

# An Encoder-Decoder Approach to the Paradigm Cell-Filling Problem

**Miikka Silfverberg**

University of Helsinki  
University of Colorado

Mans Hulden  
University of Colorado

# Paradigm Cell-Filling Problem

Partial inflection table

V;INF	speak
V;PRES	?
V;PAST	?
V;PAST;PCPLE	spoken
V;PRES;PCPLE	?

Farrell Ackerman, James P. Blevins, and Robert Malouf. 2009. Parts and wholes: Implicative patterns in inflectional paradigms. In James P. Blevins and Juliette Blevins, editors, *Analogy in Grammar*, pages 54–82. Oxford University Press.

# Paradigm Cell-Filling Problem

Partial inflection table

V;INF	speak
V;PRES	?
V;PAST	?
V;PAST;PCPLE	spoken
V;PRES;PCPLE	?

⇒

Completed inflection table

V;INF	speak
V;PRES	speaks
V;PAST	spoke
V;PAST;PCPLE	spoken
V;PRES;PCPLE	speaking

Farrell Ackerman, James P. Blevins, and Robert Malouf. 2009. Parts and wholes: Implicative patterns in inflectional paradigms. In James P. Blevins and Juliette Blevins, editors, *Analogy in Grammar*, pages 54–82. Oxford University Press.

# Paradigm Cell-Filling Problem

koirakaan, koirankaan, koiraakaan, koirassakaan, koirastakaan, koiraankaan, koirallakaan, koiraltakaan, koirallekaan,  
koiranakaan, koiraksikaan, koirattakaan, koirineenkaan, koirinkaan, koirako, koiranko, koiraako, koirassako, koirastako,  
koiraanko, koirallako, koiraltako, koiralleko, koiranako, koiraksiko, koirattako, koirineenko, koirinko, koirasikaan,  
koiranikaan, koiransakaan, koirammekaan, koirannekaan, koiraanikaan, koiraasikaan, koiraansakaan, koiraammekaan,  
koiraannekaan, koirassanikaan, koirassasikaan, koirassansakaan, koirassammekaan, koirassannekaan, koirastanikaan,  
koirastasikaan, koirastansakaan, koirastammekaan, koirastannekaan, koirallanikaan, koirallasikaan, koirallansakaan,  
koirallammekaan, koirallannekaan, koirananikaan, koiranasikaan, koiranansakaan, koiranammekaan, koiranannekaan,  
koiraksenikaan, koiraksesikaan, koiraksensakaan, koiraksemmekaan, koiraksennekaan, koirattanikaan, koirattasikaan,  
koirattansakaan, koirattammekaan, koirattannekaan, koirinenikaan, koirinesikaan, koirinensakaan, koirinemmekaan,  
koirinennekaan, koirasiko, koiraniko, koiransakokaan, koiranneko, koiraaniko, koiraasiko, koiraansako,  
koiraammeko, koiraanneko, koirassaniko, koirassansako, koirassammeko, koirassanneko,  
koirastaniko, koirastasiko, koirastansakokaan, koirastanneko, koirallaniko, koirallasiko,  
koirallansako, koirallammeko, koirallanneko, koirallansakokaan, koiranammeko, koirananneko,  
koirakseniko, koiraksesiko, koiraksensakokaan, koirattaniko, koirattasiko, koirattansako,  
koirattammeko, koirattanneko, koirattansakokaan, koirinensako, koirinemmeko,  
koirinenneko, koirasikaanko, koiranikaanko, koiransakokaan, koirannekaanko, koiraanikaanko, koiraansakokaan, koiraammekaanko,  
koiraannekaanko, koirassanikaanko, koirassansakokaan, koirassammekaanko, koirassannekaanko,  
koirastanikaanko, koirastasikaanko, koirastansakokaan, koirastannekaanko, koirallanikaanko,  
koirallammekaanko, koirallannekaanko, koirallansakokaan, koiranammekaanko, koiranannekaanko,  
koiranikaanko, koiranasikaanko, koiranansakokaan, koiranammekaanko, koiranannekaanko,  
koiraksenikaanko, koiraksesikaanko, koiraksensakokaan, koiraksemmekokaan, koiraksennekaanko, koirattasikaanko, koirattansakaa  
koirattammekaanko, koirattannekaanko, koirattansakokaan, koirinensikaanko, koirinensakokaan,  
koirinemmekokaan, koirinennekaanko, koiranikokaan, koiransakokaan, koirannekokaan, koiraansakokaan, koiraammekokaan,  
koiraannekokaan, koirassanikokaan, koirassammekokaan, koirassannekokaan,  
koirastanikokaan, koirastasikokaan, koirastansakokaan, koirastannekokaan, koirallanikokaan,  
koirallasikokaan, koirallansakokaan, koirallammekokaan, koirallannekokaan, koirananikokaan, koiranasikokaan,  
koiranansakokaan, koiranammekokaan, koiranannekokaan, koiraksenikokaan, koiraksesikokaan, koiraksensakokaan,  
koiraksemmekokaan, koiraksennekokaan, koirattanikokaan, koirattasikokaan, koirattansakokaan, koirattammekokaan,  
koirattannekokaan, koirinenikokaan, koirinesikokaan, koirinensakokaan, koirinemmekokaan...



Finnish  
“dog”

<https://herrkramski.wordpress.com/2014/11/06/finnish-is-not-a-hard-language-pt-1/>

# Contribution

1. We investigate PCFP in three different settings.
2. We present two neural models for the PCFP task.
3. We present new data sets for PCFP.



# Three Settings

# Three Settings

$n > 1$

walk	INF
walks	PRES
?	PAST
?	PCPLE
?	INF
smiles	PRES
?	PAST
smiling	PCPLE

# Three Settings

$n > 1$

$n = 1$

walk	INF
walks	PRES
?	PAST
?	PCPLE

?	INF
smiles	PRES
?	PAST
smiling	PCPLE

?	INF
?	PRES
walked	PAST
?	PCPLE

?	INF
smiles	PRES
?	PAST
?	PCPLE



# Three Settings

$n > 1$

$n = 1$

by frequency

walk	INF
walks	PRES
?	PAST
?	PCPLE

?	INF
smiles	PRES
?	PAST
smiling	PCPLE

?	INF
?	PRES
walked	PAST
?	PCPLE

?	INF
smiles	PRES
?	PAST
?	PCPLE

walk	INF
walks	PRES
?	PAST
walking	PCPLE

smile	INF
smiles	PRES
?	PAST
?	PCPLE

# General Approach

- In the  $n > 1$  and “frequent words” settings we train a LSTM encoder-decoder model with attention.

# General Approach

- In the  $n > 1$  and “frequent words” settings we train a LSTM encoder-decoder model with attention.

# General Approach

- In the  $n > 1$  and “frequent words” settings we train a LSTM encoder-decoder model with attention.
- In the  $n = 1$  setting, we apply adaptive character dropout and then train an encoder-decoder.

# Training for $n > 1$ Given Forms

Input

Output

walk	INF
walks	PRES
?	PAST
?	PCPLE
?	INF
smiles	PRES
?	PAST
smiling	PCPLE

# Training for $n > 1$ Given Forms

Input

Output

walk	INF
walks	PRES
?	PAST
?	PCPLE

?	INF
smiles	PRES
?	PAST
smiling	PCPLE

# Training for $n > 1$ Given Forms

walk	INF
walks	PRES
?	PAST
?	PCPLE
?	INF
smiles	PRES
?	PAST
smiling	PCPLE

Input

walk

Output

# Training for $n > 1$ Given Forms

walk	INF
<b>walks</b>	<b>PRES</b>
?	PAST
?	PCPLE
?	INF
smiles	PRES
?	PAST
smiling	PCPLE

Input  
walk+INF

Output



# Training for $n > 1$ Given Forms

walk	INF
<b>walks</b>	<b>PRES</b>
?	PAST
?	PCPLE
?	INF
smiles	PRES
?	PAST
smiling	PCPLE

Input

walk+INF>PRES

Output

# Training for $n > 1$ Given Forms

walk	INF
walks	PRES
?	PAST
?	PCPLE

?	INF
smiles	PRES
?	PAST
smiling	PCPLE

Input

walk+INF>PRES

Output

walks

# Training for $n > 1$ Given Forms

walk	INF
walks	PRES
?	PAST
?	PCPLE

?	INF
smiles	PRES
?	PAST
smiling	PCPLE

Input

walk+INF>PRES

walks+PRES>INF

Output

walks

walk

# Training for $n > 1$ Given Forms

walk	INF
walks	PRES
?	PAST
?	PCPLE

?	INF
smiles	PRES
?	PAST
smiling	PCPLE

Input	Output
walk+INF>PRES	walks
walks+PRES>INF	walk
smiles+PRES>PCPLE	smiling
smiling+PCPLE>PRES	smiles

# Training for $n > 1$ Given Forms

walk	INF
walks	PRES
?	PAST
?	PCPLE

?	INF
smiles	PRES
?	<b>PAST</b>
smiling	PCPLE

Input	Output
walk+INF>PRES	walks
walks+PRES>INF	walk
smiles+PRES>PCPLE	smiling
smiling+PCPLE>PRES	smiles

# Training for $n > 1$ Given Forms

walk	INF
walks	PRES
?	PAST
?	PCPLE

?	INF
smiles	PRES
?	<b>PAST</b>
smiling	PCPLE

Input	Output
walk+INF>PRES	walks
walks+PRES>INF	walk
smiles+PRES>PCPLE	smiling
smiling+PCPLE>PRES	smiles
smiles+PRES>PAST	?
smiling+PCPLE>PAST	?

# Training for $n > 1$ Given Forms

walk	INF
walks	PRES
?	PAST
?	PCPLE

?	INF
smiles	PRES
?	<b>PAST</b>
smiling	PCPLE

Input	Output
walk+INF>PRES	walks
walks+PRES>INF	walk
smiles+PRES>PCPLE	smiling
smiling+PCPLE>PRES	smiles
smiles+PRES>PAST	smiled
smiling+PCPLE>PAST	smiled

# Training for $n > 1$ Given Forms

walk	INF
walks	PRES
?	PAST
?	PCPLE

?	INF
smiles	PRES
<b>smiled</b>	<b>PAST</b>
smiling	PCPLE

Input	Output
walk+INF>PRES	walks
walks+PRES>INF	walk
smiles+PRES>PCPLE	smiling
smiling+PCPLE>PRES	smiles
smiles+PRES>PAST	smiled
smiling+PCPLE>PAST	smiled
↓	
smiled	



# Training for $n > 1$ Given Forms

walk	INF
walks	PRES
?	PAST
?	PCPLE

?	INF
smiles	PRES
<b>smiled</b>	<b>PAST</b>
smiling	PCPLE

Input	Output
walk+INF>PRES	walks
walks+PRES>INF	walk
smiles+PRES>PCPLE	smiling
smiling+PCPLE>PRES	smiles

smiles+PRES>PAST	smiled
smiling+PCPLE>PAST	smiled



smiled

Katharina Kann and Hinrich Schütze. 2016. Single-model encoder-decoder with explicit morphological representation for reinflection. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 555–560, Berlin, Germany. Association for Computational Linguistics.

# Case $n = 1$

?	INF
?	PRES
walked	PAST
?	PCPLE
?	INF
smiles	PRES
?	PAST
?	PCPLE

Input

Output

# Case n = 1

?	INF
?	PRES
walked	PAST
?	PCPLE

?	INF
smiles	PRES
?	PAST
?	PCPLE

Input

walked+PAST>PAST

smiles+PRES>PRES

Output

walked

smiles

# Case n = 1

?	INF
?	PRES
walked	PAST
?	PCPLE
?	INF
smiles	PRES
?	PAST
?	PCPLE

Input	Output
walked+PAST>PAST	walked
smiles+PRES>PRES	smiles
walked+PAST>PRES	?

# Case n = 1

?	INF
?	PRES
walked	PAST
?	PCPLE

?	INF
smiles	PRES
?	PAST
?	PCPLE

Input	Output
<del>walked+PAST&gt;PAST</del>	<del>walked</del>
<del>smiles+PRES&gt;PRES</del>	<del>smiles</del>

walked+PAST>PRES	?
------------------	---

# What about Stemming?

walked+PAST>PAST	walked
smiles+PRES>PRES	smiles

# What about Stemming?

walked+PAST>PAST      walked

smiles+PRES>PRES      smiles

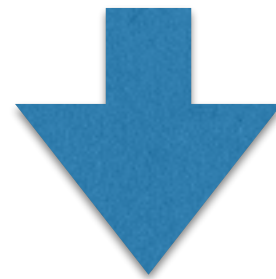
# What about Stemming?

walked+PAST>PAST

walked

smiles+PRES>PRES

smiles



walk>PAST

walked

smile>PRES

smiles



# Probabilistic Stemming Using a Language Model

**PAST** forms:

walked

slept

heard

missed

loved

ate

taped

# Probabilistic Stemming Using a Language Model

**PAST** forms:

walk**ed**

slept**t**

heard**d**

miss**ed**

lov**ed**

ate

tap**ed**

# Probabilistic Stemming Using a Language Model

**PAST** forms:

walk**ed**

slept**t**

heard**d**

miss**ed**

lov**ed**

ate

tap**ed**

- We train a character language model for each label (INF, PRES, PAST,...)

# Probabilistic Stemming Using a Language Model

**PAST** forms:

walk**ed**

slept**t**

heard**d**

miss**ed**

lov**ed**

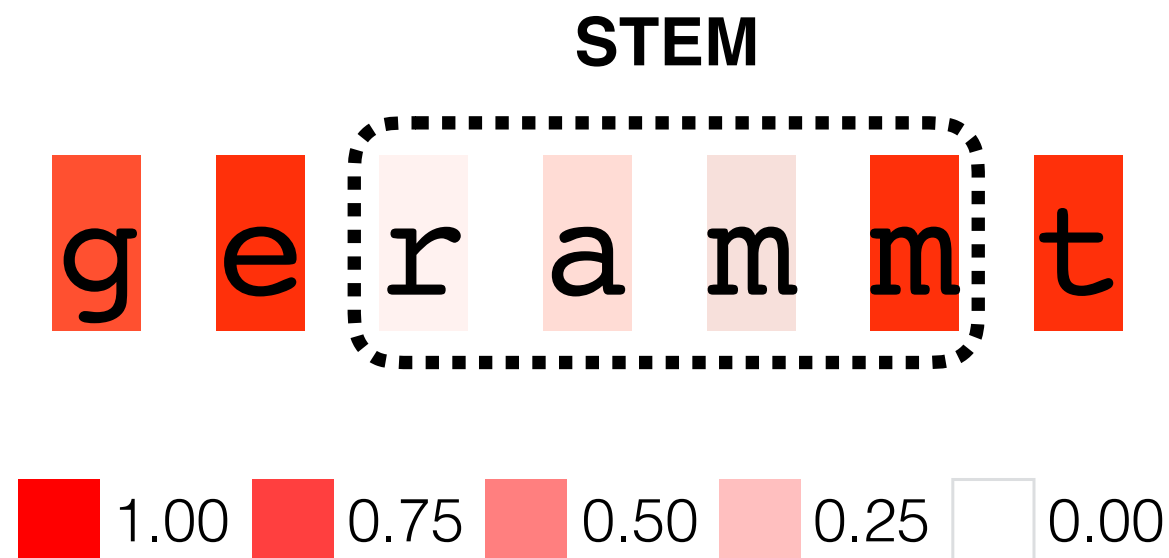
ate

tap**ed**

- We train a character language model for each label (INF, PRES, PAST,...)
- Language model confidence is used for identifying word stems and affixes.

# Language Model Confidences

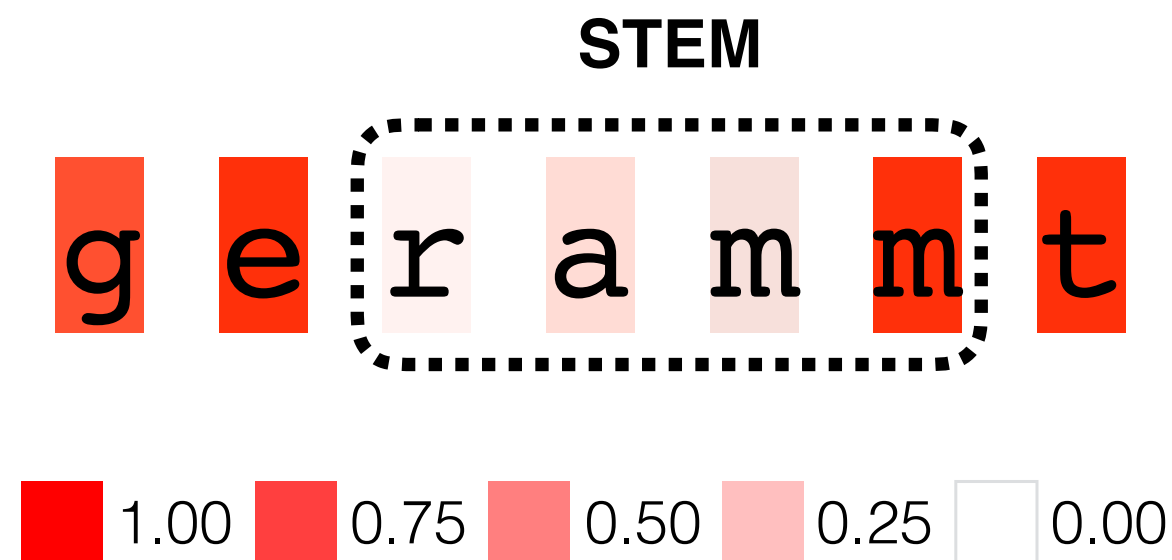
Confidence for predicting character  $x_t$  based on  $x_1, \dots, x_{t-1}$ :



German  
"rammed"

# Language Model Confidences

Confidence for predicting character  $x_t$  based on  $x_1, \dots, x_{t-1}$ :



German  
“rammed”

Characters belonging to the word stem typically have **low** language model confidence.

# Adaptive Character Dropout

During training we drop characters from the input form based on  $LM_{PAST}$  confidence:

walked+PAST

# Adaptive Character Dropout

During training we drop characters from the input form based on  $LM_{PAST}$  confidence:

walk  +PAST



# Adaptive Character Dropout

During training we drop characters from the input form based on  $LM_{PAST}$  confidence:

walk  +PAST

Because it is problematic to determine which characters belong to the stem, we drop characters probabilistically.

# Training Examples Before Character Dropout

résigner>INF -> résigner  
résigner>INF -> résigner  
résigner>INF -> résigner  
résigner>INF -> résigner  
résigner>INF -> résigner  
résigner>INF -> résigner  
résigner>INF -> résigner  
résigner>INF -> résigner  
résigner>INF -> résigner  
résigner>INF -> résigner

French  
“resign”

# Training Examples After Character Dropout

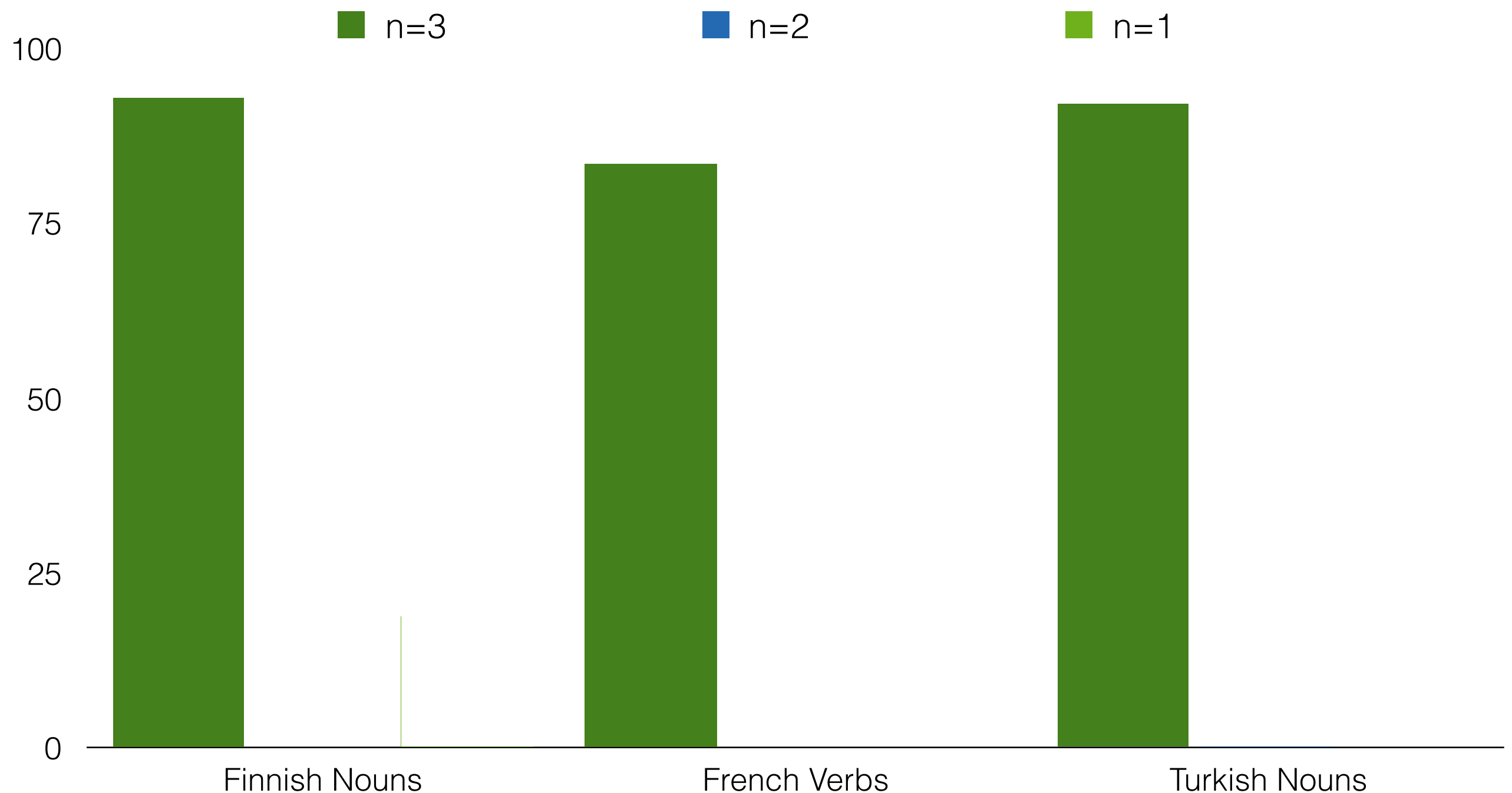
ré	i	>INF	->	résigner
r	s	>INF	->	résigner
rés	g	>INF	->	résigner
résign		>INF	->	résigner
résig		>INF	->	résigner
r	si	>INF	->	résigner
rés	i	>INF	->	résigner
r	si	>INF	->	résigner
rés	g	e	>INF	résigner
résig		>INF	->	résigner

French  
“resign”

# Data Used in Experiments

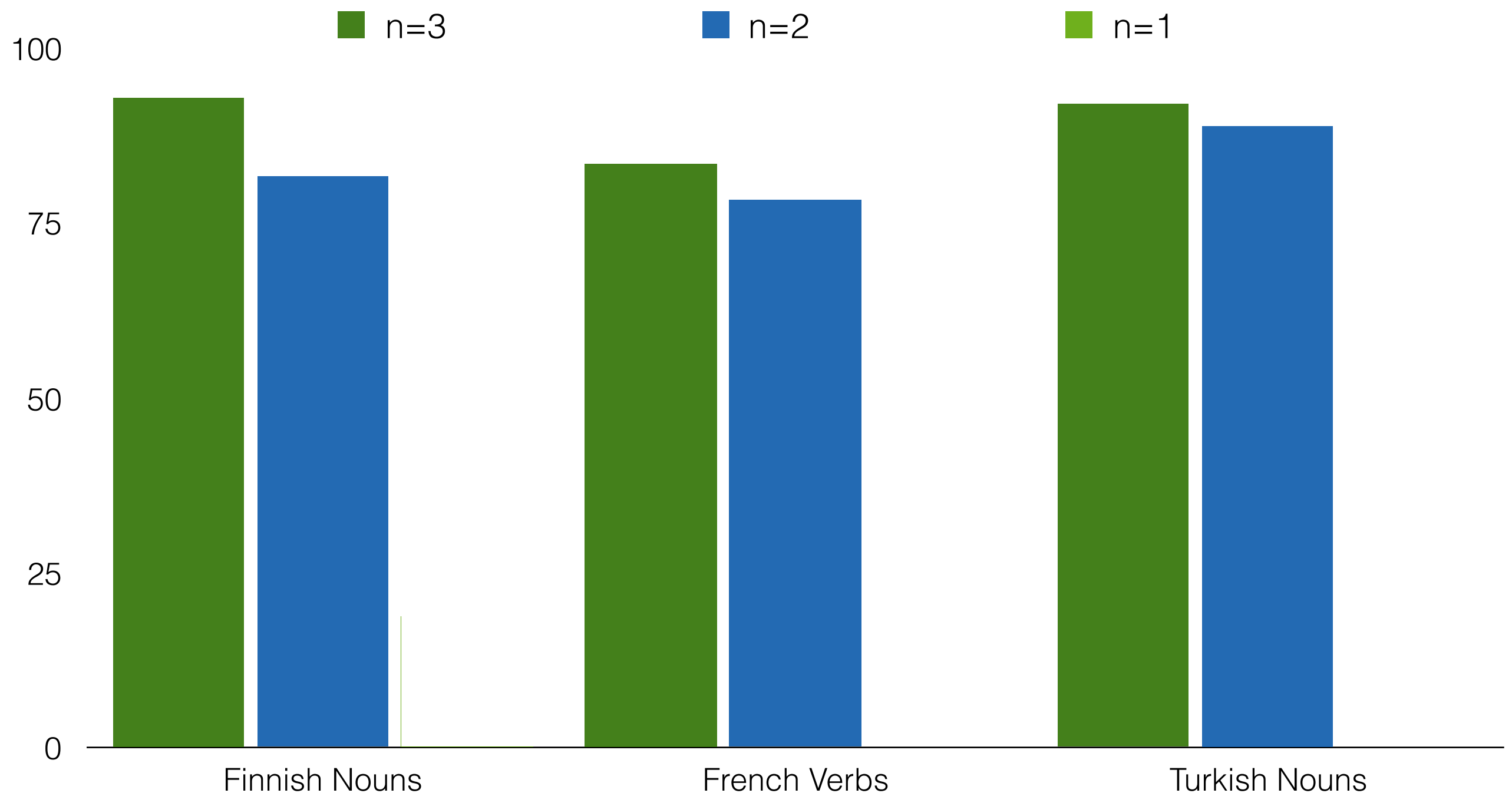
We present experiments for noun and verb tables for:  
Finnish, French, Georgian, German, Latin, Latvian,  
Spanish and Turkish.

# Results



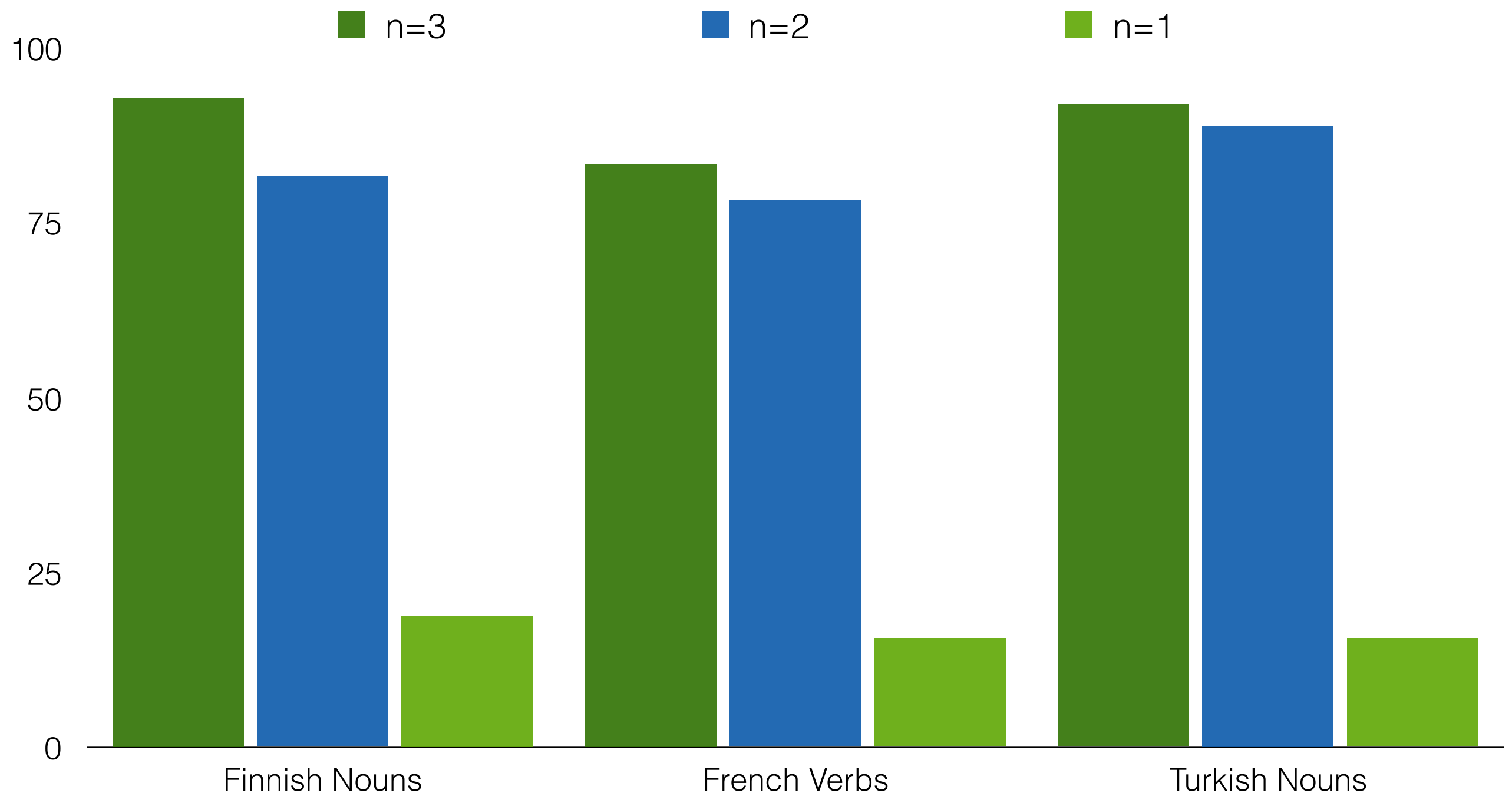
We use 1000 inflection tables for each language.

# Results



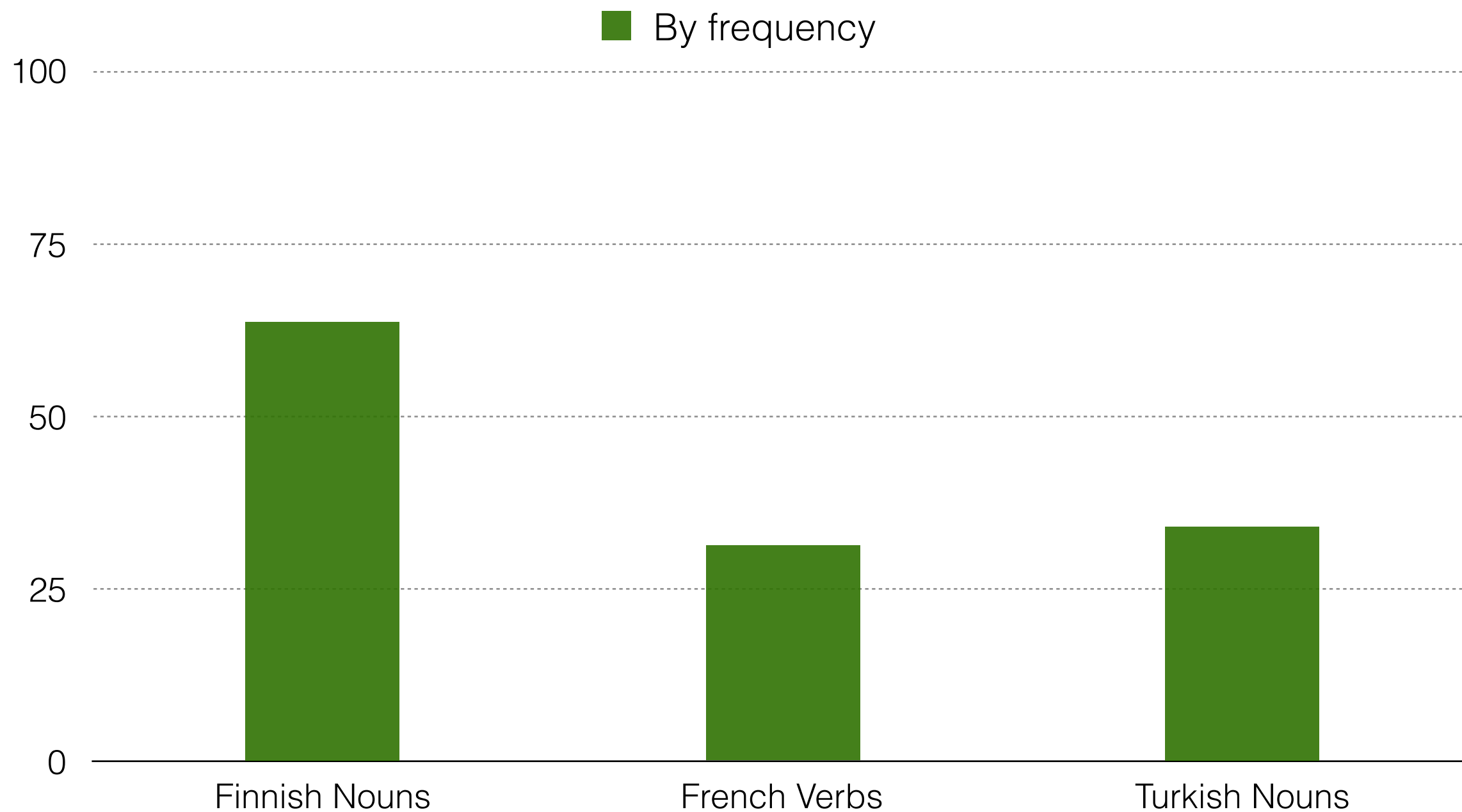
We use 1000 inflection tables for each language.

# Results



We use 1000 inflection tables for each language.

# Results





# Why are Results Low When Top 10k Forms are Given?

Difficult to learn rare forms.

# Why are Results Low When Top 10k Forms are Given?

Difficult to learn rare forms.

Major problem: Syncretism

# Why are Results Low When Top 10k Forms are Given?

Difficult to learn rare forms.

Major problem: Syncretism

**Spießer**

N, ACC, PL

**Spießer**

N, ACC, SG

**Spießer**

N, DAT, SG

**Spießer**

N, GEN, PL

N, DAT, PL

**Spießer**

N, NOM, PL

**Spießer**

N, NOM, SG

N, GEN, SG

German  
“philistine”

# Why are Results Low When Top 10k Forms are Given?

Difficult to learn rare forms.

Major problem: Syncretism

<b>Spießer</b>	N, ACC, PL
<b>Spießer</b>	N, ACC, SG
<b>Spießer</b>	N, DAT, SG
<b>Spießer</b>	N, GEN, PL
Spießern	N, DAT, PL
<b>Spießer</b>	N, NOM, PL
<b>Spießer</b>	N, NOM, SG
Spießers	N, GEN, SG

German  
“philistine”

# Conclusions & Future Work

- We can learn inflectional morphology even when only given one example per lexeme.
- If every table has two or more forms, accuracy is around 90%.
- Predicting rare forms based on frequent ones is difficult.
- Future work: Need more realistic data set for the L1 learning scenario.

# Thank you!