

Supervised Learning of Complete Morphological Paradigms



Greg Durrett and John DeNero
UC Berkeley / Google

Morphological Inflection

train (de)

Zug

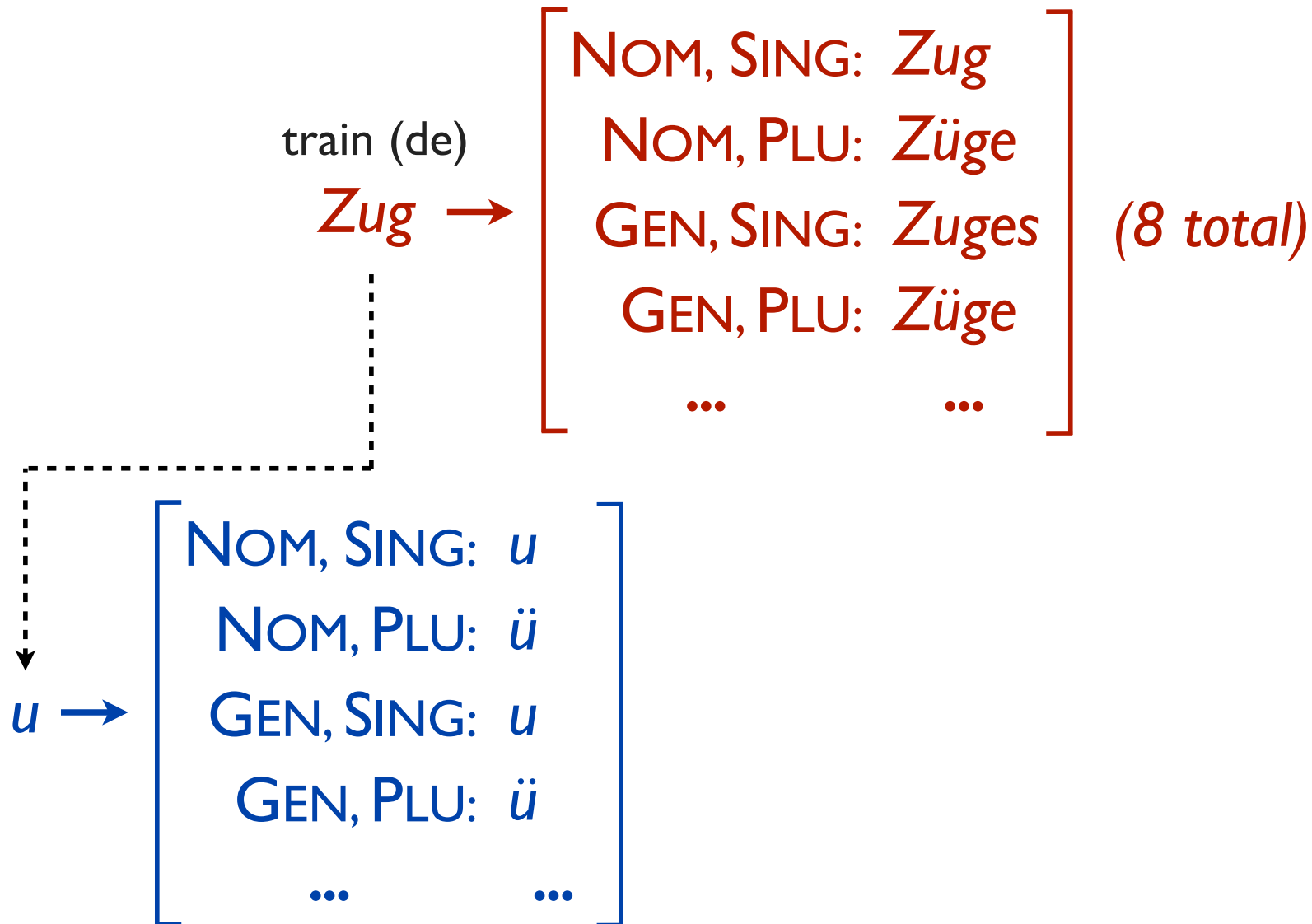
Morphological Inflection

train (de)
Zug →

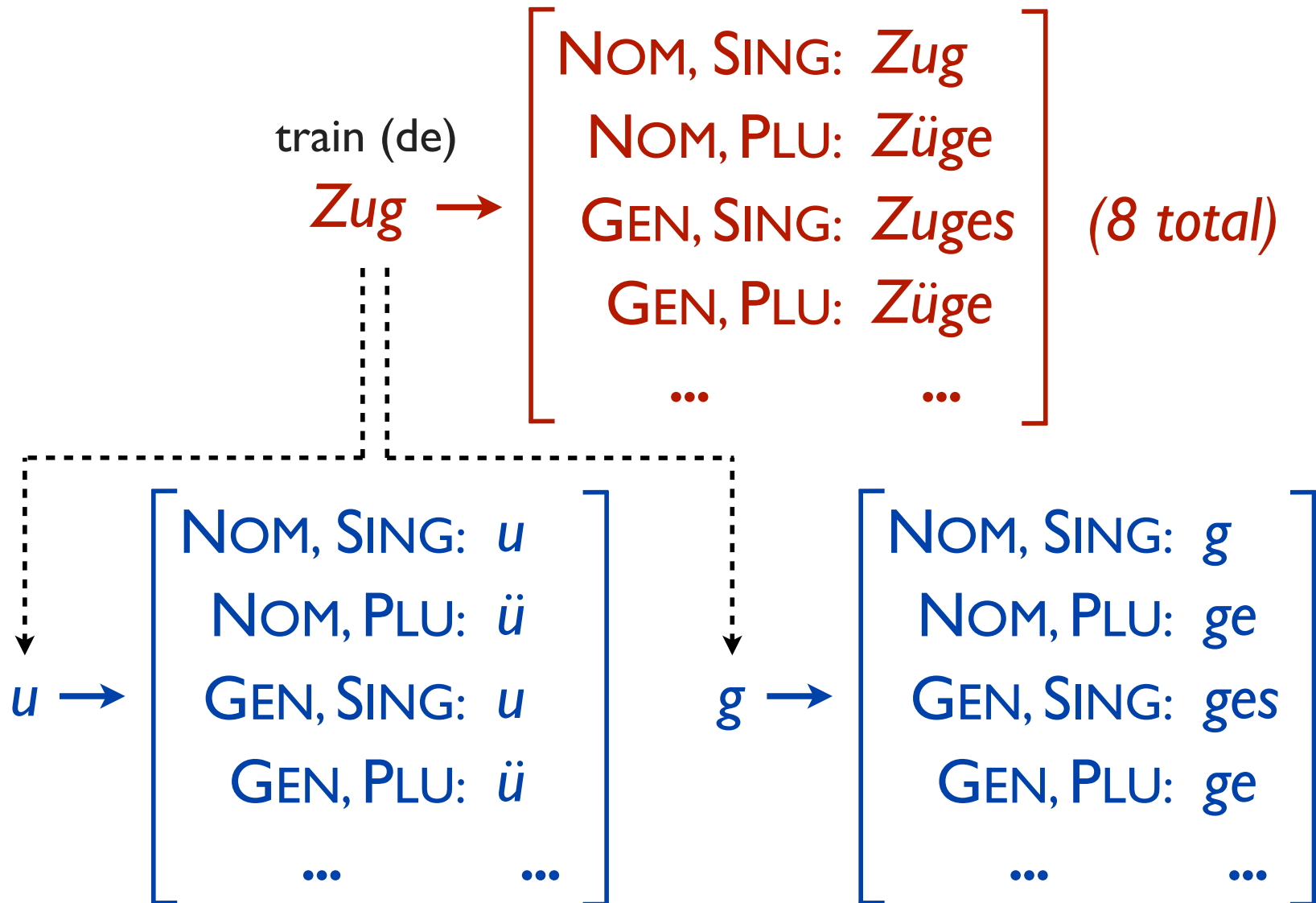
NOM, SING:	<i>Zug</i>
NOM, PLU:	<i>Züge</i>
GEN, SING:	<i>Zuges</i>
GEN, PLU:	<i>Züge</i>
...	...

(8 total)

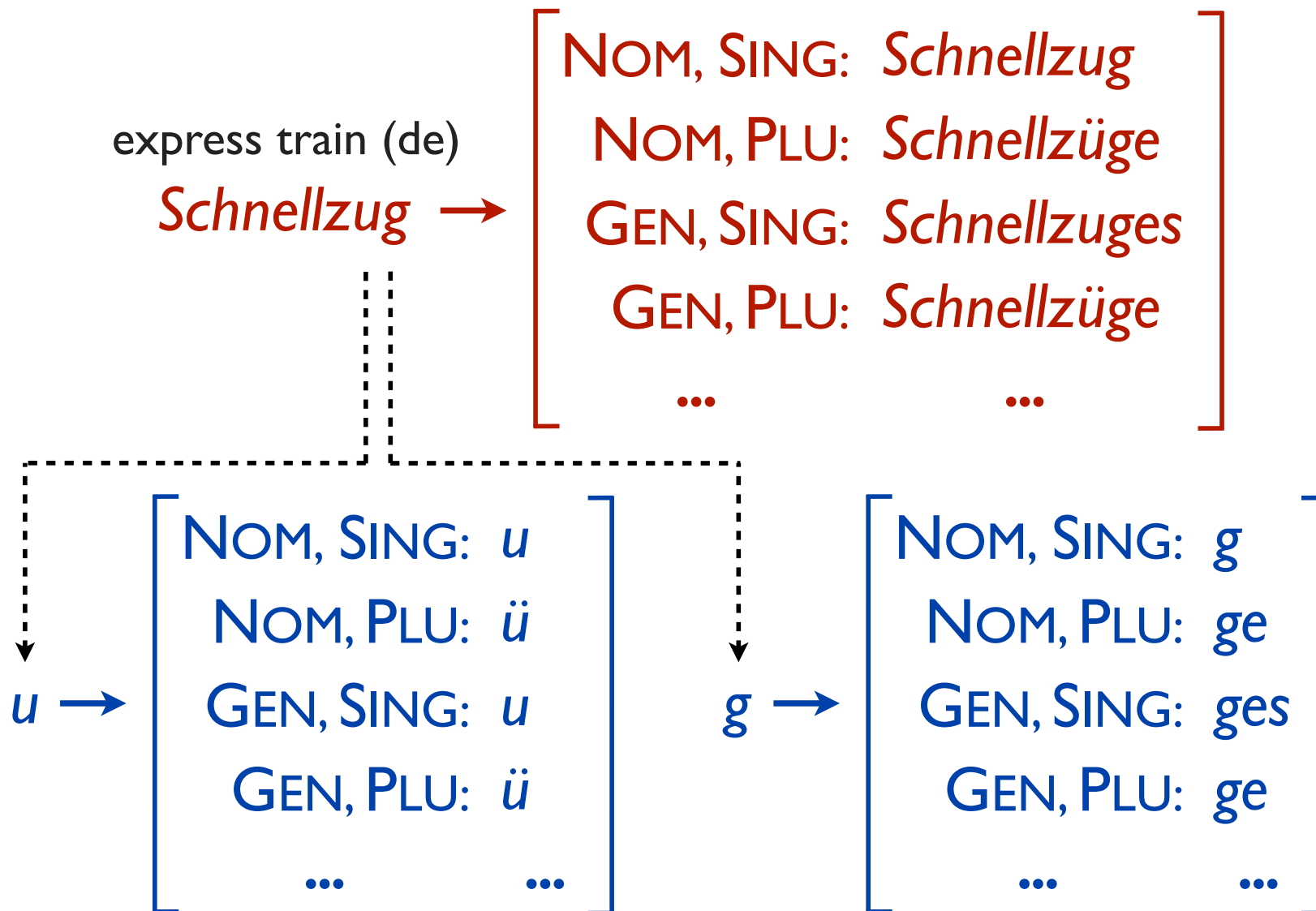
Morphological Inflection



Morphological Inflection



Morphological Inflection



Morphological Inflection

egg white (fi)

valkuainen →

NOM, SING: *valkuainen*

GEN, SING: *valkuaisen*

PART, SING: *valkuaista*

...

...

culprit (fi)

syyllinen →

NOM, SING: *syyllinen*

GEN, SING: *syllisen*

PART, SING: *syllistä*

...

...

Morphological Inflection

egg white (fi)

valkuainen →

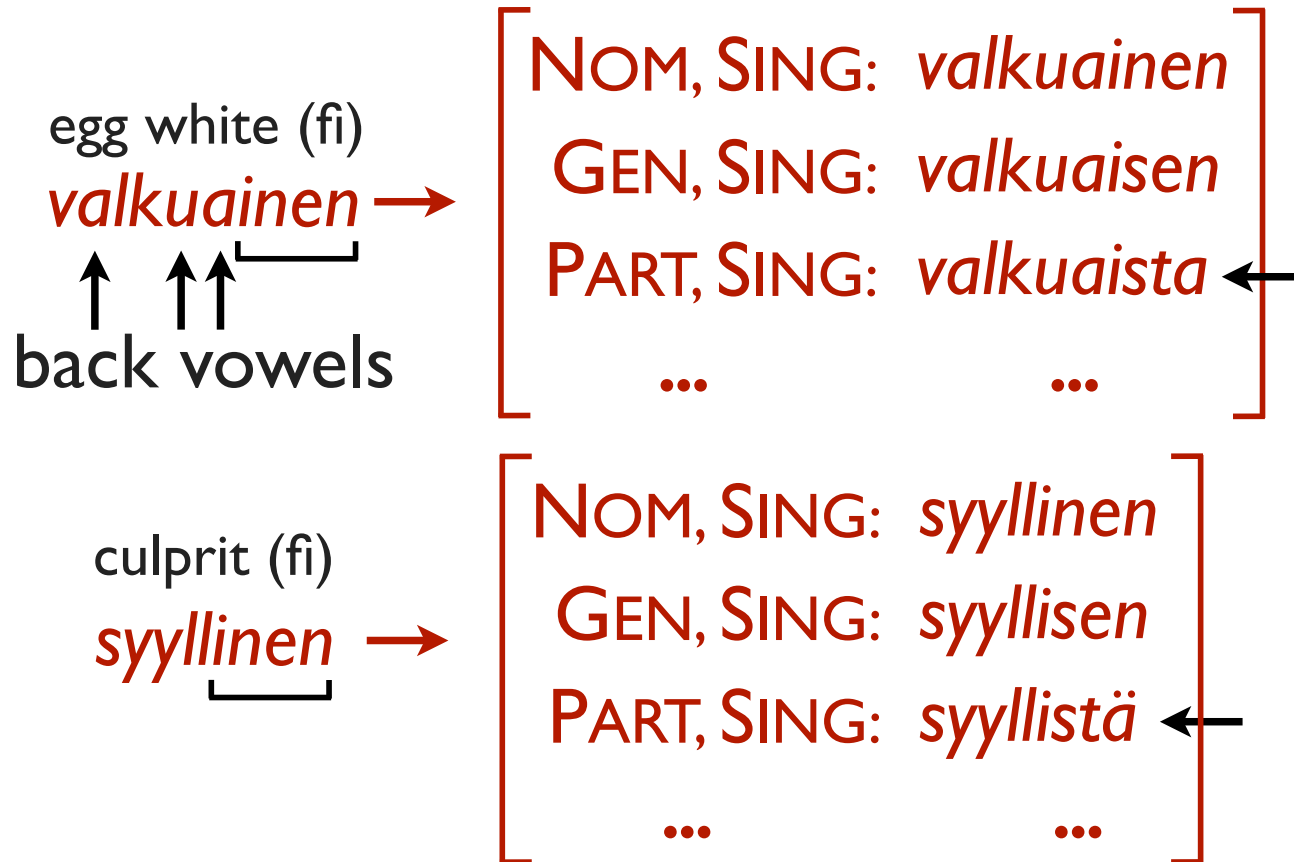
[
NOM, SING: *valkuainen*
GEN, SING: *valkuaisen*
PART, SING: *valkuaista* ←
... ..
]

culprit (fi)

syyllinen →

[
NOM, SING: *syyllinen*
GEN, SING: *syllisen*
PART, SING: *syllistä* ←
... ..
]

Morphological Inflection



Morphological Inflection

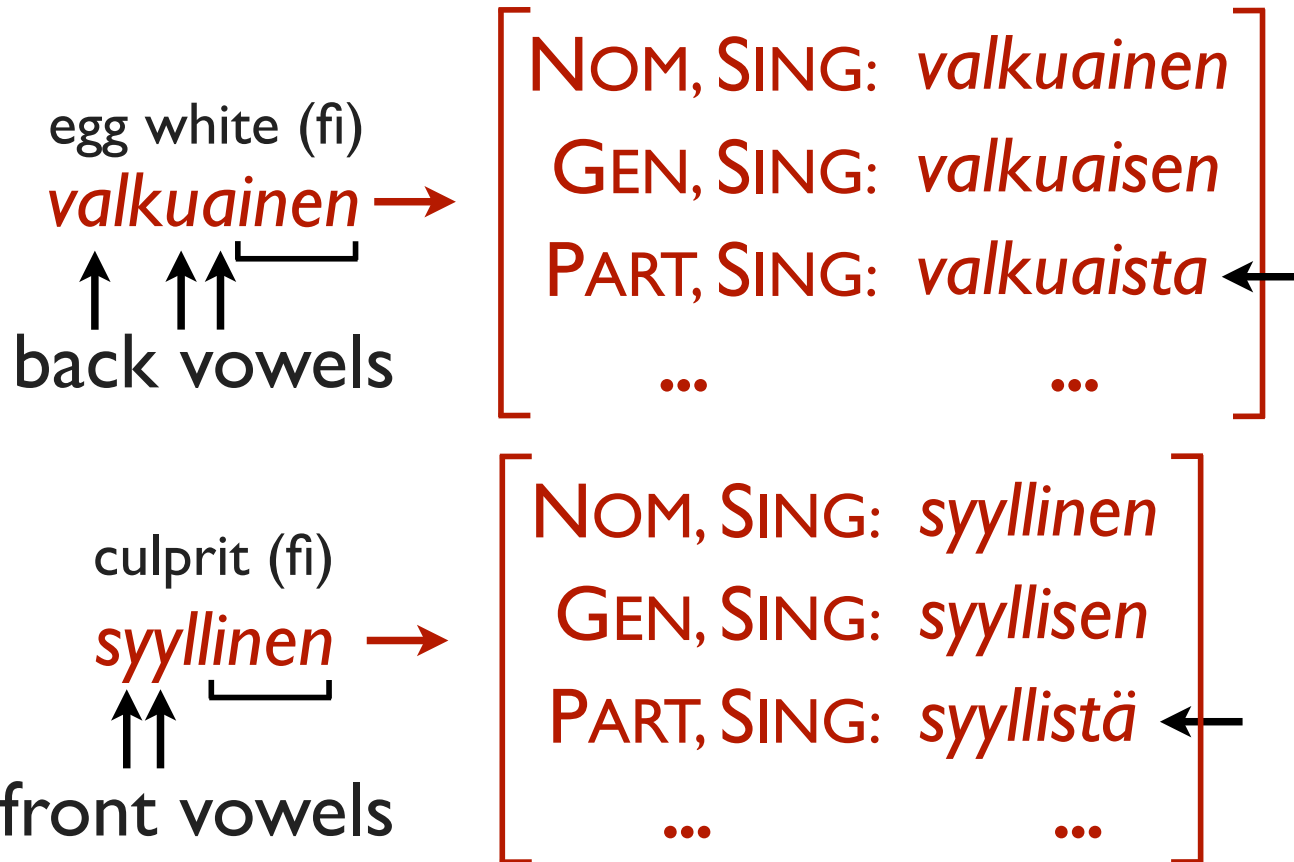
egg white (fi)
valkuainen →
↑ ↑ ↑
back vowels

[
NOM, SING: *valkuainen*
GEN, SING: *valkuaisen*
PART, SING: *valkuaista* ←
... ..
]

culprit (fi)
syyllinen →
↑ ↑
front vowels

[
NOM, SING: *syyllinen*
GEN, SING: *syyllisen*
PART, SING: *syyllistä* ←
... ..
]

Morphological Inflection



- ▶ Suffix morphology can depend on context in a complex way

Morphological Inflection

Given

Base form
(Word)

Morphological Inflection

Given

Base form
(Word)

Predict

Inflection table (variants)

Attributes(1):	Inflected form(1)
Attributes(2):	Inflected form(2)
...	...

Morphological Inflection

Given

Base form
(Word)

Predict

Inflection table (variants)

Attributes(1):	Inflected form(1)
Attributes(2):	Inflected form(2)
...	...

- ▶ Language/POS independence

Morphological Inflection

Given

Base form
(Word)

Predict

Inflection table (variants)

Attributes(1):	Inflected form(1)
Attributes(2):	Inflected form(2)
...	...

- ▶ Language/POS independence
- ▶ Handle multi-part paradigms

Morphological Inflection

Given

Base form
(Word)

Predict

Inflection table (variants)

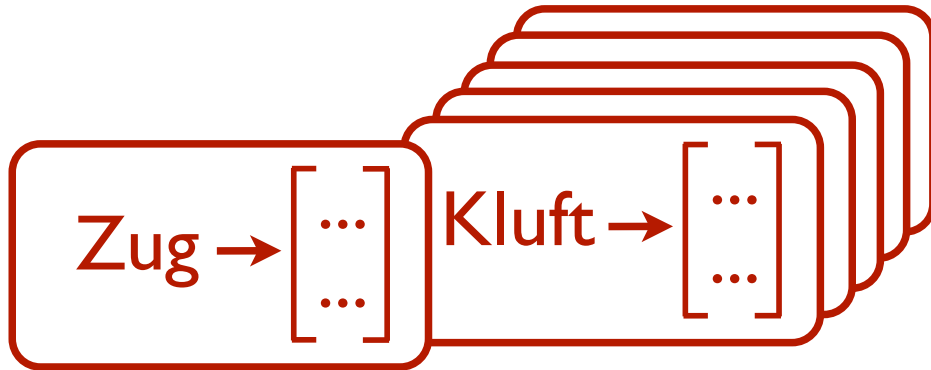
Attributes(1):	Inflected form(1)
Attributes(2):	Inflected form(2)
...	...

- ▶ Language/POS independence
- ▶ Handle multi-part paradigms
- ▶ Rich way to predict inflections

Outline

Outline

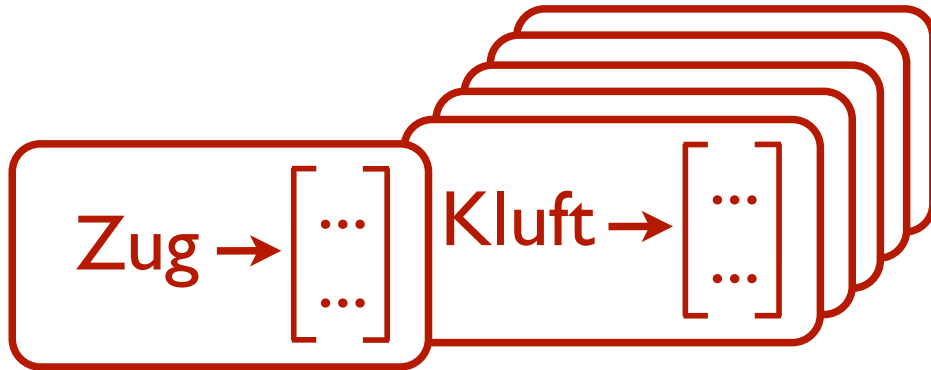
- ▶ Rule extraction (>1000)



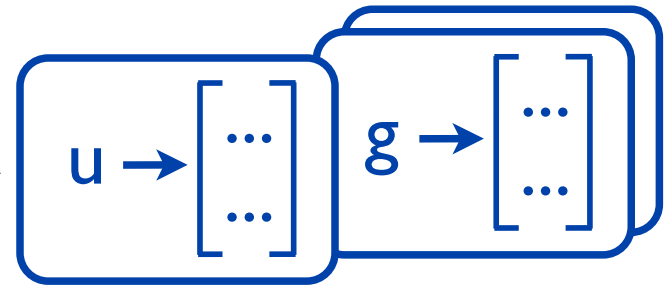
Outline

▶ Rule extraction

(> 1000)



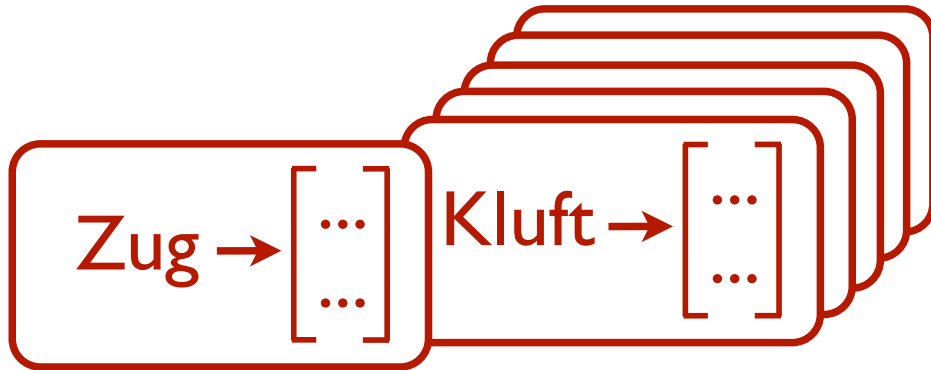
(< 50)



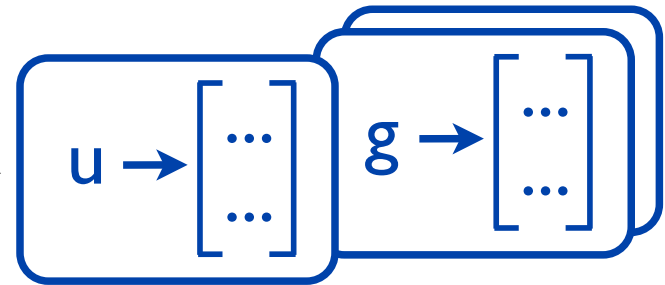
Outline

▶ Rule extraction

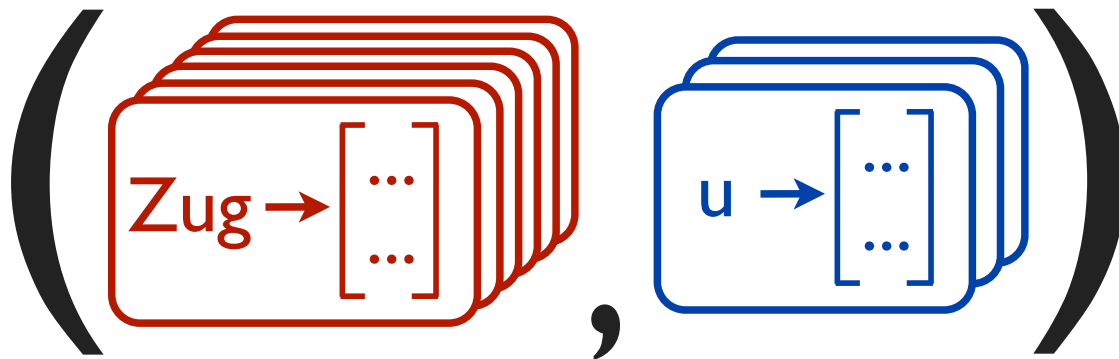
(> 1000)



(< 50)

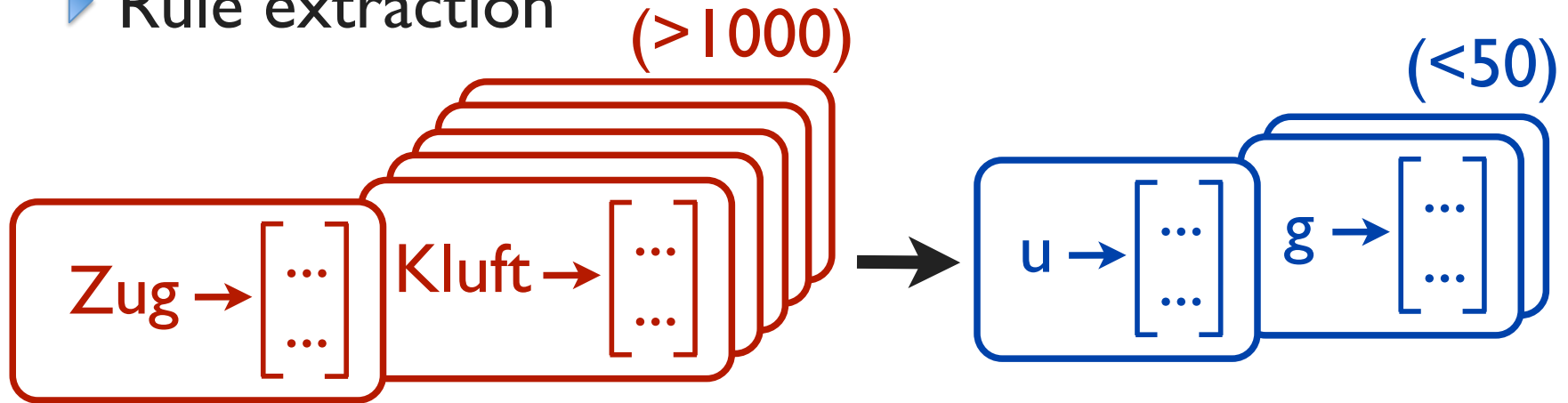


▶ Paradigm prediction

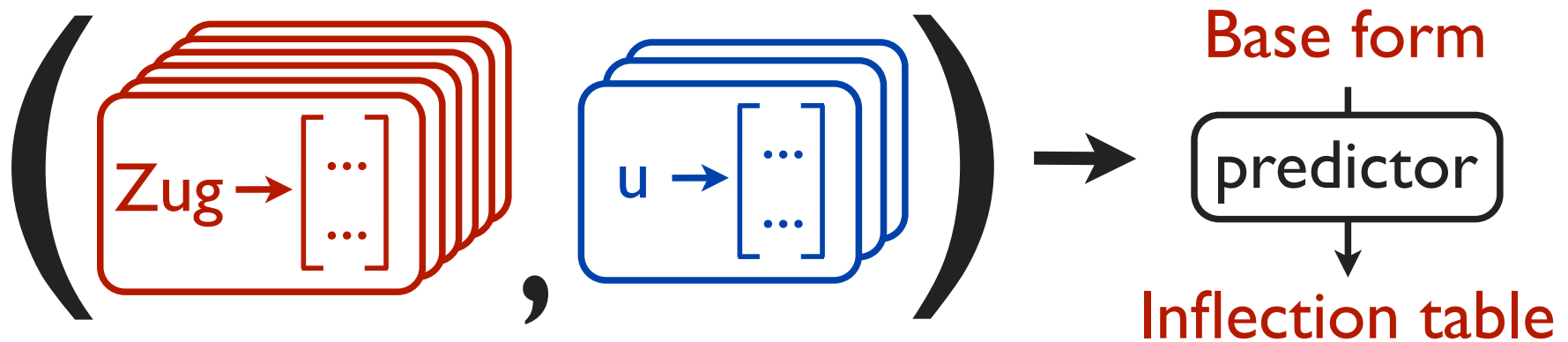


Outline

▶ Rule extraction



▶ Paradigm prediction



Wiktionary Data

- ▶ Thousands of full inflection tables for many languages and parts of speech

Wiktionary Data

- ▶ Thousands of full inflection tables for many languages and parts of speech

<http://en.wiktionary.org/wiki/Zug>

declension of Zug [hide ▲]					
	singular			plural	
<i>m</i> gender	indef.	def.	noun	def.	noun
nominative	ein	der	Zug	die	Züge
genitive	eines	des	Zuges, Zugs	der	Züge
dative	einem	dem	Zug, Zuge (<i>archaic</i>)	den	Zügen
accusative	einen	den	Zug	die	Züge

Wiktionary Data

- ▶ Thousands of full inflection tables for many languages and parts of speech

<http://en.wiktionary.org/wiki/Zug>

declension of Zug [hide ▲]					
	singular			plural	
<i>m gender</i>	indef.	def.	noun	def.	noun
nominative	ein	der	Zug	die	Züge
genitive	eines	des	Zuges, Zugs	der	Züge
dative	einem	dem	Zug, Zuge (<i>archaic</i>)	den	Zügen
accusative	einen	den	Zug	die	Züge

2763 more...

Previous Work

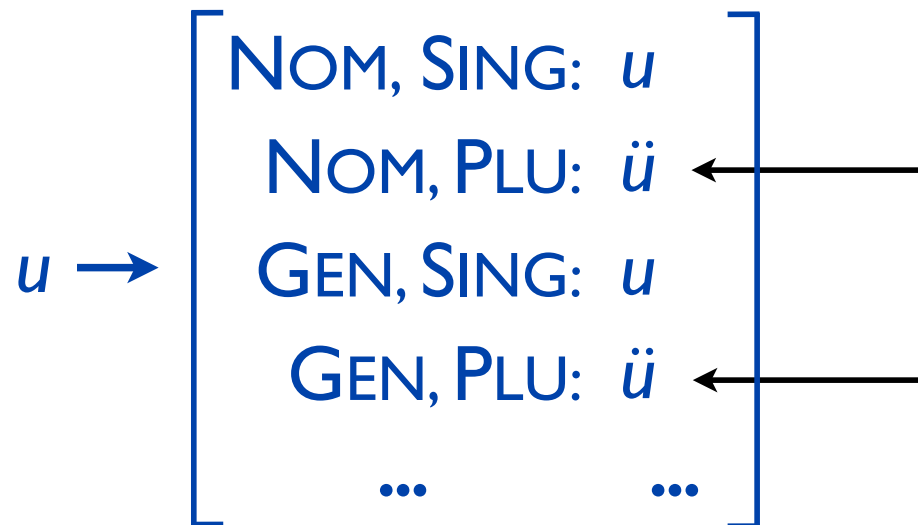
- ▶ [Dreyer and Eisner \(2011\)](#) develops a semi-supervised Bayesian model

Previous Work

- ▶ [Dreyer and Eisner \(2011\)](#) develops a semi-supervised Bayesian model
- ▶ [Wicentowski \(2004\)](#) predicts one variant at a time (the lemma)

Previous Work

- ▶ [Dreyer and Eisner \(2011\)](#) develops a semi-supervised Bayesian model
- ▶ [Wicentowski \(2004\)](#) predicts one variant at a time (the lemma)



Rule Extraction

Rule Extraction

to insist (de)
dringen →

1P, PRES, SING: <i>dringe</i>
2P, PRES, SING: <i>dringst</i>
3P, PRES, SING: <i>dringt</i>
1P, PAST, SING: <i>drang</i>
...
...

(27 total)

Rule Extraction

to insist (de)

dringen



1 P, PRES, SING: *dringe*
2 P, PRES, SING: *dringst*
3 P, PRES, SING: *dringt*
1 P, PAST, SING: *drang*
... ..

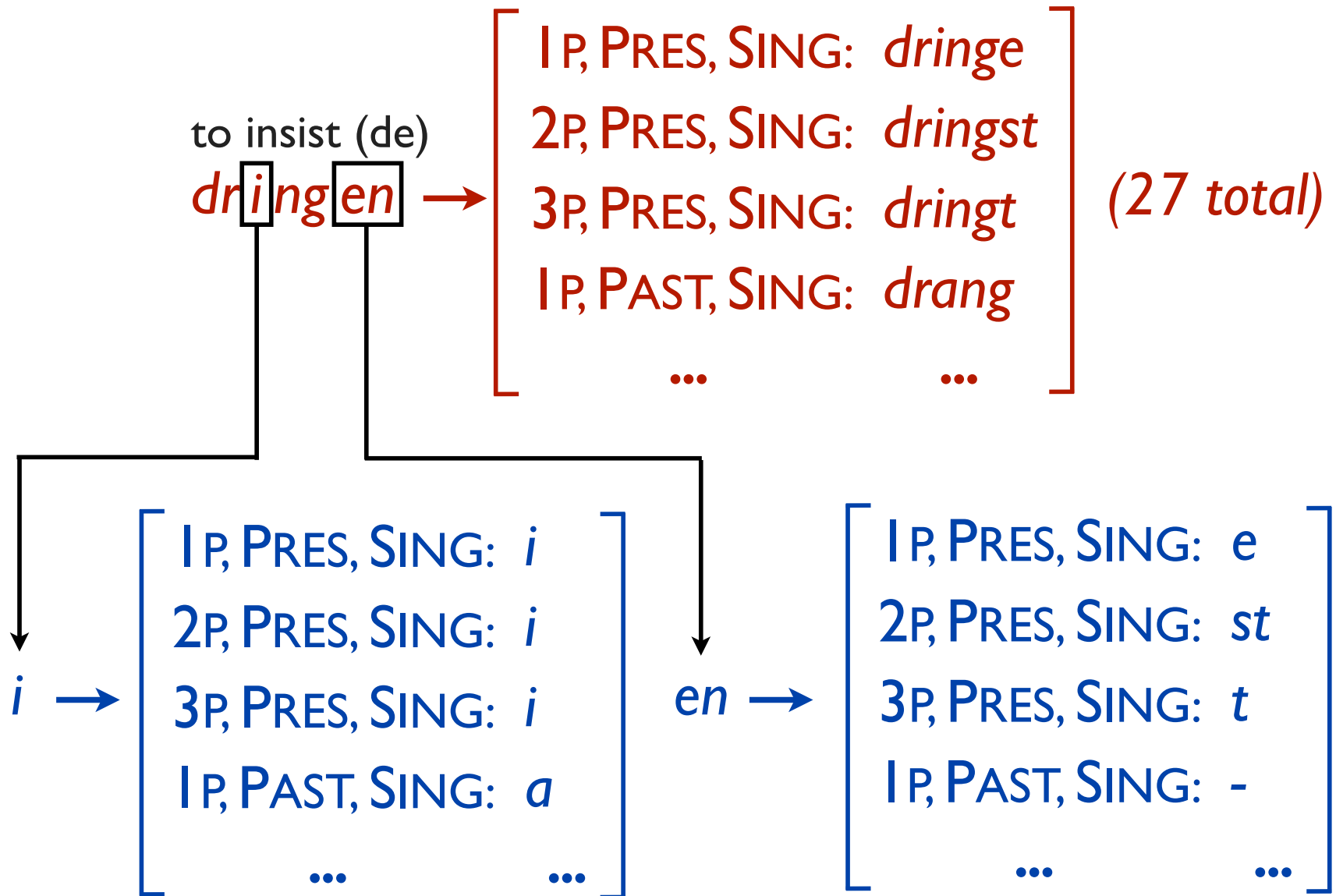
(27 total)

i



1 P, PRES, SING: *i*
2 P, PRES, SING: *i*
3 P, PRES, SING: *i*
1 P, PAST, SING: *a*
... ..

Rule Extraction



Alignment to Base Form

INFINITIVE

d r i n g e n

d r i n g e

I P, PRES, SING

Alignment to Base Form

INFINITIVE

<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	<i>n</i>
						D
<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	

I, PRES, SING

Alignment to Base Form

INFINITIVE	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	<i>n</i>
I, PRES, SING	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	

Alignment to Base Form

INFINITIVE

d r i n g e n

INFINITIVE

d r i n g e n

I P, PRES, SING

d r i n g e

d r i n g s t

2P, PRES, SING

Alignment to Base Form

INFINITIVE
d r i n g e n
| | | | | **S S**
d r i n g s t
2P, PRES, SING

INFINITIVE *d r i n g e n*
I P, PRES, SING *d r i n g e* **n**

Alignment to Base Form

INFINITIVE	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	<i>n</i>
1P, PRES, SING	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	
INFINITIVE	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	<i>n</i>
2P, PRES, SING	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>s</i>	<i>t</i>

Alignment to Base Form

INFINITIVE *d r i n g e n*
| | | | | **S D**
d r i n g t
3P, PRES, SING

INFINITIVE *d r i n g e n*
1P, PRES, SING *d r i n g e*
INFINITIVE *d r i n g e n*
2P, PRES, SING *d r i n g s t*

Alignment to Base Form

INFINITIVE	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	<i>n</i>
1P, PRES, SING	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	
INFINITIVE	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	<i>n</i>
2P, PRES, SING	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>s</i>	<i>t</i>
INFINITIVE	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	<i>n</i>
3P, PRES, SING	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>t</i>	

Alignment to Base Form

d r i n g e n
| | **S** | | **D D**
d r a n g
1P, PAST, SING

INFINITIVE

INFINITIVE *d r i n g e n*
1P, PRES, SING *d r i n g e*

INFINITIVE *d r i n g e n*
2P, PRES, SING *d r i n g s t*

INFINITIVE *d r i n g e n*
3P, PRES, SING *d r i n g t*

Alignment to Base Form

INFINITIVE	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	<i>n</i>
1P, PRES, SING	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	
INFINITIVE	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	<i>n</i>
2P, PRES, SING	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>s</i>	<i>t</i>
INFINITIVE	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	<i>n</i>
3P, PRES, SING	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>t</i>	
INFINITIVE	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	<i>n</i>
1P, PAST, SING	<i>d</i>	<i>r</i>	<i>a</i>	<i>n</i>	<i>g</i>		

Concatenation

INFINITIVE	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	<i>n</i>
I P, PRES, SING	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	
INFINITIVE	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	<i>n</i>
2P, PRES, SING	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>s</i>	<i>t</i>
INFINITIVE	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	<i>n</i>
3P, PRES, SING	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>t</i>	
INFINITIVE	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	<i>n</i>
I P, PAST, SING	<i>d</i>	<i>r</i>	<i>a</i>	<i>n</i>	<i>g</i>		

Concatenation

INFINITIVE	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	<i>n</i>
I P, PRES, SING	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	
INFINITIVE	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	<i>n</i>
2P, PRES, SING	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>s</i>	<i>t</i>
INFINITIVE	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	<i>n</i>
3P, PRES, SING	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>t</i>	
INFINITIVE	<i>d</i>	<i>r</i>	<i>i</i>	<i>n</i>	<i>g</i>	<i>e</i>	<i>n</i>
I P, PAST, SING	<i>d</i>	<i>r</i>	<i>a</i>	<i>n</i>	<i>g</i>		

Concatenation


INFINITIVE *d r i n g e n* *d r i n g e n*
1P, PRES, SING *d r i n g e* n

INFINITIVE *d r i n g e n*
2P, PRES, SING *d r i n g s t*

INFINITIVE *d r i n g* *e n*
3P, PRES, SING *d r i n g* *t*

INFINITIVE *d r* *i* *n g* *e n*
1P, PAST, SING *d r* *a* *n g*

Concatenation


INFINITIVE *d r i n g e n* *d r i n g e n*
1P, PRES, SING *d r i n g e* 

INFINITIVE *d r i n g e n*
2P, PRES, SING *d r i n g s t*

INFINITIVE *d r i n g e n*
3P, PRES, SING *d r i n g t*

INFINITIVE *d r i n g e n*
1P, PAST, SING *d r a n g*

Concatenation

INFINITIVE	<i>d r i n g e n</i>	<i>d r i n g e n</i>
I P, PRES, SING	<i>d r i n g e</i>	
INFINITIVE	<i>d r i n g e n</i>	
2P, PRES, SING	<i>d r i n g s t</i>	
INFINITIVE	<i>d r i n g e n</i>	
3P, PRES, SING	<i>d r i n g t</i>	
INFINITIVE	<i>d r i n g e n</i>	
I P, PAST, SING	<i>d r a n g</i>	

Concatenation

INFINITIVE	<i>d r i n g e n</i>	<i>d r <u>i</u> n g <u>e</u> n</i>
I P, PRES, SING	<i>d r i n g e</i>	
INFINITIVE	<i>d r i n g e n</i>	
2P, PRES, SING	<i>d r i n g s t</i>	
INFINITIVE	<i>d r i n g e n</i>	
3P, PRES, SING	<i>d r i n g t</i>	
INFINITIVE	<i>d r i n g e n</i>	
I P, PAST, SING	<i>d r a n g</i>	

Concatenation

INFINITIVE	<i>d r i n g e n</i>	<i>d r <u>i</u> n g <u>e</u> n</i>
I P, PRES, SING	<i>d r i n g e</i>	
INFINITIVE	<i>d r i n g e n</i>	
2P, PRES, SING	<i>d r i n g s t</i>	
INFINITIVE	<i>d r i n g e n</i>	
3P, PRES, SING	<i>d r i n g t</i>	
INFINITIVE	<i>d r i n g e n</i>	
I P, PAST, SING	<i>d r a n g</i>	

Concatenation

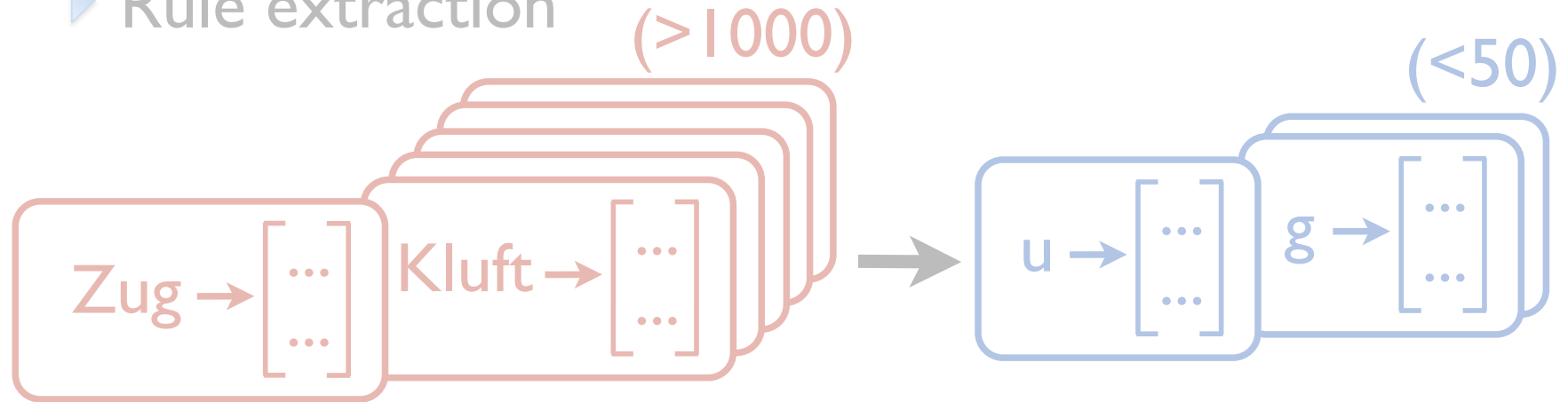
INFINITIVE	<i>d r i n g e n</i>	<i>d r <u>i</u> n g <u>e</u> n</i>
1P, PRES, SING	<i>d r i n g e</i>	
INFINITIVE	<i>d r i n g e n</i>	
2P, PRES, SING	<i>d r i n g s t</i>	<i>d r i n g e</i>
		<i>d r i n g s t</i>
INFINITIVE	<i>d r i n g e n</i>	<i>d r i n g t</i>
3P, PRES, SING	<i>d r i n g t</i>	<i>d r a n g</i>
INFINITIVE	<i>d r i n g e n</i>	
1P, PAST, SING	<i>d r a n g</i>	

Concatenation

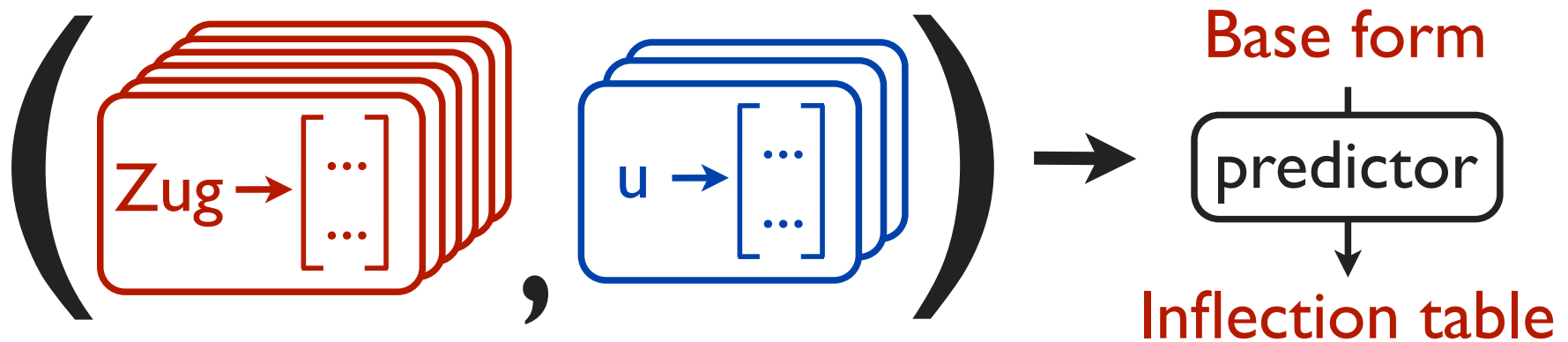
INFINITIVE	<i>d r i n g e n</i>	<i>d r</i>	<i>i</i>	<i>n g e n</i>
1P, PRES, SING	<i>d r i n g e</i>			
INFINITIVE	<i>d r i n g e n</i>			
2P, PRES, SING	<i>d r i n g s t</i>	<i>d r</i>	<i>i</i>	<i>n g e</i>
		<i>d r</i>	<i>i</i>	<i>n g s t</i>
INFINITIVE	<i>d r i n g e n</i>	<i>d r</i>	<i>i</i>	<i>n g t</i>
3P, PRES, SING	<i>d r i n g t</i>	<i>d r</i>	<i>a</i>	<i>n g</i>
INFINITIVE	<i>d r i n g e n</i>			
1P, PAST, SING	<i>d r a n g</i>			

Outline

▶ Rule extraction



▶ Paradigm prediction

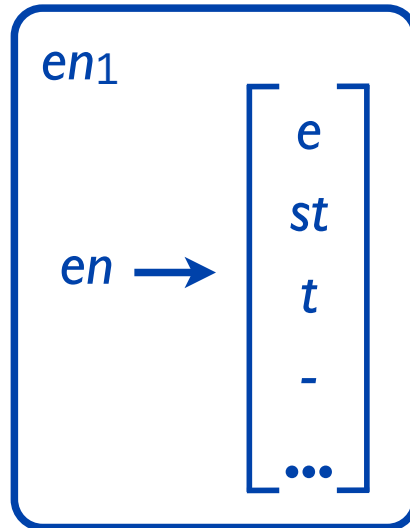
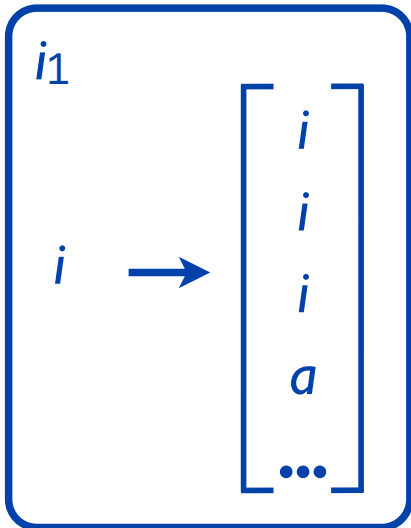


Paradigm Prediction

w i n d e n to wind (de)

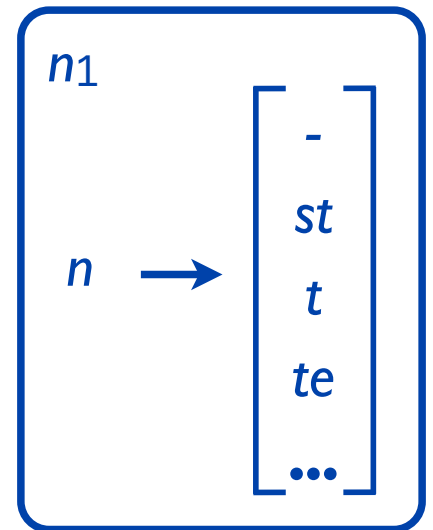
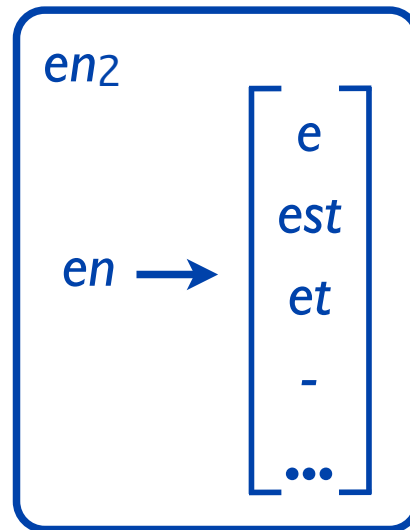
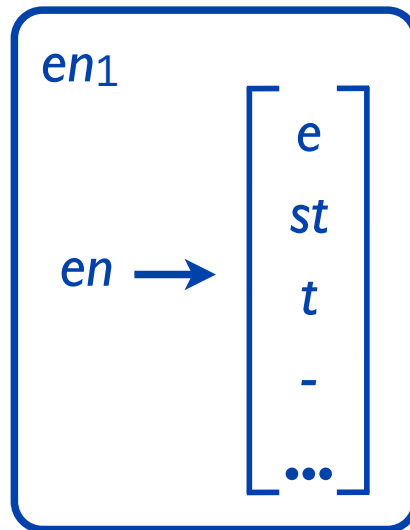
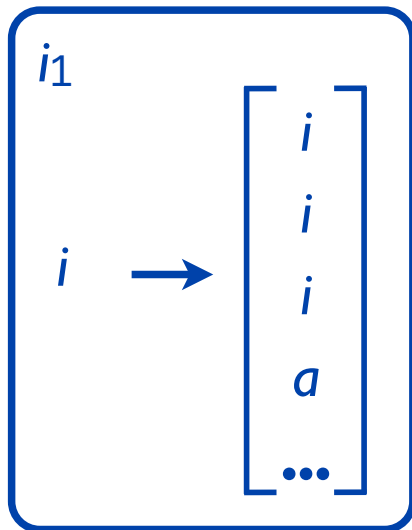
Paradigm Prediction

w i n d e n to wind (de)



Paradigm Prediction

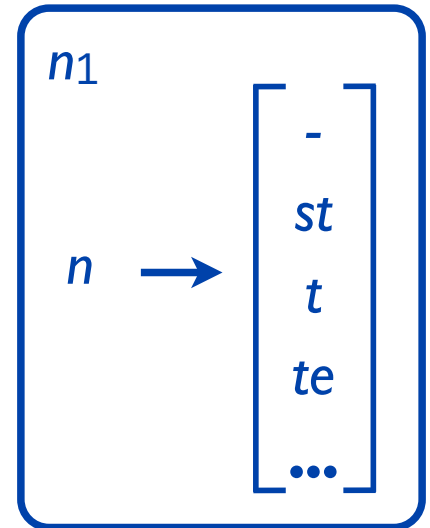
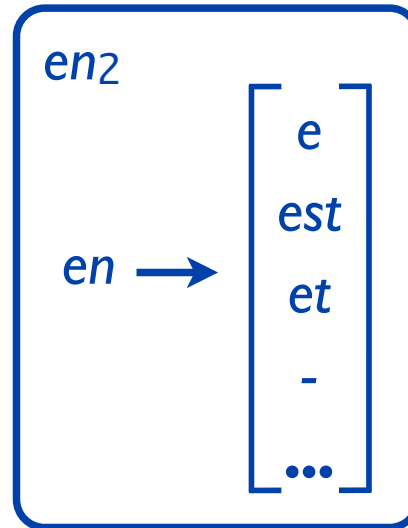
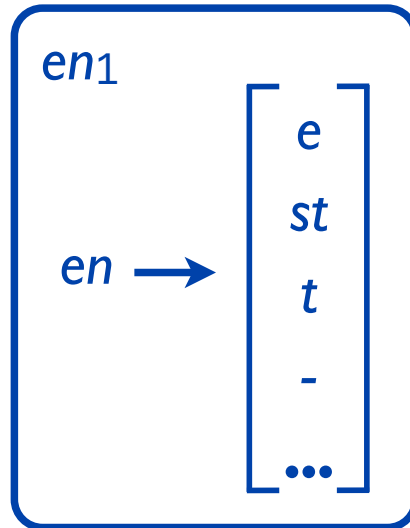
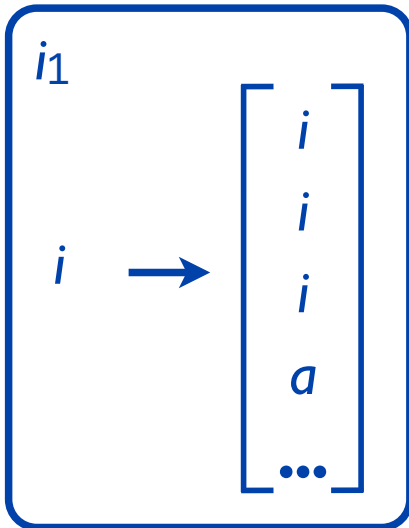
w i n d e n to wind (de)



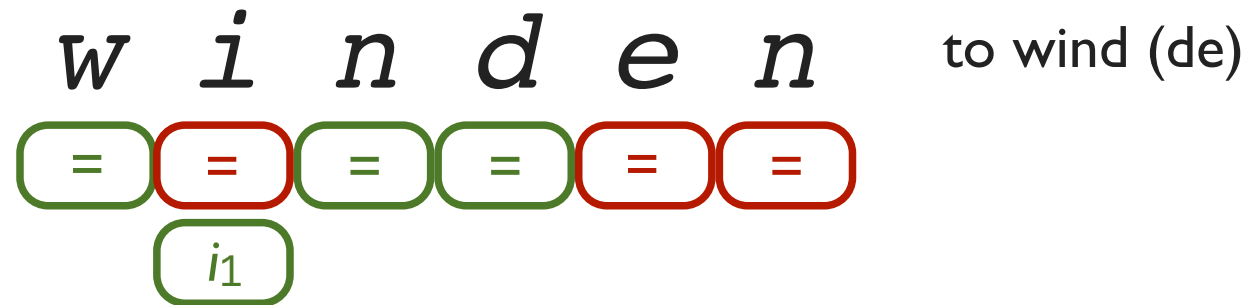
Paradigm Prediction

w i n d e n to wind (de)
= = =
*i*₁

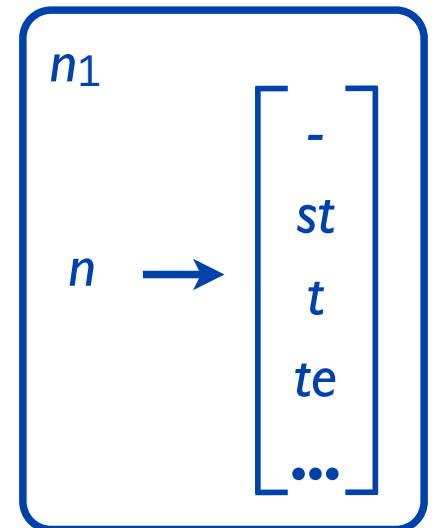
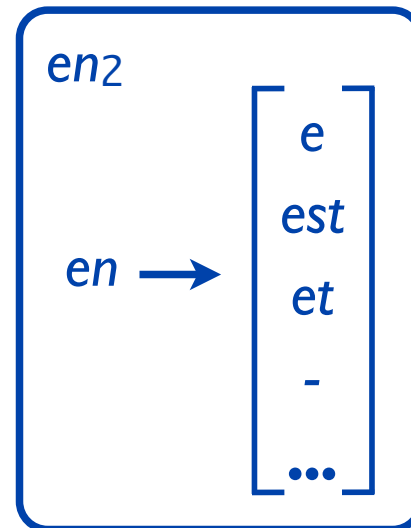
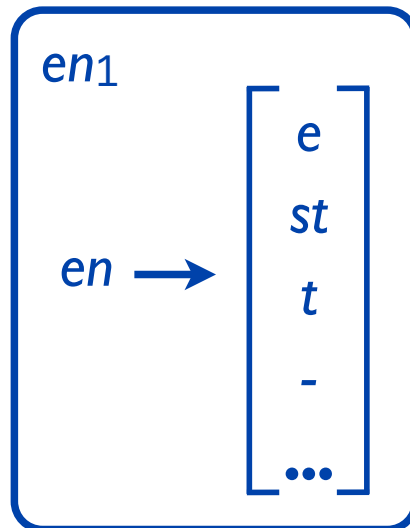
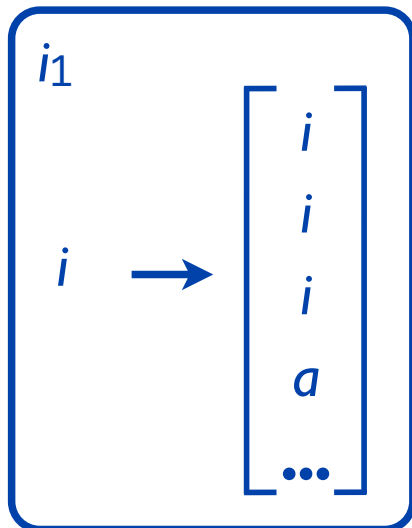
*en*₂



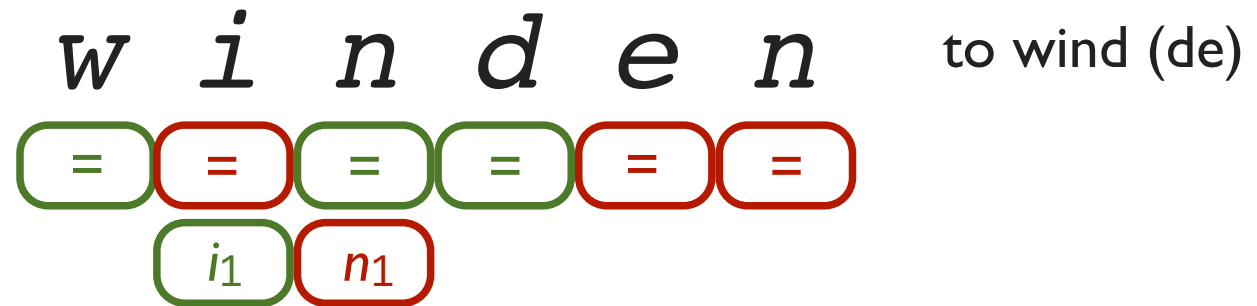
Paradigm Prediction



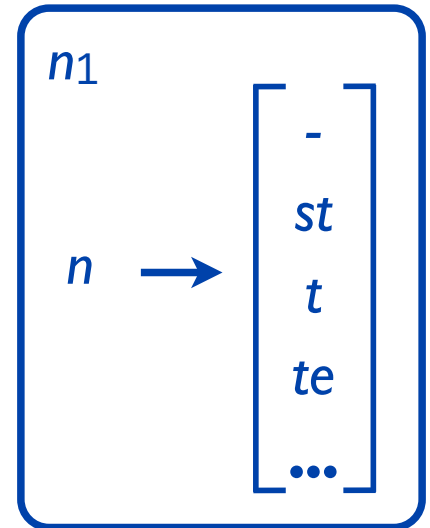
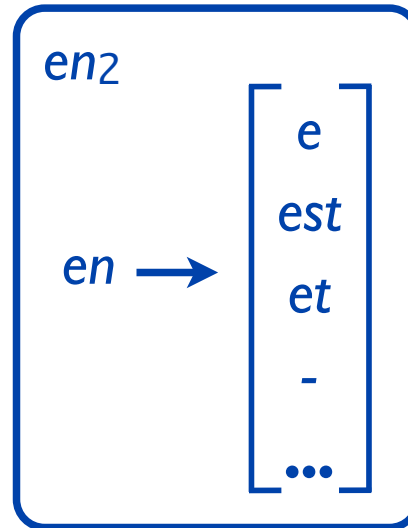
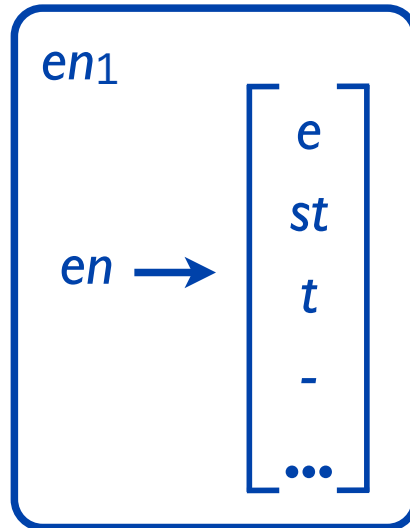
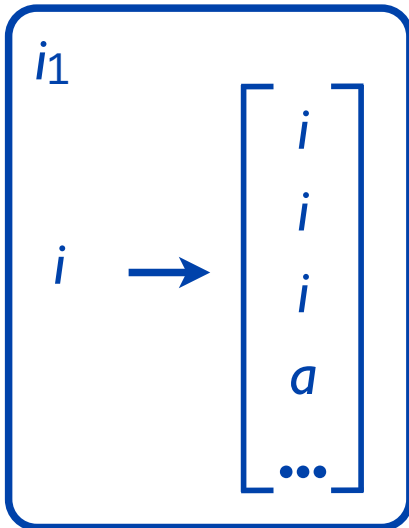
en_2



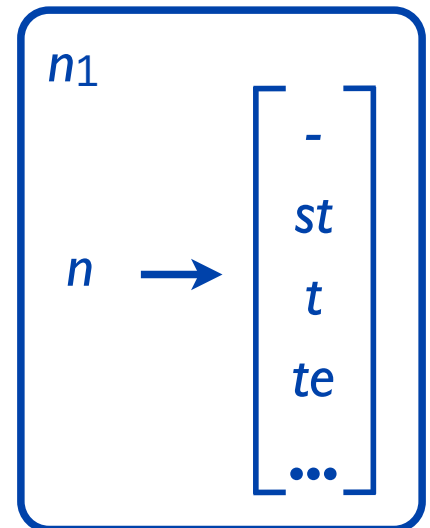
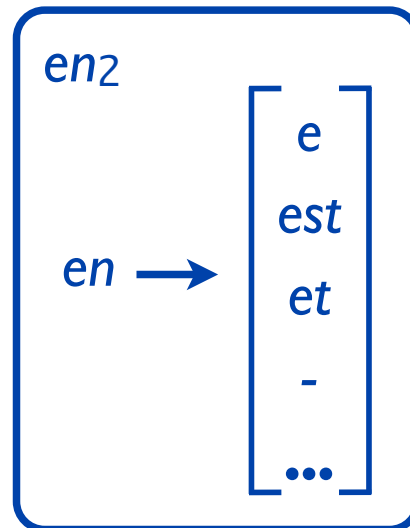
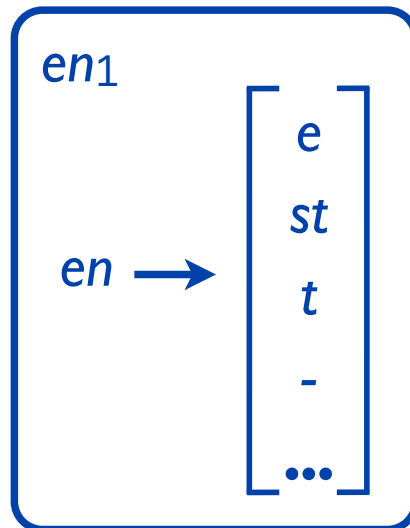
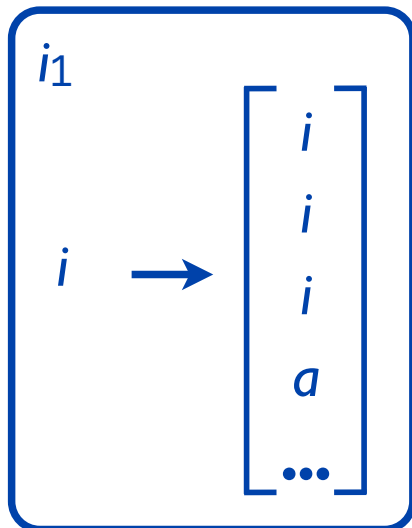
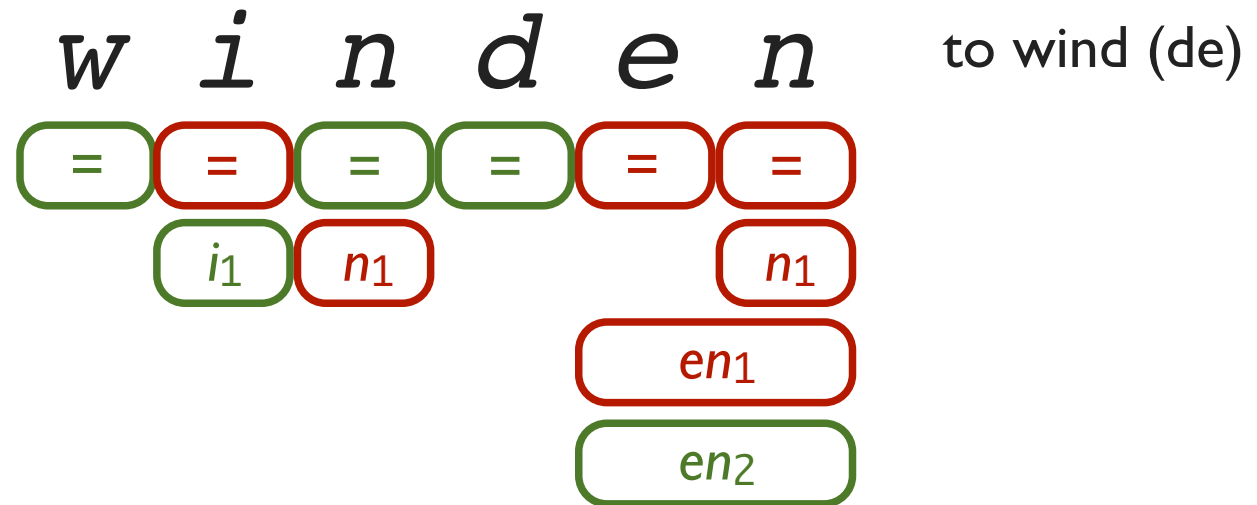
Paradigm Prediction



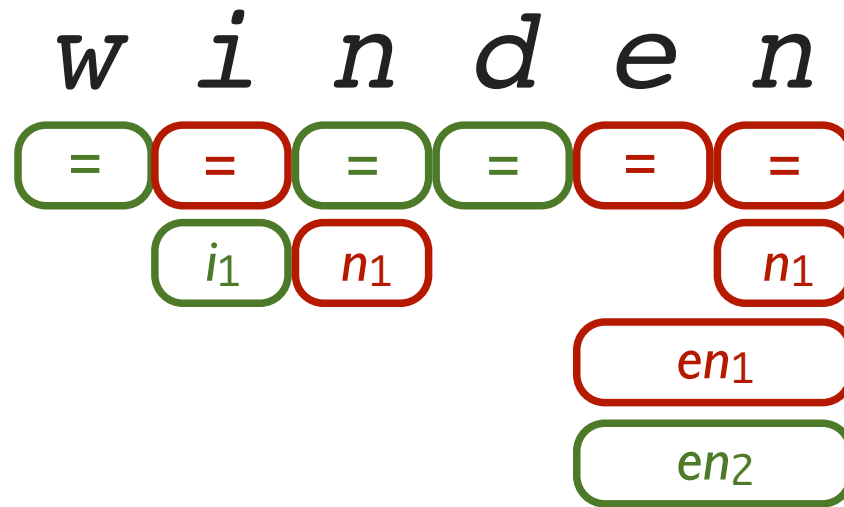
*en*₂



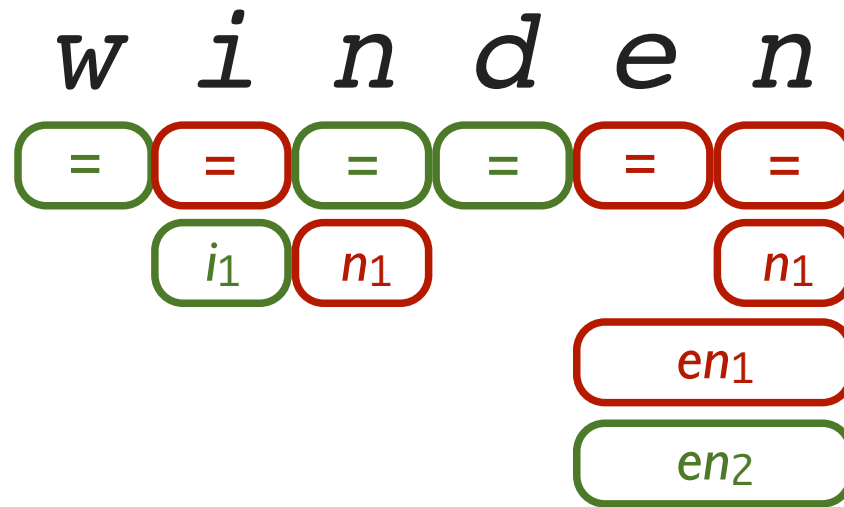
Paradigm Prediction



Paradigm Prediction

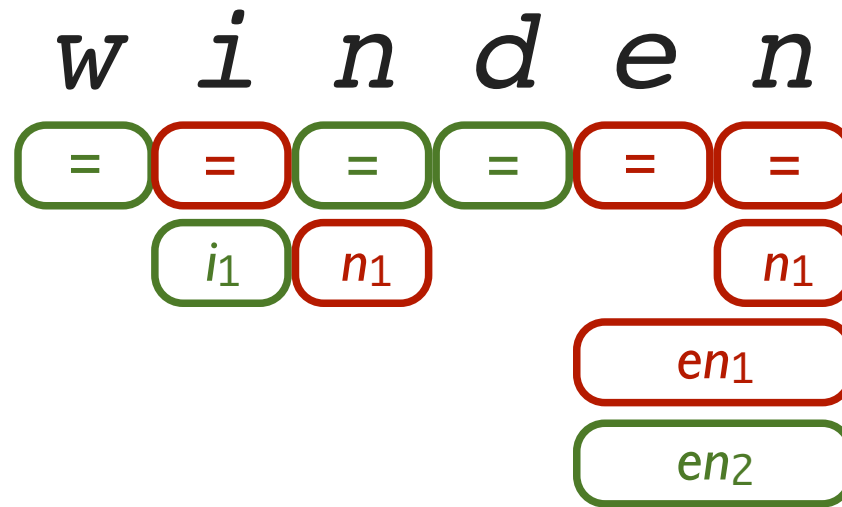


Paradigm Prediction



- ▶ Paths through this lattice are hypotheses

Paradigm Prediction



- ▶ Paths through this lattice are hypotheses
- ▶ Model with semi-Markov CRF
(Sarawagi and Cohen 2004)

Prediction Features

w i n d e n
*i*₁

Prediction Features

w i n d e n

i₁

binden

verbinden

...

Prediction Features

w i n d e n

i₁

binden

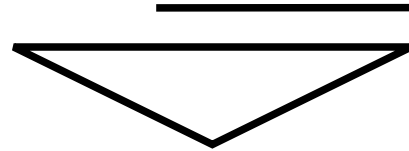
verbinden

...

Prediction Features

w i n d e n

i₁



[i₁] nd

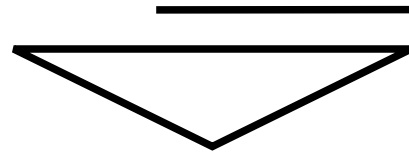
binden
verbinden

...

Prediction Features

w i n d e n

i₁



[i₁] nd

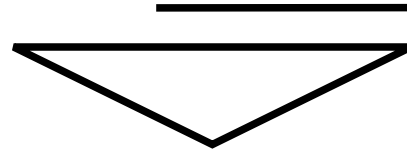
binden
verbinden
...

- ▶ Rule identity conjoined with 1- through 4-grams at offsets up to +/-5

Prediction Features

w i n d e n

i₁



[i₁] nd

binden
verbinden
...

- ▶ Rule identity conjoined with 1- through 4-grams at offsets up to +/-5
- ▶ Coarse features shared between different rules

Learning

- ▶ The gold inflection table of every training example can be produced using our rules

Learning

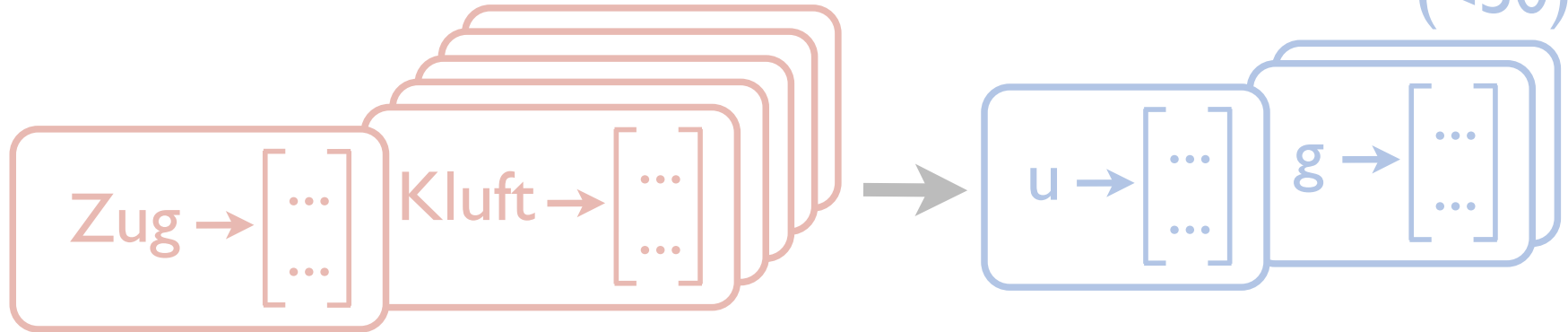
- ▶ The gold inflection table of every training example can be produced using our rules
- ▶ Optimize conditional log-likelihood of the correct rule sequence

Outline

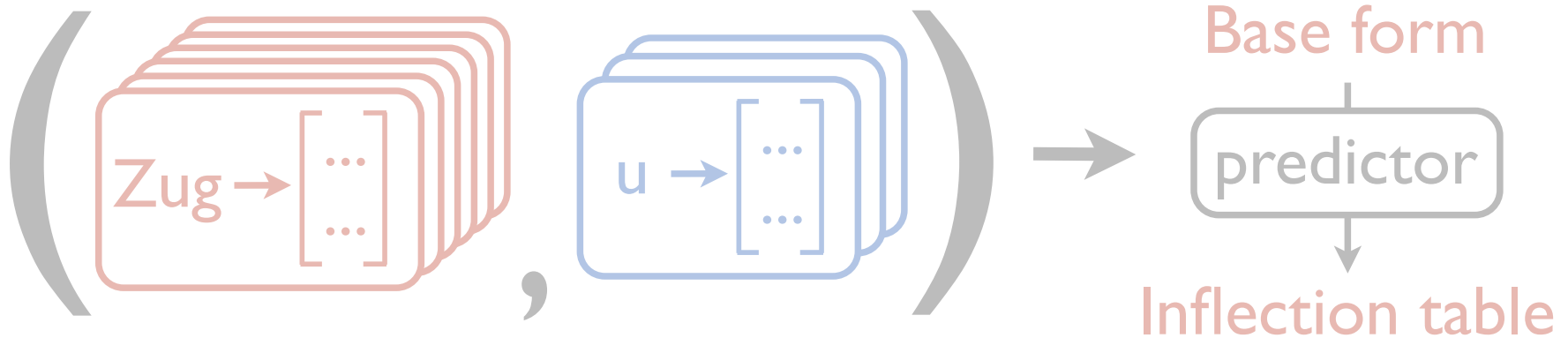
▶ Rule extraction

(> 1000)

(< 50)



▶ Paradigm prediction



Wiktionary Evaluation Setup

	# lines/table	Train size	Test size
German nouns	8	2364	200
German verbs	27	1627	200
Spanish verbs	57	3655	200
Finnish nouns/adjs	28	6000	200
Finnish verbs	53	6849	200

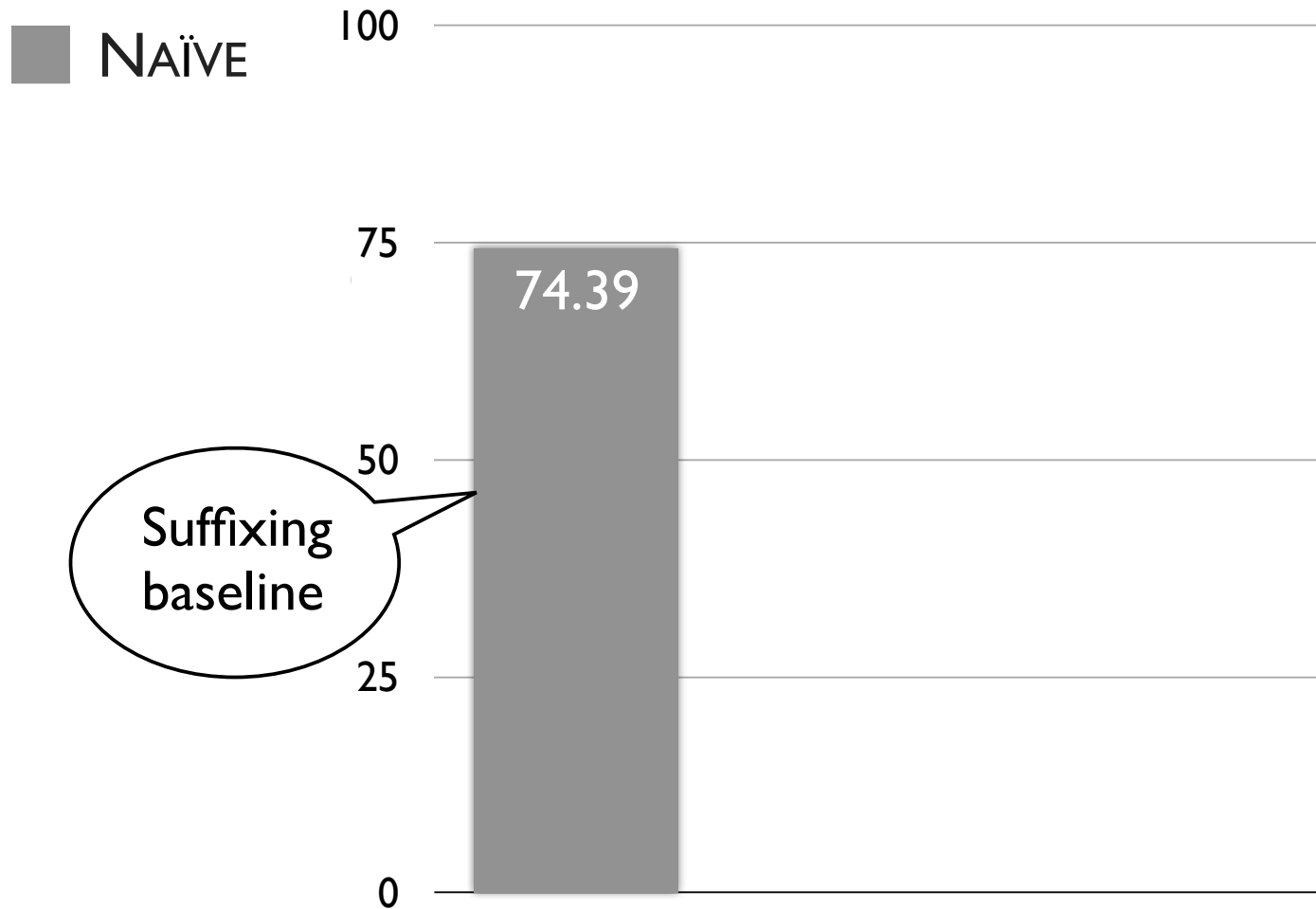
Wiktionary Evaluation Setup

	# lines/table	Train size	Test size
German nouns	8	2364	200
German verbs	27	1627	200
Spanish verbs	57	3655	200
Finnish nouns/adjs	28	6000	200
Finnish verbs	53	6849	200

- ▶ Most common 200 inflection tables are not chosen for the test set; we can memorize these

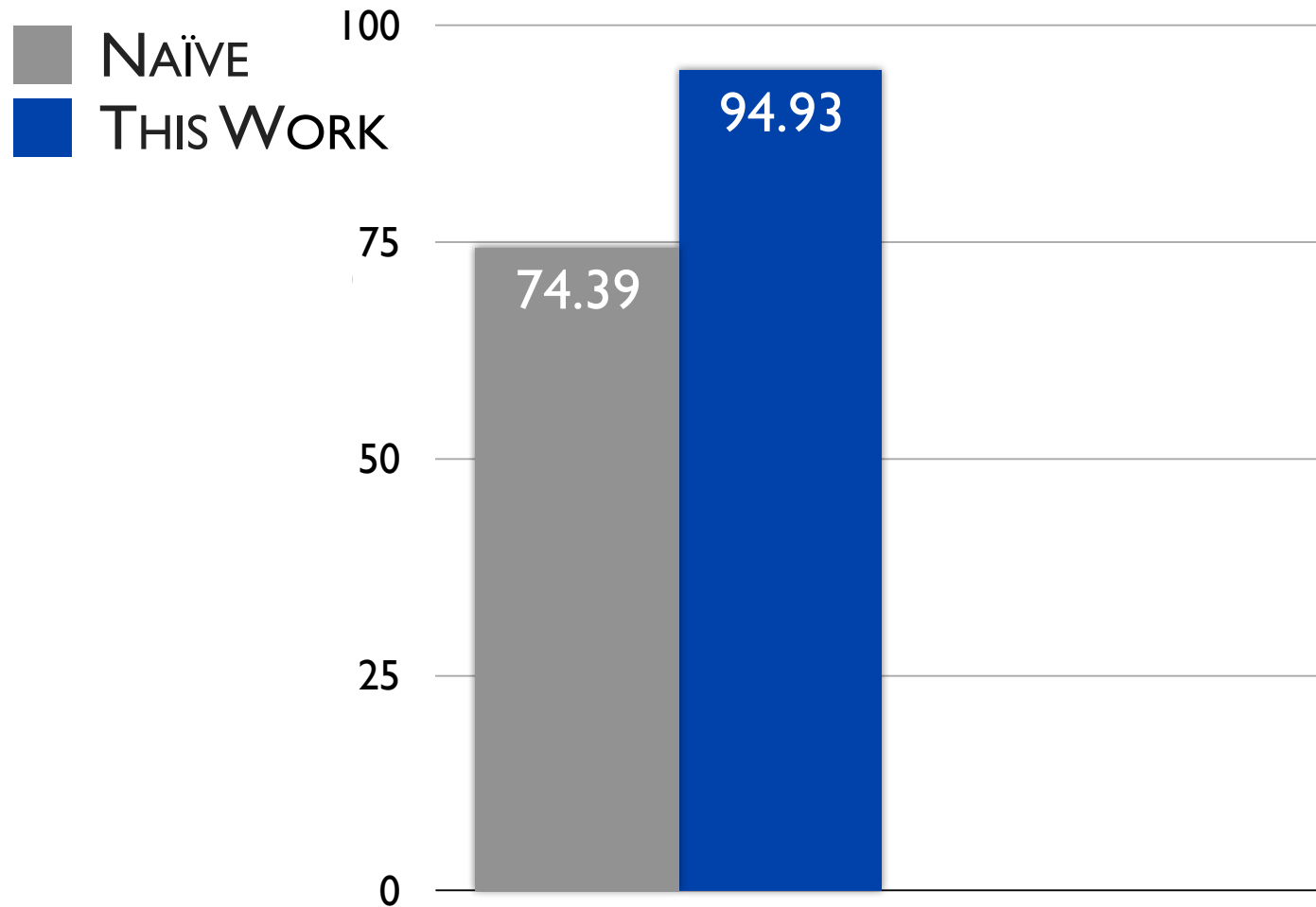
Wiktionary Results

Average inflected form accuracy



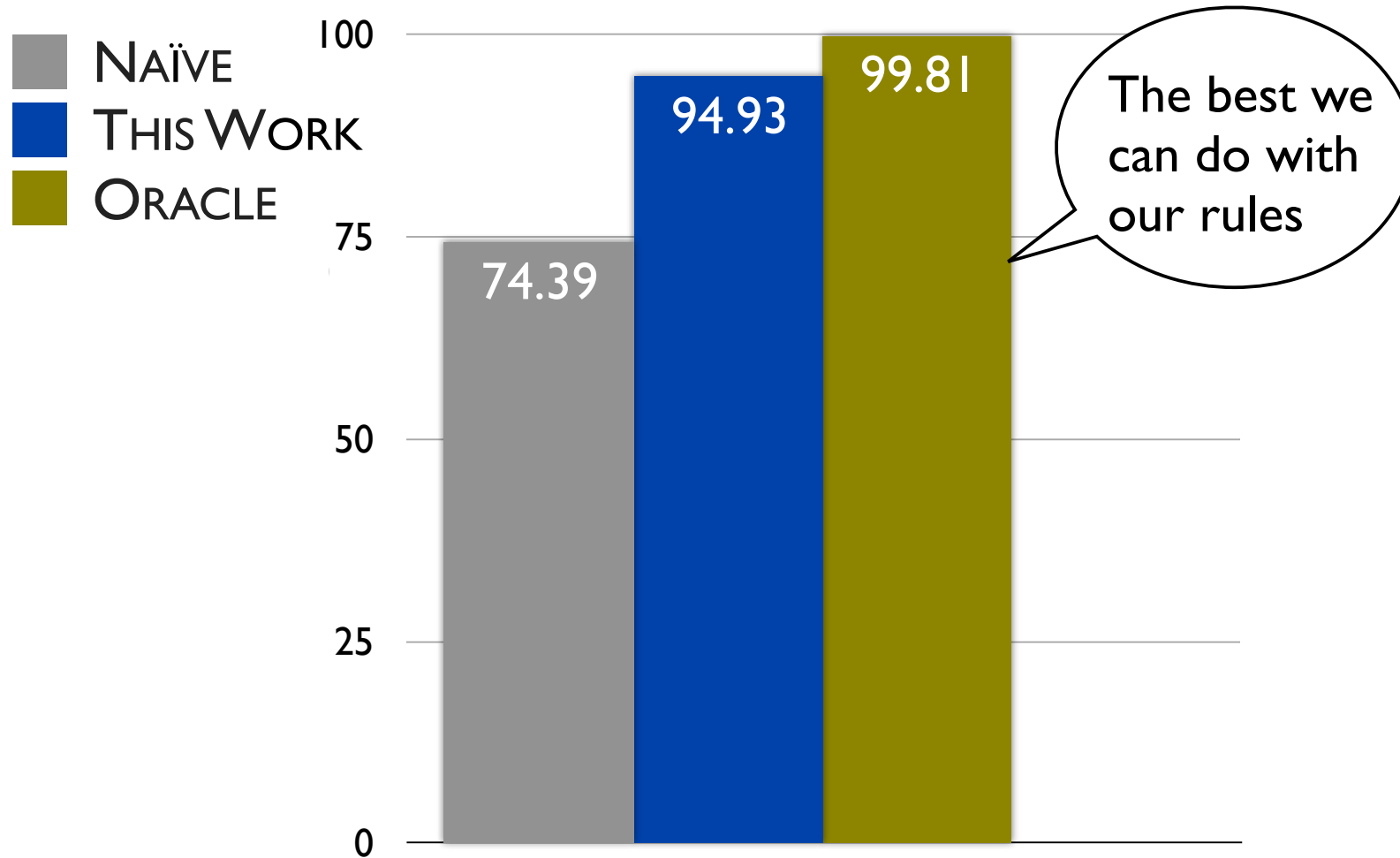
Wiktionary Results

Average inflected form accuracy



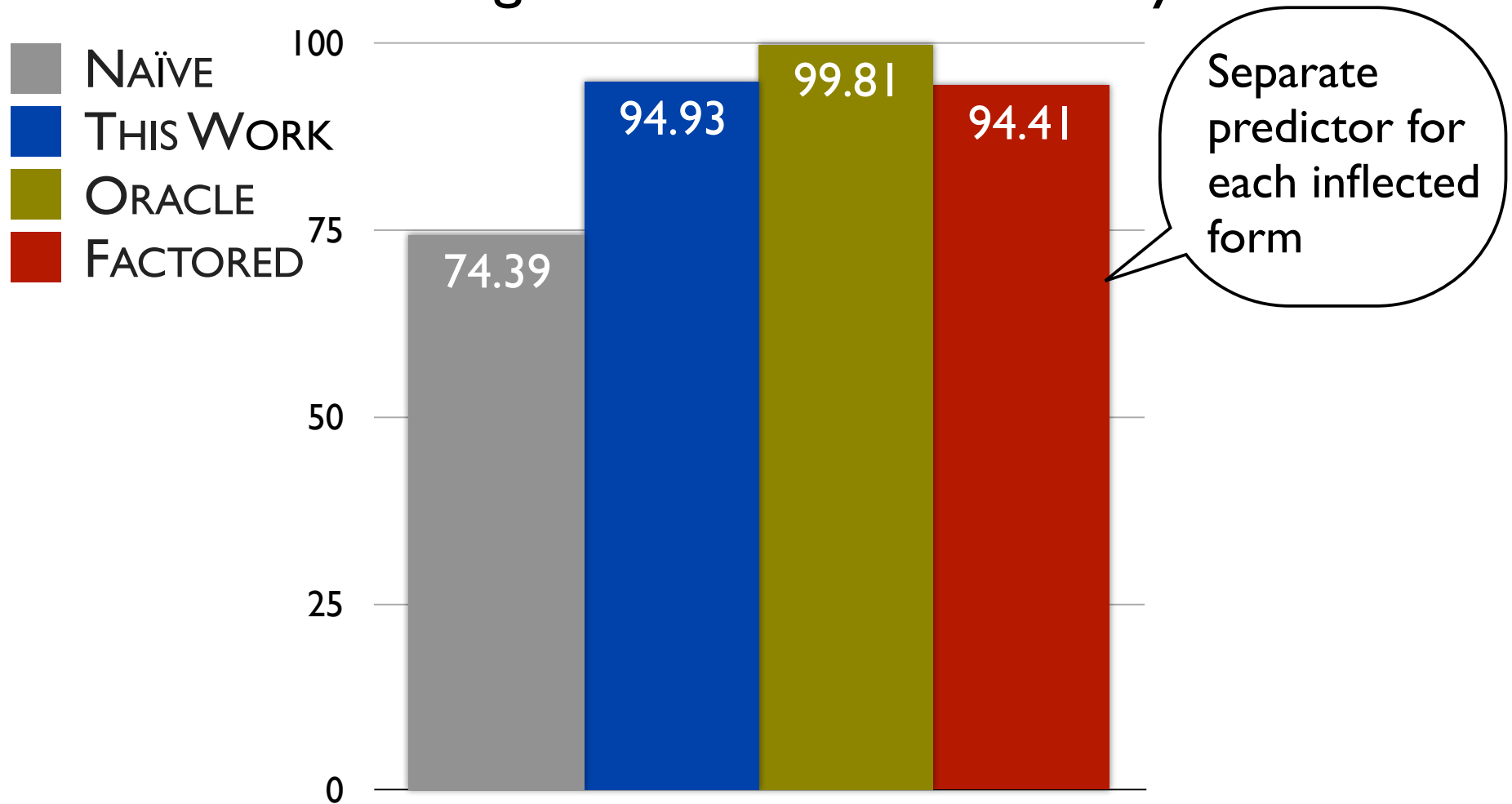
Wiktionary Results

Average inflected form accuracy



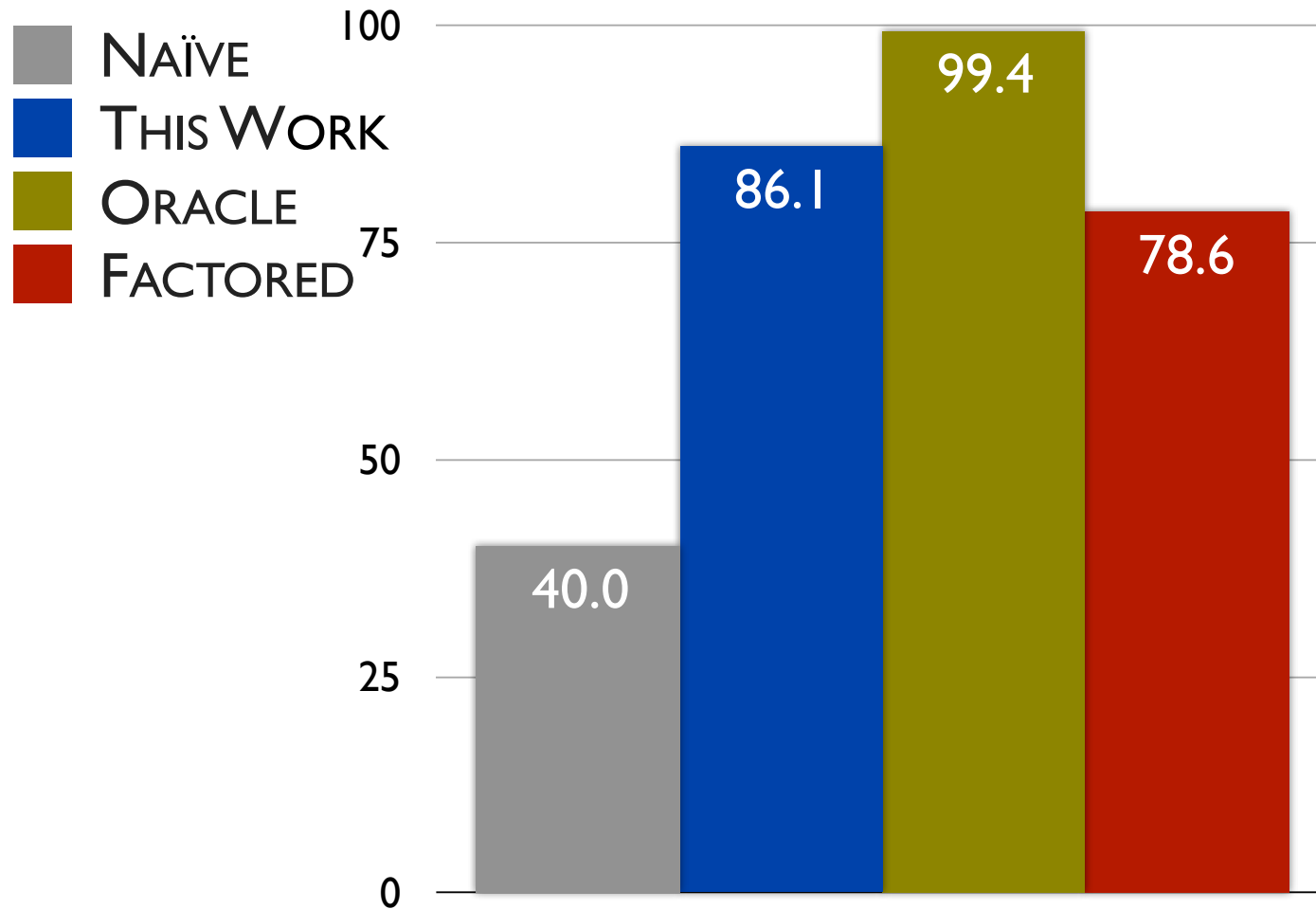
Wiktionary Results

Average inflected form accuracy



Wiktionary Results

Average whole table accuracy



Dreyer and Eisner (2011)

Dreyer and Eisner (2011)

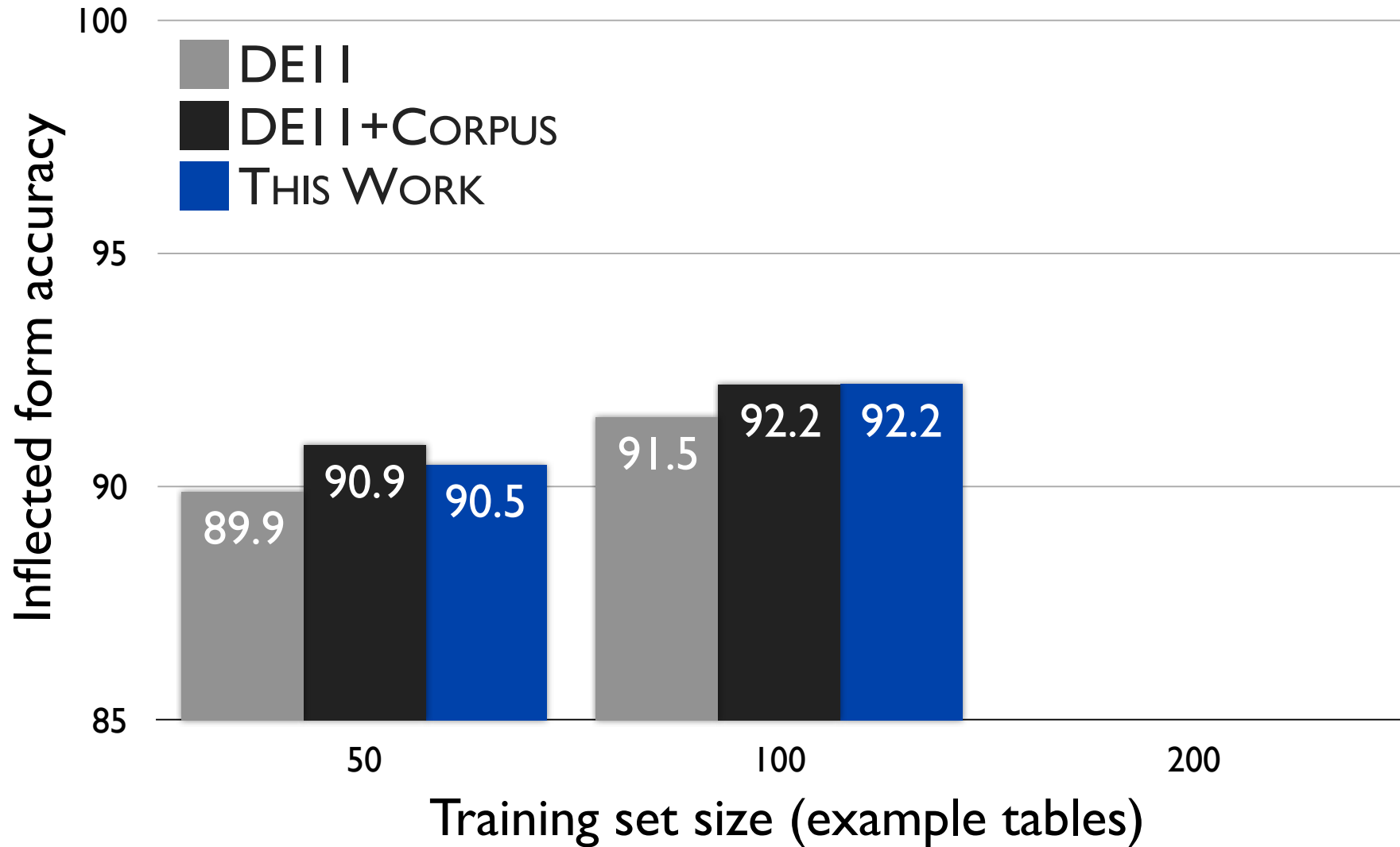
- ▶ Evaluate on German verbs in CELEX
 - ▶ Results averaged over 10 random train/test splits
 - ▶ Small train sets (50 or 100 observed tables)
 - ▶ Large test sets (5415 verbs)

Dreyer and Eisner (2011)

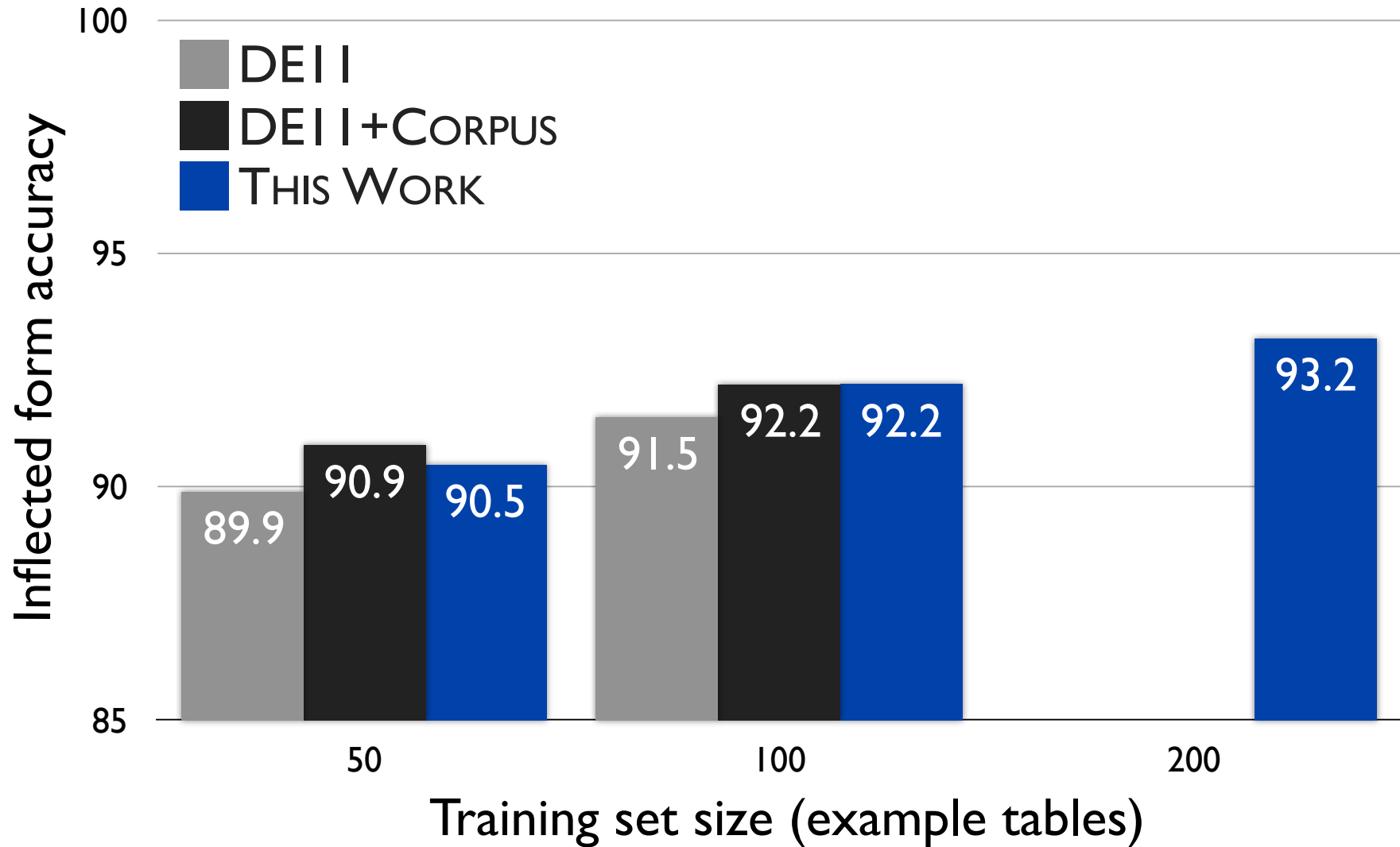
- ▶ Evaluate on German verbs in CELEX
 - ▶ Results averaged over 10 random train/test splits
 - ▶ Small train sets (50 or 100 observed tables)
 - ▶ Large test sets (5415 verbs)

- ▶ Hierarchical Bayesian model of inflection
 - ▶ Type-level transducers of variants can be trained in a supervised fashion
 - ▶ Additionally incorporate unlabeled text with a token-level model

CELEX Results



CELEX Results



Conclusion

- ▶ Morphological inflection rules can be learned from supervised data, which is widely available

Conclusion

- ▶ Morphological inflection rules can be learned from supervised data, which is widely available
- ▶ Structured prediction of entire tables at once is effective for inflecting unseen base forms

Conclusion

- ▶ Morphological inflection rules can be learned from supervised data, which is widely available
- ▶ Structured prediction of entire tables at once is effective for inflecting unseen base forms
- ▶ Code and Wiktionary data is available at <http://eecs.berkeley.edu/~gdurrett>

Conclusion

- ▶ Morphological inflection rules can be learned from supervised data, which is widely available
- ▶ Structured prediction of entire tables at once is effective for inflecting unseen base forms
- ▶ Code and Wiktionary data is available at <http://eecs.berkeley.edu/~gdurrett>

Thank you!

Accuracy Breakdown

