual words*~~
- ***patterns* of word behavior** – *how do we know they know this?*

present	*past*
love	loved
sing	sang
grow	grew
spling	**splang**

Introduction

Two things that speakers know about their language:

- ~~the behavior of *individual words*~~
- ***patterns* of word behavior** – *how do we know they know this?*

present	*past*
love	loved
sing	sang
grow	grew
spling	**splang**
spling	**splinged**

Introduction

Two things that speakers know about their language:

- ~~the behavior of *individual words*~~
- ***patterns* of word behavior** – *how do we know they know this?*

present	*past*
love	loved
sing	sang
grow	grew
spling	**splang**
spling	**splinged**

By asking speakers to make forms of made-up words, we can identify what patterns they have learned and use productively (Berko, 1958)

Introduction

We know that speakers learn variable *phonological* patterns and apply them variably to new words (e.g. Ernestus and Baayen, 2003; Albright and Hayes, 2003; Hayes et al., 2009; Becker et al., 2011; Gouskova et al., 2015)

Introduction

We know that speakers learn variable *phonological* patterns and apply them variably to new words (e.g. Ernestus and Baayen, 2003; Albright and Hayes, 2003; Hayes et al., 2009; Becker et al., 2011; Gouskova et al., 2015)

… But languages have other patterns too!

Introduction

We know that speakers learn variable *phonological* patterns and apply them variably to new words (e.g. Ernestus and Baayen, 2003; Albright and Hayes, 2003; Hayes et al., 2009; Becker et al., 2011; Gouskova et al., 2015)

… But languages have other patterns too!

present	*past*	*participle*
love	loved	loved
sing	sang	sung
grow	grew	grown

Introduction

We know that speakers learn variable *phonological* patterns and apply them variably to new words (e.g. Ernestus and Baayen, 2003; Albright and Hayes, 2003; Hayes et al., 2009; Becker et al., 2011; Gouskova et al., 2015)

… But languages have other patterns too! – correlations *between* a word's various behaviors

present	*past*	*participle*
love	loved	loved
sing	sang	sung
grow	grew	grown

Introduction

We know that speakers learn variable *phonological* patterns and apply them variably to new words (e.g. Ernestus and Baayen, 2003; Albright and Hayes, 2003; Hayes et al., 2009; Becker et al., 2011; Gouskova et al., 2015)

… But languages have other patterns too! – correlations *between* a word's various behaviors

present	*past*	*participle*
love	loved	loved
sing	sang	sung
grow	grew	grown
spling		**splung**
scring		**scringed**

Introduction

We know that speakers learn variable *phonological* patterns and apply them variably to new words (e.g. Ernestus and Baayen, 2003; Albright and Hayes, 2003; Hayes et al., 2009; Becker et al., 2011; Gouskova et al., 2015)

… But languages have other patterns too! – correlations *between* a word's various behaviors

present	*past*	*participle*
love	loved	loved
sing	sang	sung
grow	grew	grown
spling	**splang**	**splung**
scring		**scringed**

Introduction

We know that speakers learn variable *phonological* patterns and apply them variably to new words (e.g. Ernestus and Baayen, 2003; Albright and Hayes, 2003; Hayes et al., 2009; Becker et al., 2011; Gouskova et al., 2015)

… But languages have other patterns too! – correlations *between* a word's various behaviors

present	*past*	*participle*
love	loved	loved
sing	sang	sung
grow	grew	grown
spling	**splang**	**splung**
scring	**scringed**	**scringed**

Main question

Question: Do speakers learn variable *morphological* patterns and apply them variably to new words, just as they do *phonological* patterns?

Main question

Question: Do speakers learn variable *morphological* patterns and apply them variably to new words, just as they do *phonological* patterns?

- Correlations between forms of a word are known to be an important feature of languages with rich morphology (e.g. Wurzel, 1989; Finkel and Stump, 2007; Halle and Marantz, 2008; Ackerman et al., 2009; Ackerman and Malouf, 2013; Bonami and Beniamine, 2016; Parker and Sims, 2020)

Main question

Question: Do speakers learn variable *morphological* patterns and apply them variably to new words, just as they do *phonological* patterns?

- Correlations between forms of a word are known to be an important feature of languages with rich morphology (e.g. Wurzel, 1989; Finkel and Stump, 2007; Halle and Marantz, 2008; Ackerman et al., 2009; Ackerman and Malouf, 2013; Bonami and Beniamine, 2016; Parker and Sims, 2020)
- . . . But speakers' knowledge of these correlations hasn't been systematically tested

Main question

Question: Do speakers learn variable *morphological* patterns and apply them variably to new words, just as they do *phonological* patterns?

- Correlations between forms of a word are known to be an important feature of languages with rich morphology (e.g. Wurzel, 1989; Finkel and Stump, 2007; Halle and Marantz, 2008; Ackerman et al., 2009; Ackerman and Malouf, 2013; Bonami and Beniamine, 2016; Parker and Sims, 2020)
- . . . But speakers' knowledge of these correlations hasn't been systematically tested

Answer: They do!

Main question

Question: Do speakers learn variable *morphological* patterns and apply them variably to new words, just as they do *phonological* patterns?

- Correlations between forms of a word are known to be an important feature of languages with rich morphology (e.g. Wurzel, 1989; Finkel and Stump, 2007; Halle and Marantz, 2008; Ackerman et al., 2009; Ackerman and Malouf, 2013; Bonami and Beniamine, 2016; Parker and Sims, 2020)
- . . . But speakers' knowledge of these correlations hasn't been systematically tested

Answer: They do!

- In an experiment with Czech nouns, we see the hypothesized patterns of speaker behavior

Theoretical implications

We can model speakers' knowledge of these patterns using the same tools we have to model the phonological patterns

- I adapt the *sublexicon* model (Allen and Becker, 2015; Gouskova et al., 2015; Becker and Gouskova, 2016)

Theoretical implications

We can model speakers' knowledge of these patterns using the same tools we have to model the phonological patterns

- I adapt the *sublexicon* model (Allen and Becker, 2015; Gouskova et al., 2015; Becker and Gouskova, 2016)

This pattern-matching module encodes paradigm structure outside of the rules and procedures of the generative grammar (cf. Ackerman and Malouf, 2013)

Experiment: Czech locative and genitive

Czech nouns have seven *cases* used in different contexts

Experiment: Czech locative and genitive

Czech nouns have seven *cases* used in different contexts

- genitive: -u/-a
  - z kostɛl-a ‘out of the church’
  - do kostɛl-a ‘into the church’
  - u kostɛl-a ‘by the church’
- locative: -u/-ɛ
  - v kostɛl-ɛ ‘in the church’
  - na kostɛl-ɛ ‘on the church’
  - o kostɛl-ɛ ‘about the church’

Experiment: Czech locative and genitive

Czech nouns have seven *cases* used in different contexts

- genitive: -u/-a
  - z kostɛl-a ‘out of the church’
  - do kostɛl-a ‘into the church’
  - u kostɛl-a ‘by the church’
- locative: -u/-ɛ
  - v kostɛl-ɛ ‘in the church’
  - na kostɛl-ɛ ‘on the church’
  - o kostɛl-ɛ ‘about the church’

Words appear with all possible pairings of genitive and locative suffixes

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: Czech locative and genitive

Czech nouns have seven *cases* used in different contexts

- genitive: -u/-a
  - z kostɛl-a ‘out of the church’
  - do kostɛl-a ‘into the church’
  - u kostɛl-a ‘by the church’
- locative: -u/-ɛ
  - v kostɛl-ɛ ‘in the church’
  - na kostɛl-ɛ ‘on the church’
  - o kostɛl-ɛ ‘about the church’

Words appear with all possible pairings of genitive and locative suffixes

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: Czech locative and genitive

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

Experiment: Czech locative and genitive

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

- Today -u is much more common
- In particular: -ɛ triggers alternation of dorsals [k x ɦ]: [jazɪk] 'language', [v jazɪts-ɛ] 'in language' (which taking -u avoids)
- Words with the older genitive (-a) often retain the older locative (-ɛ) as well

Experiment: Czech locative and genitive

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

- Today -u is much more common
- In particular: -ɛ triggers alternation of dorsals [k x ɦ]: [jazɪk] 'language', [v jazɪts-ɛ] 'in language' (which taking -u avoids)
- Words with the older genitive (-a) often retain the older locative (-ɛ) as well

		locative			
		-u	-u∼-ɛ	-ɛ	% -u
genitive	-u	9686	523	21	94.7%
	-u∼-a	145	18	3	87.3%
	-a	32	18	31	39.5%
	% -u	98.2%	93.6%	38.1%	

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

Basic results

Speakers chose locative -ɛ more often when paired with genitive -a – they have learned the correlation between them!

		locative		
		-u	-ɛ	% -u
genitive	-u	2532	672	79.0%
	-a	667	426	61.0%
	% -u	79.1%	61.2%	

Phonological frequency matching

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

Phonological frequency matching

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

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

Phonological frequency matching

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

- final C place + final C manner + final coda complexity + final V length + final V front + final V height + word length
- Given *word*, predicts odds of -ɛ as coefficient *phon_odds*
- Phonology is only slightly predictive of locative suffix ($R^2 = .09$)
- In part because the data are so skewed (94% of nouns take -u)

Phonological frequency matching

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

- final C place + final C manner + final coda complexity + final V length + final V front + final V height + word length
- Given *word*, predicts odds of -ɛ as coefficient *phon_odds*
- Phonology is only slightly predictive of locative suffix ($R^2 = .09$)
- 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 types from the Czech National Corpus (Křen et al., 2022)

- final C place + final C manner + final coda complexity + final V length + final V front + final V height + word length
- Given *word*, predicts odds of -ɛ as coefficient *phon_odds*
- Phonology is only slightly predictive of locative suffix ($R^2 = .09$)
- 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 + phon_odds + preposition | participant) + (1 | word) + phon_odds + preposition

Phonological frequency matching

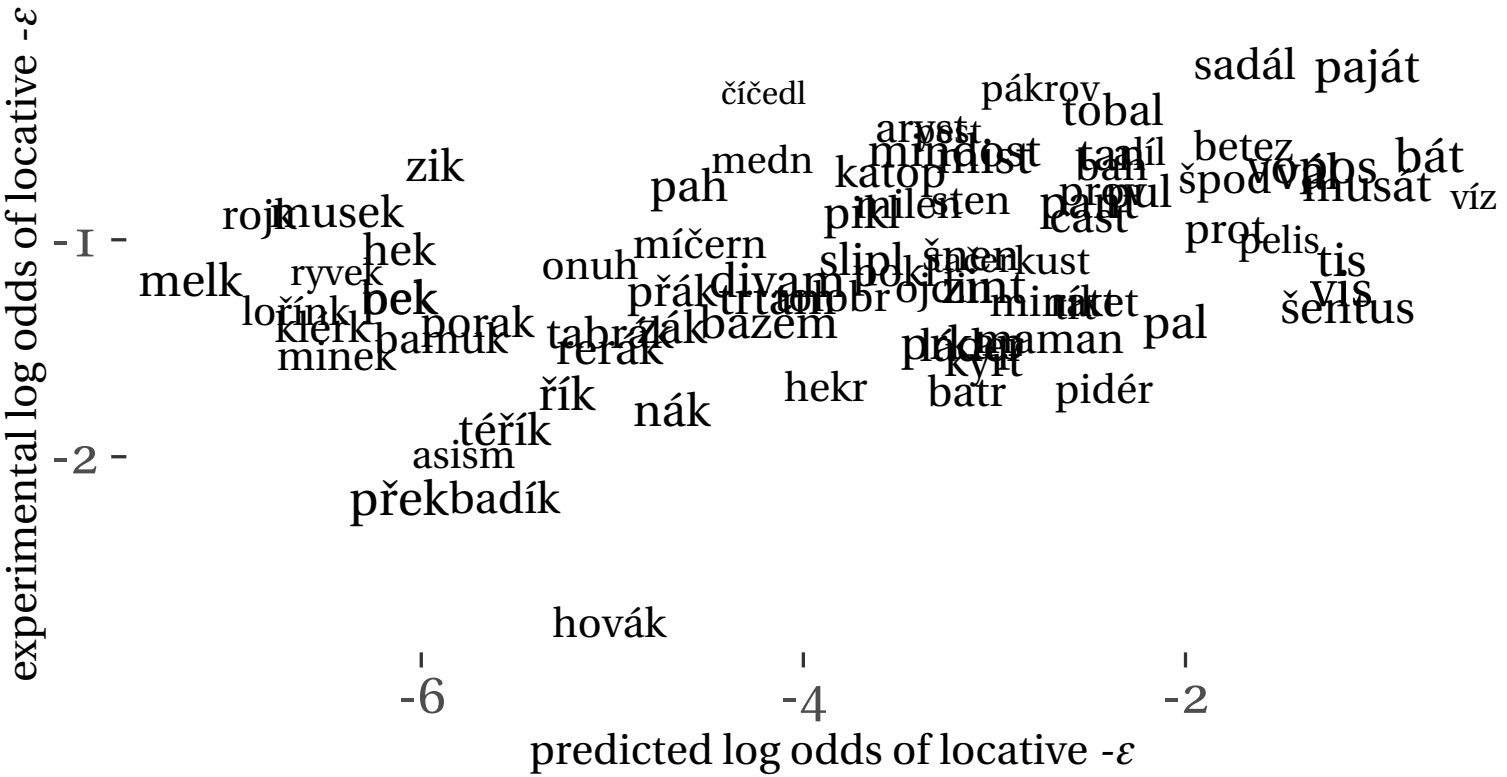
Baseline: phonological model trained on noun types from the Czech National Corpus (Křen et al., 2022)

- final C place + final C manner + final coda complexity + final V length + final V front + final V height + word length
- Given *word*, predicts odds of -ɛ as coefficient ***phon_odds***
- Phonology is only slightly predictive of locative suffix ($R^2 = .09$)
- In part because the data are so skewed (94% of nouns take -u)

Then: predict experimental results from phonological model

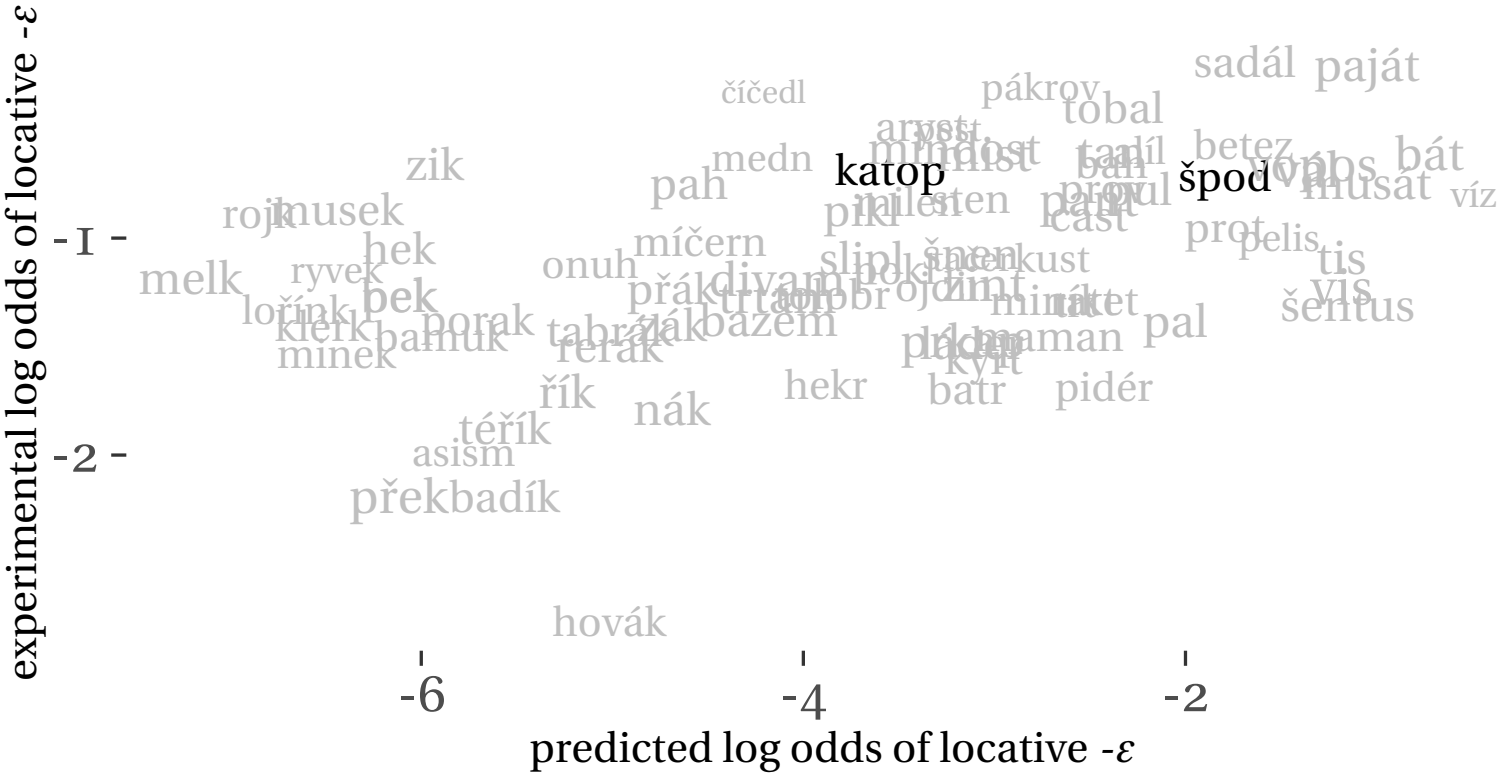
- Given nonce word phonology, syntactic context, and participant, predicts odds of -ɛ
- (1 + **phon_odds** + preposition | participant) + (1 | word) + **phon_odds** + preposition

Results: phonological frequency matching



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

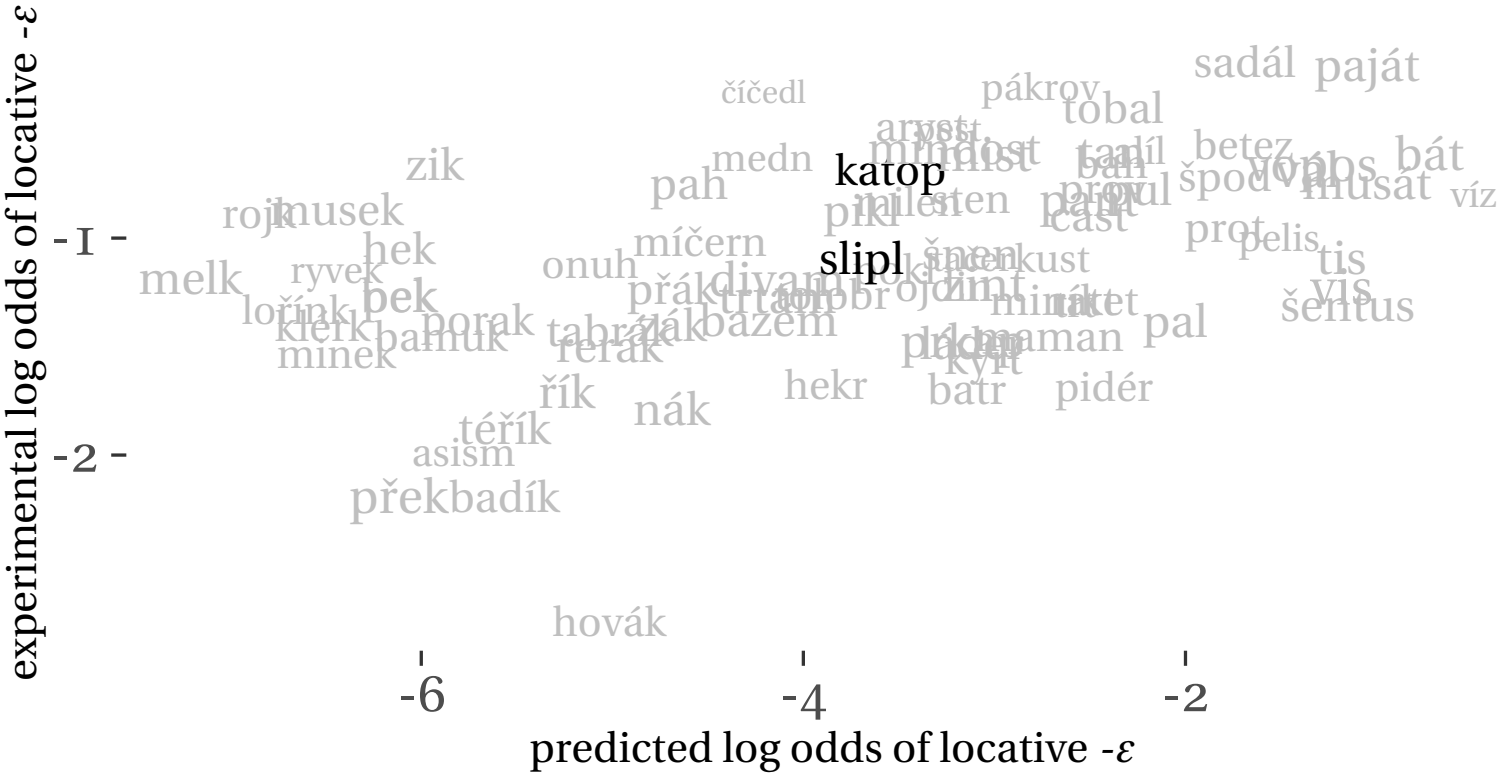
Results: phonological frequency matching



predicted: katopjɛ < ʃpoɟɛ

actual: katopjɛ = ʃpoɟɛ

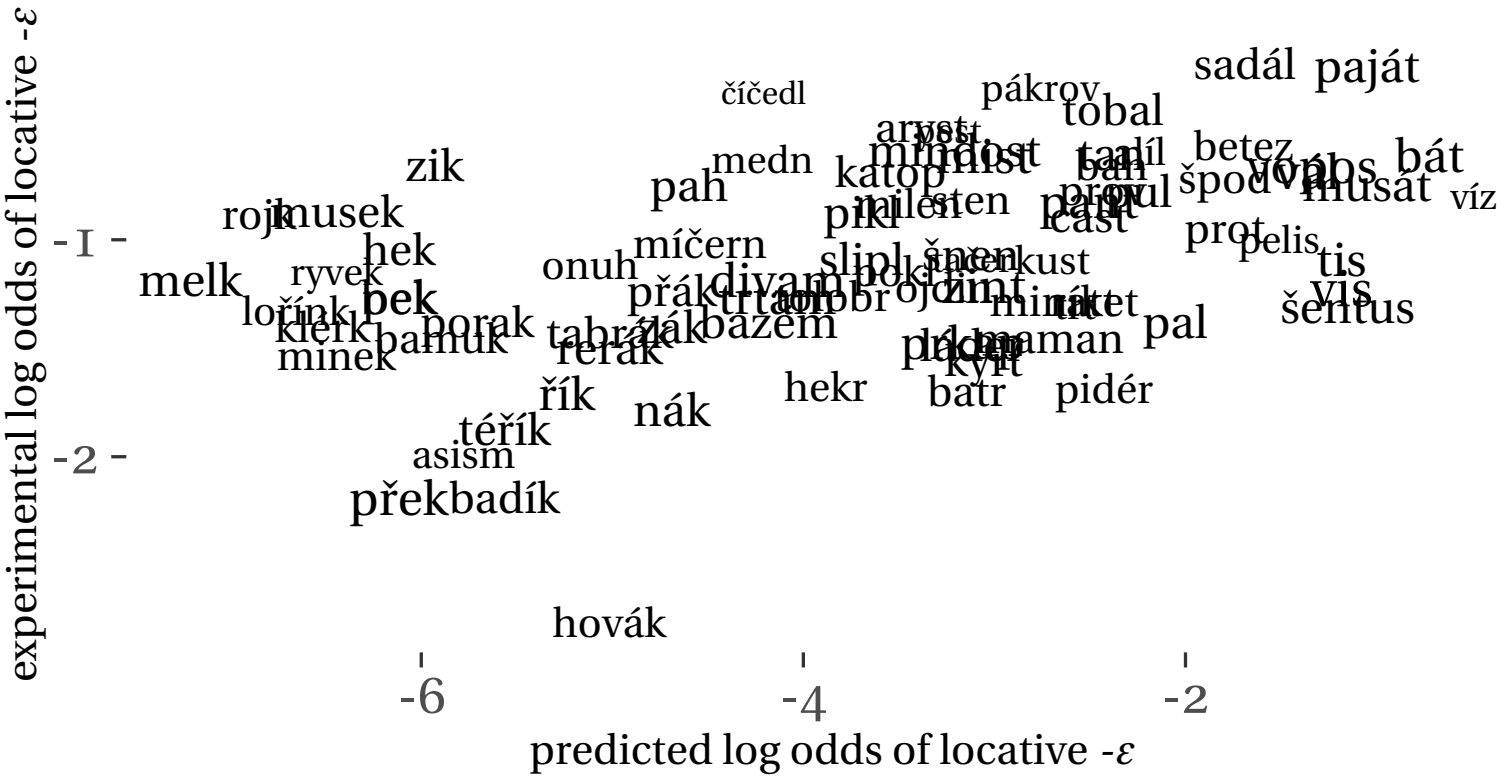
Results: phonological frequency matching



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)

- final C place + final C manner + final coda complexity + final V length + final V front + final V height + word length
- Given *word*, predicts odds of -ɛ as coefficient *phon_odds*
- Phonology is only slightly predictive of locative suffix ($R^2 = .09$)
- 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 + phon_odds + preposition | participant) + (1 | word) + phon_odds + preposition

Sensitivity to morphology

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

- final C place + final C manner + final coda complexity + final V length + final V front + final V height + word length
- Given *word*, predicts odds of -ɛ as coefficient *phon_odds*
- Phonology is only slightly predictive of locative suffix ($R^2 = .09$)
- 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 + phon_odds + preposition | participant) + (1 | word) + phon_odds + preposition

Sensitivity to morphology

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

- final C place + final C manner + final coda complexity + final V length + final V front + final V height + word length
- Given *word*, predicts odds of -ɛ as coefficient *phon_odds*
- Phonology is only slightly predictive of locative suffix ($R^2 = .09$)
- 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 + phon_odds + preposition | participant) + (1 | word) + phon_odds + preposition

Sensitivity to morphology

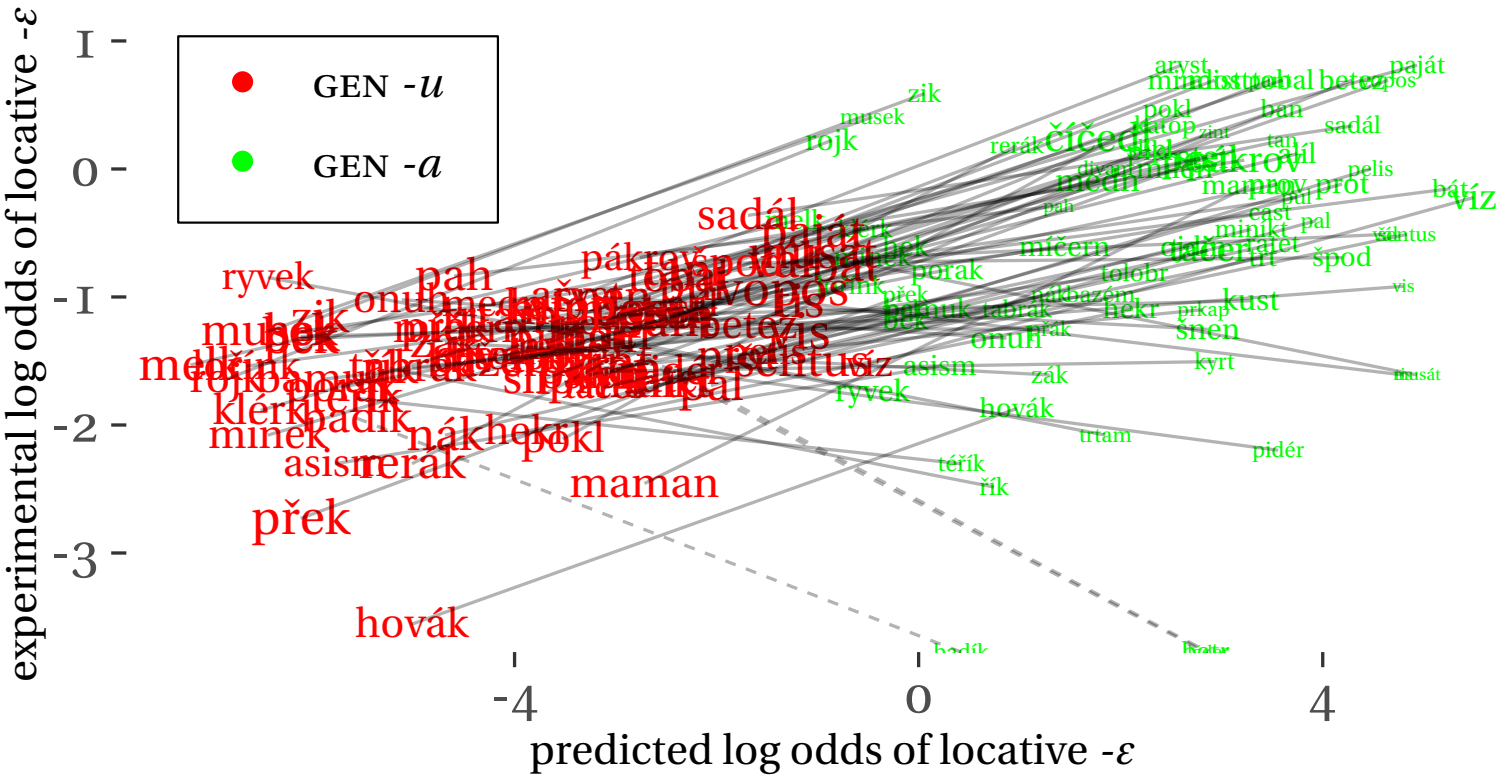
Baseline: phonological model trained on noun tokens from the Czech National Corpus (Křen et al., 2022)

- final C place + final C manner + final coda complexity + final V length + final V front + final V height + word length
- Given *word*, predicts odds of -ɛ as coefficient *phon_odds*
- Phonology is only slightly predictive of locative suffix ($R^2 = .09$)
- In part because the data are so skewed (94% of nouns take -u)

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

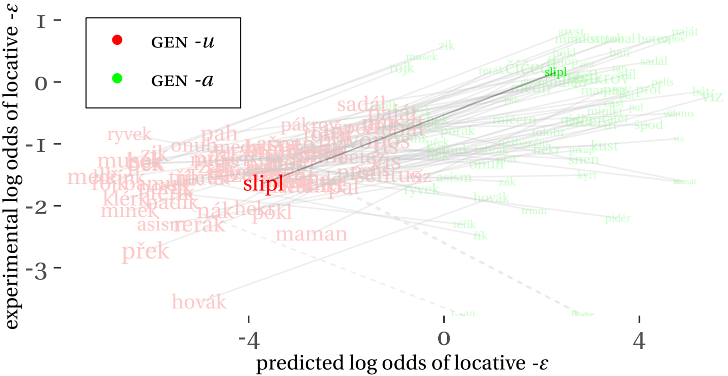
- Given nonce word phonology, syntactic context, **genitive**, and participant, predicts odds of -ɛ
- (1 + phon_odds + preposition **+ genitive** | participant) + (1 | word) + phon_odds + preposition **+ genitive**

Results: sensitivity to morphology



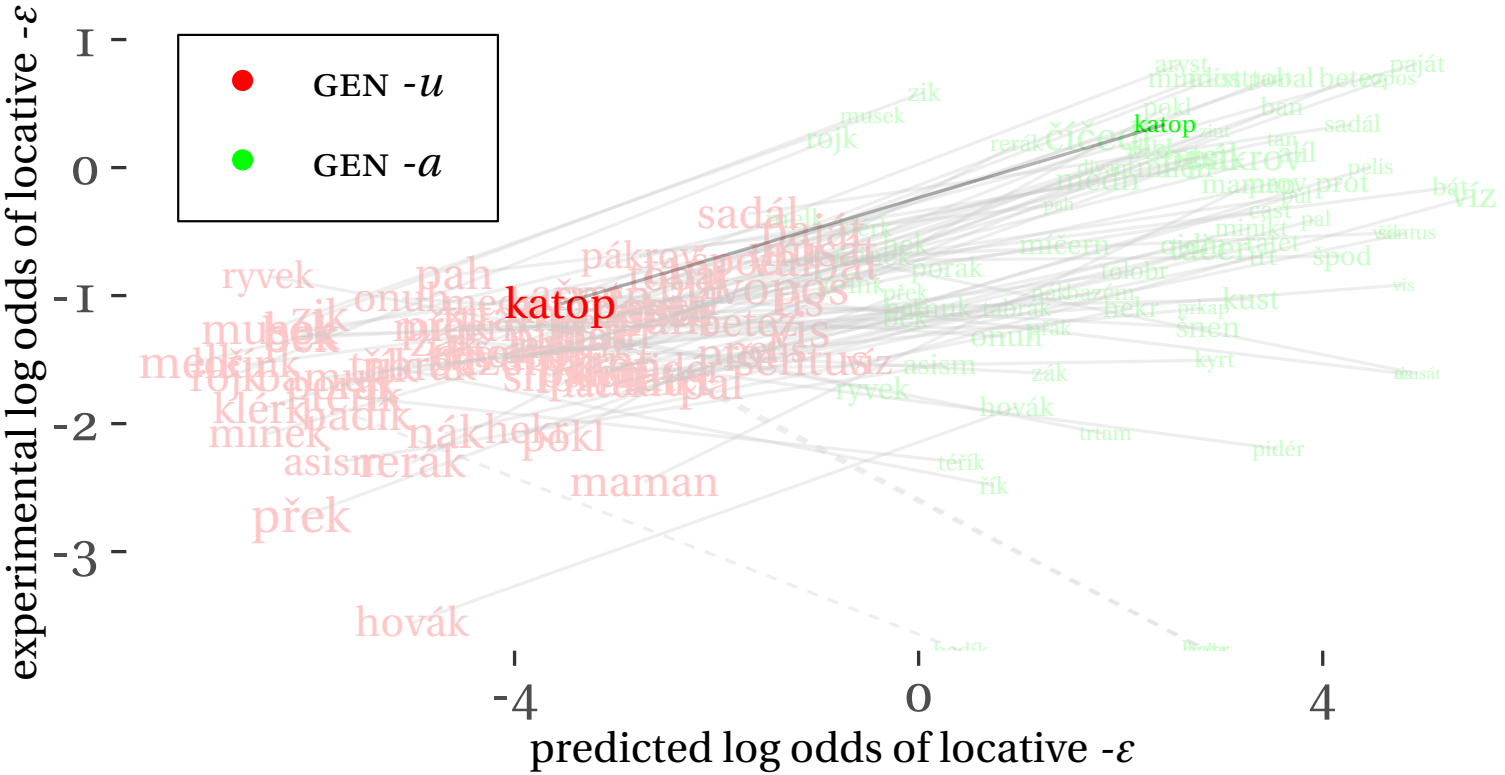
Target condition: most nonce words had a *much higher* rate of -ε when also assigned genitive as -a

Results: sensitivity to morphology



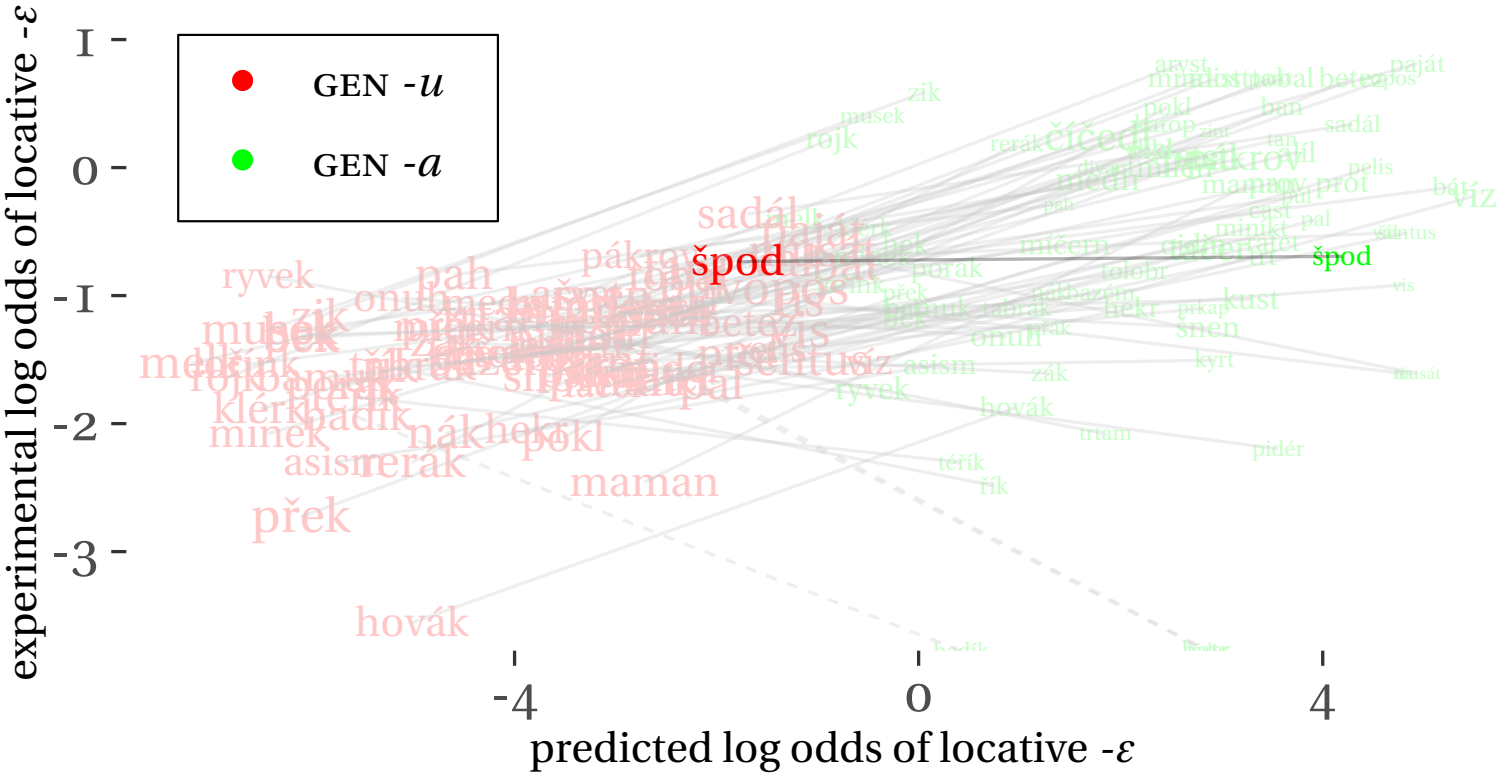
predicted: slıplu, slıplɛ < slıpla, slıplɛ
actual: slıplu, slıplɛ < slıpla, slıplɛ

Results: sensitivity to morphology



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

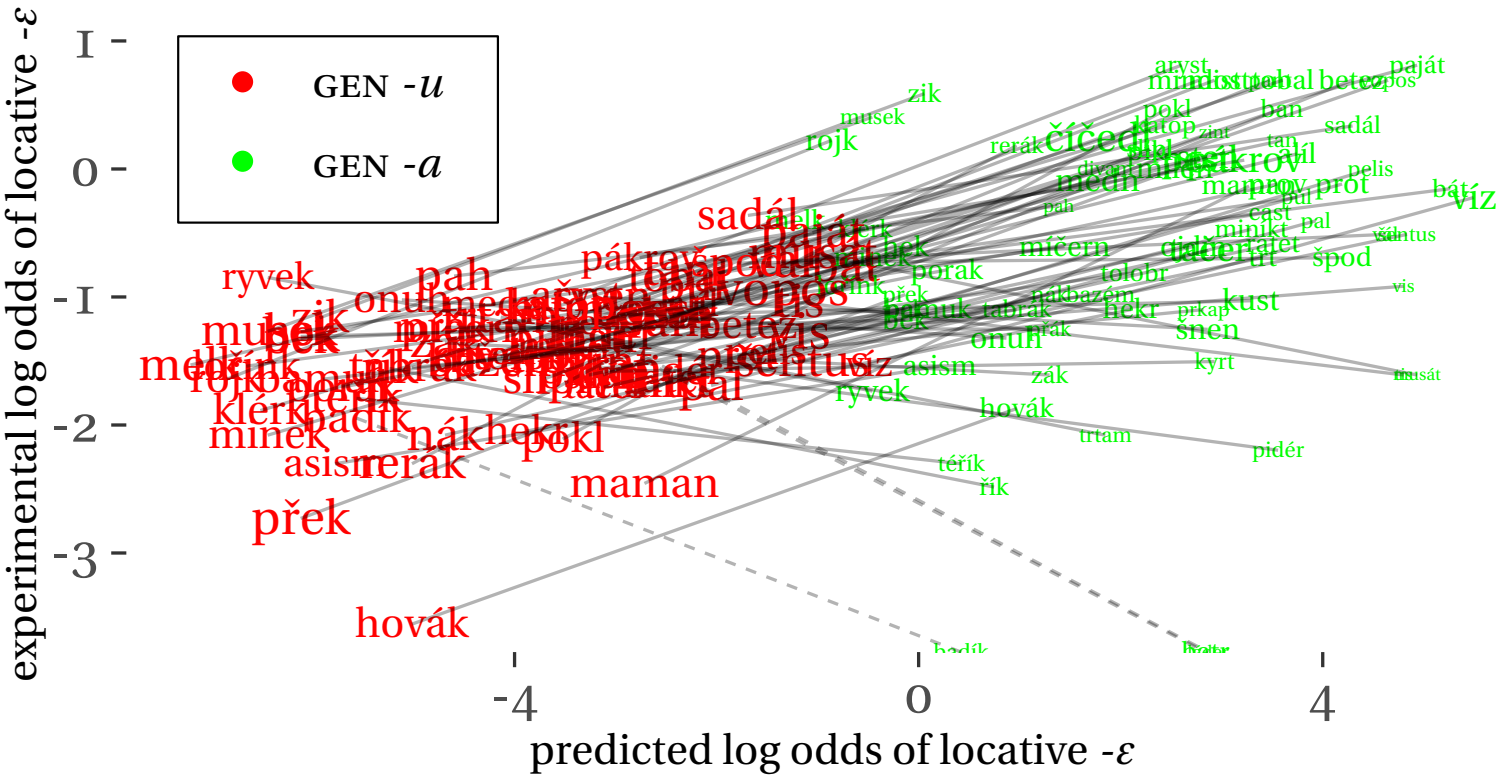
Results: sensitivity to morphology



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 -ε 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
- They assigned -ɛ more to nonce words with genitive -a

Results: interpretation

Interpretation of results:

- Speakers have *learned* a correlation between genitive -a and locative -ɛ from their lexicon and *apply* it productively for novel locatives

Results: interpretation

Interpretation of results:

1. Speakers have *learned* a correlation between genitive -a and locative -ɛ from their lexicon and *apply* it productively for novel locatives
2. Speakers are subject to a priming effect of genitive -u → locative -u

Results: interpretation

Interpretation of results:

1. Speakers have *learned* a correlation between genitive -a and locative -ɛ from their lexicon and *apply* it productively for novel locatives
2. Speakers are subject to a priming effect of genitive -u → locative -u

The two interpretations differ in their predictions on the same task applied to *real* words that allow both variants:

1. No effect for real words, which already have stored locative behavior and do not require productive generation using analogy
2. Similar effect for real words, which show the same surface allomorphs

Experiment: real words

Stimuli: words variable in the genitive and (usually) the locative

		genitive		locative	
	noun	tokens	% -u	tokens	% -u
komiːn	‘chimney’	13 992	18.2%	8965	13.1%
bɛtlɛːm	‘nativity scene’	4150	51.0%	2783	74.7%
siːr	‘cheese’	2365	21.6%	1027	100.0%

Experiment: real words

Stimuli: words variable in the genitive and (usually) the locative

		genitive		locative	
	noun	tokens	% -u	tokens	% -u
komiːn	‘chimney’	13 992	18.2%	8965	13.1%
bɛtlɛːm	‘nativity scene’	4150	51.0%	2783	74.7%
siːr	‘cheese’	2365	21.6%	1027	100.0%

Results:

Experiment: real words

Stimuli: words variable in the genitive and (usually) the locative

		genitive		locative	
	noun	tokens	% -u	tokens	% -u
komiːn	‘chimney’	13 992	18.2%	8965	13.1%
bɛtlɛːm	‘nativity scene’	4150	51.0%	2783	74.7%
siːr	‘cheese’	2365	21.6%	1027	100.0%

Results:

- priming effect in the genitive: komiːn-u → komiːn-u

Experiment: real words

Stimuli: words variable in the genitive and (usually) the locative

		genitive		locative	
	noun	tokens	% -u	tokens	% -u
komiːn	‘chimney’	13 992	18.2%	8965	13.1%
bɛtlɛːm	‘nativity scene’	4150	51.0%	2783	74.7%
siːr	‘cheese’	2365	21.6%	1027	100.0%

Results:

- priming effect in the genitive: komiːn-u → komiːn-u
- no priming effect in the locative: komiːn-u ↛ komiːn-u

Experiment: real words

Stimuli: words variable in the genitive and (usually) the locative

		genitive		locative	
	noun	tokens	% -u	tokens	% -u
komiːn	‘chimney’	13 992	18.2%	8965	13.1%
bɛtlɛːm	‘nativity scene’	4150	51.0%	2783	74.7%
siːr	‘cheese’	2365	21.6%	1027	100.0%

Results:

- priming effect in the genitive: komiːn-u → komiːn-u
- no priming effect in the locative: komiːn-u ↛ komiːn-u

The genitive–locative correlation found in the previous study really is a cooccurrence relation learned from the lexicon!

Accounting for what speakers know

Two things that speakers know about their language:

Accounting for what speakers know

Two things that speakers know about their language:

- the behavior of *individual words*

Accounting for what speakers know

Two things that speakers know about their language:

- the behavior of *individual words*
  - words are stored with symbolic *diacritic features* that index their behavior (e.g. Lightner, 1965; Lieber, 1980; Corbett and Baerman, 2006)

Accounting for what speakers know

Two things that speakers know about their language:

- the behavior of *individual words*
  - words are stored with symbolic *diacritic features* that index their behavior (e.g. Lightner, 1965; Lieber, 1980; Corbett and Baerman, 2006)

    'problem' /problɛːm$_{[\text{GEN: u, LOC: u}]}$/
    'church' /kostɛl$_{[\text{GEN: a, LOC: ɛ}]}$/

Accounting for what speakers know

Two things that speakers know about their language:

- the behavior of *individual words*
  - words are stored with symbolic *diacritic features* that index their behavior (e.g. Lightner, 1965; Lieber, 1980; Corbett and Baerman, 2006)

    'problem' /problɛːm$_{[\text{GEN: u, LOC: u}]}$/
    'church' /kostɛl$_{[\text{GEN: a, LOC: ɛ}]}$/
  - these features are implicated in *rules of realization* (spellout)

Accounting for what speakers know

Two things that speakers know about their language:

- the behavior of *individual words*
  - words are stored with symbolic *diacritic features* that index their behavior (e.g. Lightner, 1965; Lieber, 1980; Corbett and Baerman, 2006)

    'problem' /problɛːm$_{[\text{GEN: u, LOC: u}]}$/
    'church' /kostɛl$_{[\text{GEN: a, LOC: ɛ}]}$/

  - these features are implicated in *rules of realization* (spellout)

    GEN ↔ u / [GEN: u] ___ LOC ↔ u / [LOC: u]___
    GEN ↔ a / [GEN: a] ___ LOC ↔ ɛ / [LOC: ɛ] ___

Accounting for what speakers know

Two things that speakers know about their language:

- the behavior of *individual words*
  - words are stored with symbolic *diacritic features* that index their behavior (e.g. Lightner, 1965; Lieber, 1980; Corbett and Baerman, 2006)

    'problem' /problɛːm$_{[\text{GEN: u, LOC: u}]}$/
    'church' /kostɛl$_{[\text{GEN: a, LOC: ɛ}]}$/

  - these features are implicated in *rules of realization* (spellout)

    GEN ↔ u / [GEN: u] ___ LOC ↔ u / [LOC: u]___
    GEN ↔ a / [GEN: a] ___ LOC ↔ ɛ / [LOC: ɛ] ___

- *patterns* of word behavior

Accounting for what speakers know

Two things that speakers know about their language:

- the behavior of *individual words*
  - words are stored with symbolic *diacritic features* that index their behavior (e.g. Lightner, 1965; Lieber, 1980; Corbett and Baerman, 2006)

    'problem' /problɛːm$_{[\text{GEN: u, LOC: u}]}$/
    'church' /kostɛl$_{[\text{GEN: a, LOC: ɛ}]}$/

  - these features are implicated in *rules of realization* (spellout)

    GEN ↔ u / [GEN: u] ___ LOC ↔ u / [LOC: u]___
    GEN ↔ a / [GEN: a] ___ LOC ↔ ɛ / [LOC: ɛ] ___

- *patterns* of word behavior
  - speakers store (gradient and categorical) generalizations over words that share a feature as weighted constraints (Allen and Becker, 2015; Gouskova et al., 2015; Becker and Gouskova, 2016)

Accounting for what speakers know

Two things that speakers know about their language:

- the behavior of *individual words*
  - words are stored with symbolic *diacritic features* that index their behavior (e.g. Lightner, 1965; Lieber, 1980; Corbett and Baerman, 2006)

    'problem' /problɛːm$_{[\text{GEN: u, LOC: u}]}$/
    'church' /kostɛl$_{[\text{GEN: a, LOC: ɛ}]}$/

  - these features are implicated in *rules of realization* (spellout)

    GEN ↔ u / [GEN: u] ___ LOC ↔ u / [LOC: u]___
    GEN ↔ a / [GEN: a] ___ LOC ↔ ɛ / [LOC: ɛ] ___

- *patterns* of word behavior
  - speakers store (gradient and categorical) generalizations over words that share a feature as weighted constraints (Allen and Becker, 2015; Gouskova et al., 2015; Becker and Gouskova, 2016)

    [LOC: ɛ]: ending in dorsal [k x ɦ] dispreferred *[dorsal]# (2)
    [LOC: u]: genitive -a dispreferred *[GEN: a] (4)

Accounting for what speakers do

Known words already have a locative feature, so the derivation proceeds without an issue:

Accounting for what speakers do

Known words already have a locative feature, so the derivation proceeds without an issue:

- underlying form: /kostɛl$_{[\text{GEN: a, LOC: ɛ}]}$/
- applied rule: LOC $\leftrightarrow$ ɛ / [LOC: ɛ]___
- output form: [kostɛl-ɛ]

Accounting for what speakers do

Known words already have a locative feature, so the derivation proceeds without an issue:

- underlying form: /kostɛl$_{[\text{GEN: a, LOC: ɛ}]}$/
- applied rule: LOC $\leftrightarrow$ ɛ / [LOC: ɛ]___
- output form: [kostɛl-ɛ]

Novel words have no locative feature, so one needs to be added:

Accounting for what speakers do

Known words already have a locative feature, so the derivation proceeds without an issue:

- underlying form: /kostɛl$_{[\text{GEN: a, LOC: ɛ}]}$/
- applied rule: LOC ↔ ɛ / [LOC: ɛ]___
- output form: [kostɛl-ɛ]

Novel words have no locative feature, so one needs to be added:

- underlying form: /zɪk/
- applied rule: ???
- output form: ???

Accounting for what speakers do

Known words already have a locative feature, so the derivation proceeds without an issue:

- underlying form: /kostɛl$_{[\text{GEN: a, LOC: ɛ}]}$/
- applied rule: LOC ↔ ɛ / [LOC: ɛ]___
- output form: [kostɛl-ɛ]

Novel words have no locative feature, so one needs to be added:

- underlying form: /zɪk$_{[\text{LOC: u}]}$/
- applied rule: LOC ↔ u / [LOC: u]___
- output form: [zɪk-u]

Accounting for what speakers do

Known words already have a locative feature, so the derivation proceeds without an issue:

- underlying form: /kostɛl$_{[\text{GEN: a, LOC: ɛ}]}$/
- applied rule: LOC ↔ ɛ / [LOC: ɛ]___
- output form: [kostɛl-ɛ]

Novel words have no locative feature, so one needs to be added:

- underlying form: /zɪk$_{[\text{LOC: ɛ}]}$/
- applied rule: LOC ↔ ɛ / [LOC: ɛ]___
- output form: [zɪts-ɛ]

Accounting for what speakers do

Known words already have a locative feature, so the derivation proceeds without an issue:

- underlying form: /kostɛl$_{[\text{GEN: a, LOC: ɛ}]}$/
- applied rule: LOC ↔ ɛ / [LOC: ɛ]___
- output form: [kostɛl-ɛ]

Novel words have no locative feature, so one needs to be added:

- underlying form: /zɪk/
- applied rule: ???
- output form: ???

Speakers must have a way of *assigning* features to lexical entries when needed

Accounting for what speakers do

We can use the stored generalizations (weighted constraints) to model feature assignment (and thus speakers' performance on the nonce word task)

Accounting for what speakers do

We can use the stored generalizations (weighted constraints) to model feature assignment (and thus speakers' performance on the nonce word task)

- See a nonce word
  - [zɪk], genitive [zɪk-a] → /zɪk$_{[\text{GEN: a}]}$/

Accounting for what speakers do

We can use the stored generalizations (weighted constraints) to model feature assignment (and thus speakers' performance on the nonce word task)

- See a nonce word
  - [zɪk], genitive [zɪk-a] → /zɪk$_{[\text{GEN: a}]}$/
- Evaluate against constraints in the two featural grammars to produce scores s
  - [LOC: ɛ] grammar: *[dorsal]# → $s([\text{LOC: ɛ}]) = -2$
  - [LOC: u] grammar: *[GEN: a] → $s([\text{LOC: u}]) = -4$

Accounting for what speakers do

We can use the stored generalizations (weighted constraints) to model feature assignment (and thus speakers' performance on the nonce word task)

- See a nonce word
  - [zɪk], genitive [zɪk-a] $\rightarrow$ /zɪk$_{[\text{GEN: a}]}$/
- Evaluate against constraints in the two featural grammars to produce scores s
  - [LOC: ɛ] grammar: *[dorsal]# $\rightarrow s([\text{LOC: ɛ}]) = -2$
  - [LOC: u] grammar: *[GEN: a] $\rightarrow s([\text{LOC: u}]) = -4$
- Assign a feature randomly based on scores using *maximum entropy* (Goldwater and Johnson, 2003; Hayes and Wilson, 2008)
  - $P([\text{LOC: ɛ}]) \propto e^{s([\text{LOC: ɛ}])} = 88.1\%$
  - $P([\text{LOC: u}]) \propto e^{s([\text{LOC: u}])} = 11.9\%$

Accounting for what speakers do

We can use the stored generalizations (weighted constraints) to model feature assignment (and thus speakers' performance on the nonce word task)

- See a nonce word
  - [zɪk], genitive [zɪk-a] → /zɪk$_{[\textsc{gen}:\ a]}$/
- Evaluate against constraints in the two featural grammars to produce scores s
  - [loc: ɛ] grammar: *[dorsal]# → $s([\textsc{loc}:\ \varepsilon]) = -2$
  - [loc: u] grammar: *[gen: a] → $s([\textsc{loc}:\ u]) = -4$
- Assign a feature randomly based on scores using *maximum entropy* (Goldwater and Johnson, 2003; Hayes and Wilson, 2008)
  - $P([\textsc{loc}:\ \varepsilon]) \propto e^{s([\textsc{loc}:\ \varepsilon])} = \mathbf{88.1}\% \rightarrow$ /zɪk$_{[\textsc{gen}:\ a,\ \textsc{loc}:\ \varepsilon]}$/
  - $P([\textsc{loc}:\ u]) \propto e^{s([\textsc{loc}:\ u])} = 11.9\%$

Accounting for what speakers do

We can use the stored generalizations (weighted constraints) to model feature assignment (and thus speakers' performance on the nonce word task)

- See a nonce word
  - [zɪk], genitive [zɪk-a] → /zɪk$_{[\text{GEN: a}]}$/
- Evaluate against constraints in the two featural grammars to produce scores s
  - [LOC: ɛ] grammar: *[dorsal]# → $s([\text{LOC: ɛ}]) = -2$
  - [LOC: u] grammar: *[GEN: a] → $s([\text{LOC: u}]) = -4$
- Assign a feature randomly based on scores using *maximum entropy* (Goldwater and Johnson, 2003; Hayes and Wilson, 2008)
  - $P([\text{LOC: ɛ}]) \propto e^{s([\text{LOC: ɛ}])} = \mathbf{88.1}\%$ → /zɪk$_{[\text{GEN: a, LOC: ɛ}]}$/
  - $P([\text{LOC: u}]) \propto e^{s([\text{LOC: u}])} = 11.9\%$
- Produce new form
  - /zɪk$_{[\text{GEN: a, LOC: ɛ}]}$/ → locative [zɪts-ɛ]

Summary

- Speakers learn and apply variable *morphological* patterns (correlations between two behaviors) just as they do variable *phonological* patterns (correlations between sounds and behavior)
- Our existing tools to account for the phonological patterns can easily handle morphological patterns as well
- My experiments provide a new tool for systematically studying the intersecting patterns, giving us a better understanding of what people know about language and how they use it

Coda: Paradigms in syntactic theories of morphology

Ackerman and Malouf (2013): theories of morphology that build up words from constituent pieces (morphemes), like Distributed Morphology leave certain questions unanswered, or even unaskable:

Coda: Paradigms in syntactic theories of morphology

Ackerman and Malouf (2013): theories of morphology that build up words from constituent pieces (morphemes), like Distributed Morphology leave certain questions unanswered, or even unaskable:

- typical questions:
  - what are the relations between the constituent parts of a word?
  - what are the relations between the constituent parts of a word and the abstract morphosyntactic structure they spell out?

Coda: Paradigms in syntactic theories of morphology

Ackerman and Malouf (2013): theories of morphology that build up words from constituent pieces (morphemes), like Distributed Morphology leave certain questions unanswered, or even unaskable:

- typical questions:
  - what are the relations between the constituent parts of a word?
  - what are the relations between the constituent parts of a word and the abstract morphosyntactic structure they spell out?
- atypical questions:
  - what are the relations between words built off of the same stem?

Coda: Paradigms in syntactic theories of morphology

Ackerman and Malouf (2013): theories of morphology that build up words from constituent pieces (morphemes), like Distributed Morphology leave certain questions unanswered, or even unaskable:

- typical questions:
  - what are the relations between the constituent parts of a word?
  - what are the relations between the constituent parts of a word and the abstract morphosyntactic structure they spell out?
- atypical questions:
  - what are the relations between words built off of the same stem?

The question of *paradigm structure* raised in this work has been almost entirely ignored by work in Distributed Morphology and related theories (but see Halle and Marantz, 2008)

Coda: Paradigms in syntactic theories of morphology

My proposal shows that piece-based theories of morphology can accommodate paradigmatic structure:

Coda: Paradigms in syntactic theories of morphology

My proposal shows that piece-based theories of morphology can accommodate paradigmatic structure:

- generative grammar builds up words from constituent parts
- pattern-matching grammars store generalizations over morpheme structure and deploy them productively when needed

Coda: Paradigms in syntactic theories of morphology

My proposal shows that piece-based theories of morphology can accommodate paradigmatic structure:

- generative grammar builds up words from constituent parts
- pattern-matching grammars store generalizations over morpheme structure and deploy them productively when needed

This two-pronged approach has a major advantage: it is *explicit* about the units of morphology and paradigm structure

Coda: Paradigms in syntactic theories of morphology

My proposal shows that piece-based theories of morphology can accommodate paradigmatic structure:

- generative grammar builds up words from constituent parts
- pattern-matching grammars store generalizations over morpheme structure and deploy them productively when needed

This two-pronged approach has a major advantage: it is *explicit* about the units of morphology and paradigm structure

- associations between words and the patterns they follow are indexed by diacritic features on lexical entries
- the content of these features is determined by their use in the grammar (e.g. providing the context for rules of realization)
- morphological dependencies are learned as cooccurrence relations between the diacritic 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.

Allen, B. and Becker, M. (2015). Learning alternations from surface forms with sublexical phonology. lingbuzz/002503.

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

Becker, M., Ketrez, N., and Nevins, A. (2011). The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish laryngeal alternations. *Language*, 87(1):84–125.

Berko, J. (1958). The child's learning of English morphology. *Word*, 14(2–3):150–177.

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

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

Ernestus, M. and Baayen, R. H. (2003). Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language*, 79(1):5–38.

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

Goldwater, S. and Johnson, M. (2003). Learning OT constraint rankings using a maximum entropy model. In Spenader, J., Eriksson, A., and Dahl, Ö., editors, *Variation within Optimality Theory: Proceedings of the Stockholm Workshop on 'Variation within Optimality Theory', April 26–27, 2003 at Department of Linguistics, Stockholm University, Sweden*, pages 111–120, Stockholm. Stockholm University.

References II

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

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 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.

Křen, M., Cvrček, V., Hnátková, M., Jelínek, T., Kocek, J., Kováříková, 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.

Lightner, T. (1965). *Segmental phonology of Modern Standard Russian*. PhD thesis, Massachusetts Institute of Technology, Cambridge.

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.

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