Investigating Inductive Biases of Neural Models via Synthetic Languages

Paula Czarnowska

University of Cambridge, pjc211@cam.ac.uk

DELPH-IN 2021

Overview

1. **Setting the scene**
   - Earlier experiments and a new problem

     Neural models are biased towards relying on word order over morphology when detecting subjects and objects.

2. **The focus of this presentation**

Overview

1. **Setting the scene**
   - Earlier experiments and a new problem

     Neural models are biased towards relying on word order over morphology when detecting subjects and objects.

2. **The focus of this presentation**

     What properties of a language 'encourage' or 'discourage' the models from relying on morphology?

PART 1: Setting the scene…

Linguistic indicators of `who did what to whom'

- Languages signal subject/objecthood in different ways…

Ada is lifting the cups

- English:

Linguistic indicators of `who did what to whom'

- Languages signal subject/objecthood in different ways…

Ada is lifting the cups

S V O

- English: word order

Linguistic indicators of `who did what to whom'

- Languages signal subject/objecthood in different ways…

Ada	**is**	lifting	the	cups
sg	3.sg			pl

- English: word order, verbal agreement

Linguistic indicators of `who did what to whom'

- Languages signal subject/objecthood in different ways…

She is lifting the cups

nom

- English: word order, verbal agreement, case markings (pronouns)

Linguistic indicators of `who did what to whom’

- Languages signal subject/objecthood in different ways…

Ada podnosi kubki

- English: word order, verbal agreement, case markings (pronouns)
- Polish:

Linguistic indicators of `who did what to whom'

- Languages signal subject/objecthood in different ways…

Ada podno**si** kubki

sg 3.sg pl

- English: word order, verbal agreement, case markings (pronouns)
- Polish: verbal agreement

Linguistic indicators of `who did what to whom'

- Languages signal subject/objecthood in different ways…

Ad**a** podnosi kub**ki**

nom ~~nom~~/acc

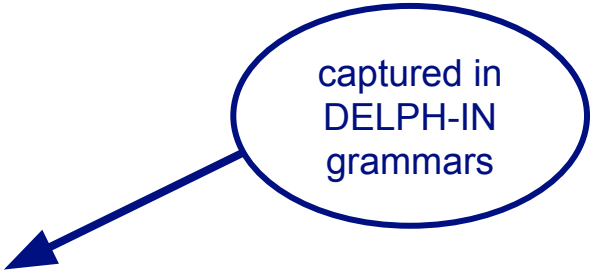
- English: word order, verbal agreement, case markings (pronouns)
- Polish: verbal agreement, case markings

Linguistic indicators of `who did what to whom'

- Languages signal subject/objecthood in different ways…

Ad**a** podnosi kub**ki**

nom ~~nom~~/acc



- English: word order, verbal agreement, case markings (pronouns)
- Polish: verbal agreement, case markings

Statistical correlates

- Features that do not encode subject/objecthood in a language may correlate with grammatical function.

Statistical correlates

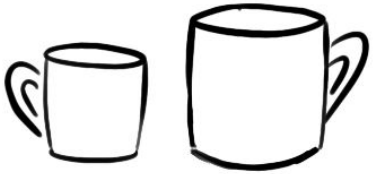
- Features that do not encode subject/objecthood in a language may correlate with grammatical function.

Ada podnosi kubki

Statistical correlates

- Features that do not encode subject/objecthood in a language may correlate with grammatical function.

Ada podnosi kubki



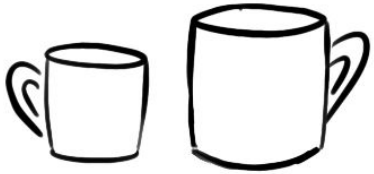
- Polish: lexical semantics

Statistical correlates

- Features that do not encode subject/objecthood in a language may correlate with grammatical function.

Ada podnosi kubki

S V O



- Polish: lexical semantics, word order

Earlier experiments

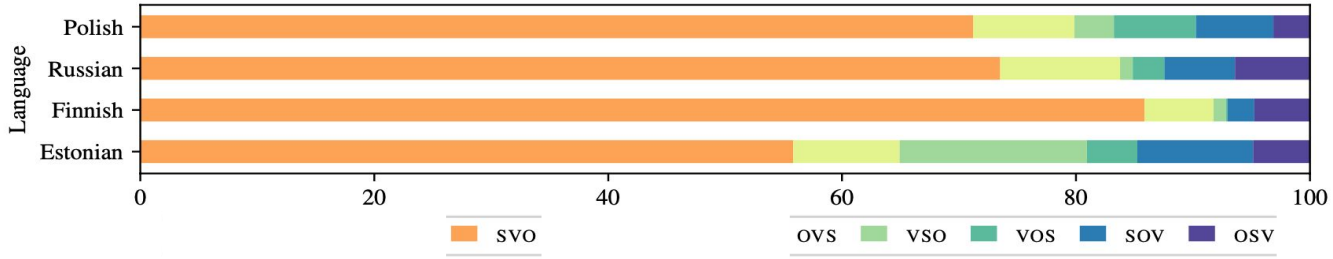
- Do neural models recognize which features are the `true indicators' and which are the correlates?

- The prominent paradigm in NN research is to feed only unprocessed text as input.

Earlier experiments: Morphology vs word order

- **Task:** dependency parsing
- **Languages:** Polish, Russian, Finnish, Estonian

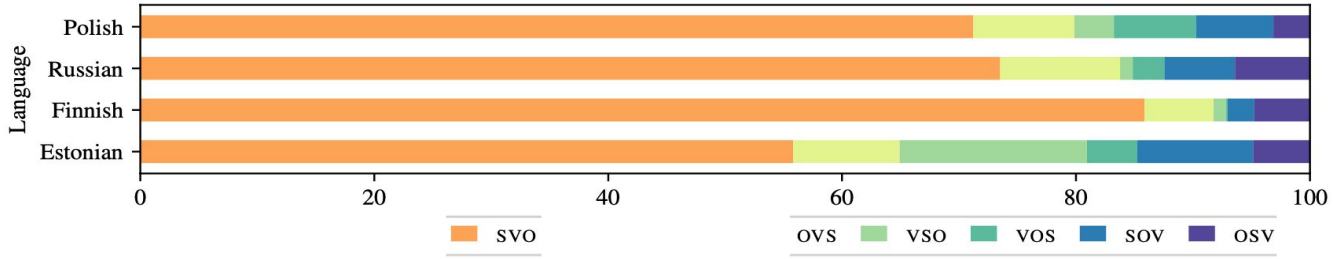
  All allow for flexible word order but the word order distribution is heavily skewed towards SVO.



Earlier experiments: Morphology vs word order

- **Task:** dependency parsing
- **Languages:** Polish, Russian, Finnish, Estonian

  All allow for flexible word order but the word order distribution is heavily skewed towards SVO.



- When parsing Polish, Russian, Finnish and Estonian...

  *...do neural dependency parsers correctly **generalize primarily based on morphology** rather than word order?*

Morphology vs word order: Methodology

- Evaluating on reordered transitive sentences from the UD treebank:

SVO	Ada podnosi kubki.
SOV	Ada kubki podnosi.
OVS	Kubki podnosi Ada.
OSV	Kubki Ada podnosi.
VSO	Podnosi Ada kubki.
VOS	Podnosi kubki Ada.

Morphology vs word order: Methodology

- Evaluating on reordered transitive sentences from the UD treebank:

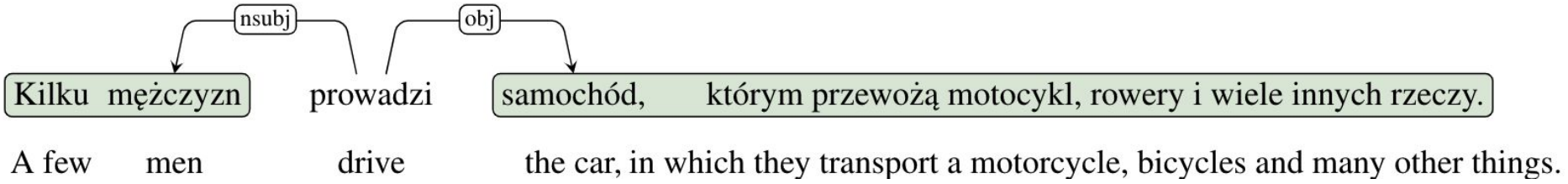
SVO	Ada podnosi kubki.
SOV	Ada kubki podnosi.
OVS	Kubki podnosi Ada.
OSV	Kubki Ada podnosi.
VSO	Podnosi Ada kubki.
VOS	Podnosi kubki Ada.

One counterfactual treebank for every word order.

Morphology vs word order: Methodology

Three steps to create a reordered treebank:

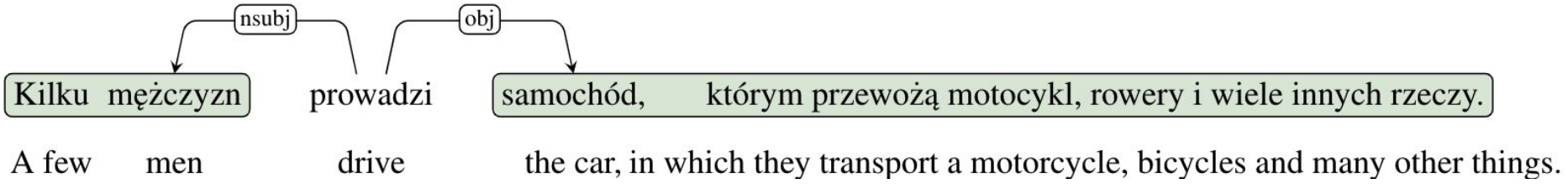
1. Filter out all non-transitive sentences.
2. Simplify sentences to maintain acceptability after the reordering.
3. Reorder.

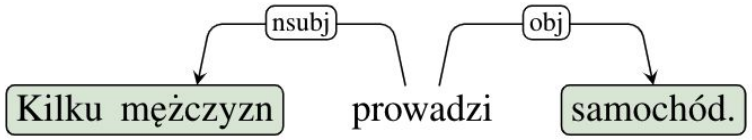


Morphology vs word order: Methodology

Three steps to create a reordered treebank:

1. Filter out all non-transitive sentences.
2. Simplify sentences to maintain acceptability after the reordering.
3. Reorder.

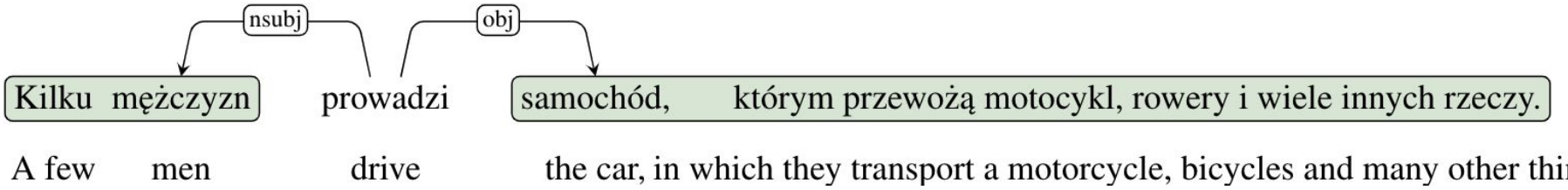


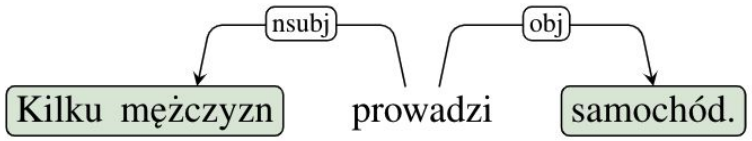


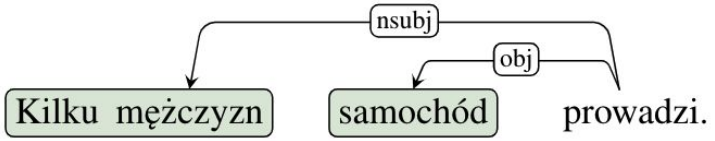
Morphology vs word order: Methodology

Three steps to create a reordered treebank:

1. Filter out all non-transitive sentences.
2. Simplify sentences to maintain acceptability after the reordering.
3. Reorder.

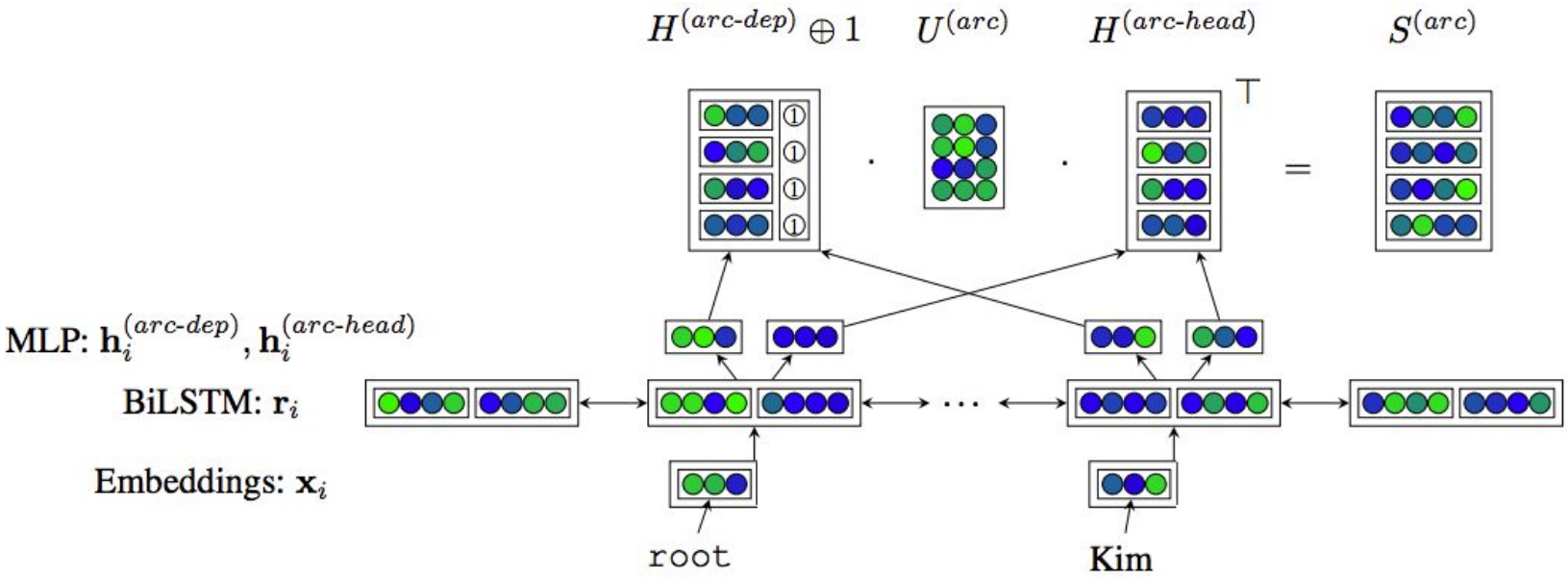






Experimental Details

- **Data**: Universal Dependencies (UD) Treebanks
- **Model:** (Dozat and Manning, 2017)



Timothy Dozat and Christopher D Manning. 2017. Deep Biaffine Attention for Neural Dependency Parsing. ICLR (2017)

Experimental Details

- **Data**: Universal Dependencies (UD) Treebanks
- **Model:** (Dozat and Manning, 2017)
- **Inputs:**
  - CNN over characters (Kim et al., 2016)
  - BERT (Devlin et al., 2019)
  - fastText (Bojanowski et al, 2017)

Timothy Dozat and Christopher D Manning. 2017. Deep Biaffine Attention for Neural Dependency Parsing. ICLR (2017).
Yoon Kim, Yacine Jernite, David Sontag, and Alexander M. Rush. 2016. Character-Aware Neural Language Models. AAAI (2016).
Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. NACL (2019).
Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2017. Enriching Word Vectors with Subword Information. TACL (2017).

Morphology vs word order: LAS results

	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	#	all	nsubj	obj	#	all	nsubj	obj	#
	Polish					Russian					Finnish				Estonian			
(Dozat and Manning, 2017) + BERT																		
svo	94.2	99.9	98.3	89.6	*295*	97.6	99.1	99.1	94.3	*632*	97.7	99.7	98.7	*127*	93.6	97.3	96.9	*258*
ovs	91.6	94.1	92.9	79.8	*295*	95.8	94.0	92.5	92.9	*632*	96.5	98.1	99.2	*127*	92.7	94.5	93.6	*258*
sov	90.7	96.2	84.7	71.0	*295*	95.3	94.8	93.1	90.8	*632*	95.4	96.3	93.2	*127*	87.6	86.7	85.8	*258*
osv	91.4	92.4	90.8	80.3	*295*	95.0	91.2	95.0	90.8	*632*	96.2	98.9	94.2	*127*	86.3	89.5	75.3	*258*
vso	84.2	71.8	73.9	70.5	*295*	91.1	78.9	81.9	78.7	*632*	94.9	90.8	96.6	*127*	90.8	89.3	92.7	*258*
vos	88.7	79.4	93.2	84.7	*295*	94.1	81.1	98.0	92.2	*632*	90.2	73.7	88.5	*127*	87.4	82.7	83.1	*258*
(Dozat and Manning, 2017) + CNN																		
svo	88.0	93.6	91.5	77.0	*295*	90.3	86.4	88.6	70.2	*632*	86.4	92.9	84.3	*127*	84.0	86.8	85.7	*258*
ovs	69.9	35.6	43.4	54.1	*295*	81.3	57.6	57.5	59.6	*632*	74.8	69.3	52.0	*127*	75.7	65.1	58.5	*258*
sov	74.2	84.7	34.0	44.3	*295*	81.3	69.9	56.5	44.7	*632*	63.1	63.8	20.5	*127*	67.1	49.2	46.5	*258*
osv	67.7	43.7	37.9	47.5	*295*	79.2	46.7	59.7	66.0	*632*	70.8	88.2	23.6	*127*	66.3	60.1	27.5	*258*
vso	66.6	32.2	43.4	44.3	*295*	75.5	38.0	50.7	48.9	*632*	68.3	48.8	50.4	*127*	73.2	54.7	69.0	*258*
vos	69.7	18.0	79.1	72.1	*295*	80.9	38.9	80.2	66.0	*632*	68.9	41.7	52.0	*127*	72.4	56.6	52.7	*258*

Morphology vs word order: LAS results

	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	#	all	nsubj	obj	#	all	nsubj	obj	#
	Polish					Russian					Finnish				Estonian			
(Dozat and Manning, 2017) + BERT																		
svo	94.2	99.9	98.3	89.6	*295*	97.6	99.1	99.1	94.3	*632*	97.7	99.7	98.7	*127*	93.6	97.3	96.9	*258*
ovs	91.6	94.1	92.9	79.8	*295*	95.8	94.0	92.5	92.9	*632*	96.5	98.1	99.2	*127*	92.7	94.5	93.6	*258*
sov	90.7	96.2	84.7	71.0	*295*	95.3	94.8	93.1	90.8	*632*	95.4	96.3	93.2	*127*	87.6	86.7	85.8	*258*
osv	91.4	92.4	90.8	80.3	*295*	95.0	91.2	95.0	90.8	*632*	96.2	98.9	94.2	*127*	86.3	89.5	75.3	*258*
vso	84.2	71.8	73.9	70.5	*295*	91.1	78.9	81.9	78.7	*632*	94.9	90.8	96.6	*127*	90.8	89.3	92.7	*258*
vos	88.7	79.4	93.2	84.7	*295*	94.1	81.1	98.0	92.2	*632*	90.2	73.7	88.5	*127*	87.4	82.7	83.1	*258*
(Dozat and Manning, 2017) + CNN																		
svo	88.0	93.6	91.5	77.0	*295*	90.3	86.4	88.6	70.2	*632*	86.4	92.9	84.3	*127*	84.0	86.8	85.7	*258*
ovs	69.9	35.6	43.4	54.1	*295*	81.3	57.6	57.5	59.6	*632*	74.8	69.3	52.0	*127*	75.7	65.1	58.5	*258*
sov	74.2	84.7	34.0	44.3	*295*	81.3	69.9	56.5	44.7	*632*	63.1	63.8	20.5	*127*	67.1	49.2	46.5	*258*
osv	67.7	43.7	37.9	47.5	*295*	79.2	46.7	59.7	66.0	*632*	70.8	88.2	23.6	*127*	66.3	60.1	27.5	*258*
vso	66.6	32.2	43.4	44.3	*295*	75.5	38.0	50.7	48.9	*632*	68.3	48.8	50.4	*127*	73.2	54.7	69.0	*258*
vos	69.7	18.0	79.1	72.1	*295*	80.9	38.9	80.2	66.0	*632*	68.9	41.7	52.0	*127*	72.4	56.6	52.7	*258*

What does it mean?

- Not due to `morphology shortage': models are constantly exposed to morphological markings.

- Models are biased towards relying on word order.

What does it mean?

- Not due to `morphology shortage': models are constantly exposed to morphological markings.

- Models are biased towards relying on word order.

Two directions from here:

1. *Practical*: How do we fix it?
2. *Theoretical*: What aspects of a morphological system make it hard for the models to rely on morphology?

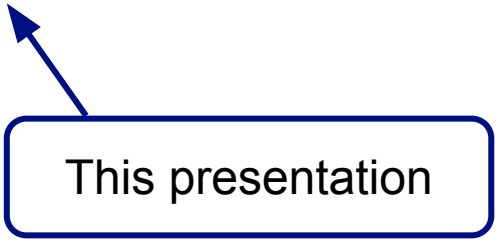
What does it mean?

- Not due to `morphology shortage': models are constantly exposed to morphological markings.

- Models are biased towards relying on word order.

Two directions from here:

1. *Practical*: How do we fix it?
2. *Theoretical*: What aspects of a morphological system make it hard for the models to rely on morphology?



PART 2: What aspects of morphological systems are hard for the NNs?

Evaluating on synthetic versions of a language

The idea:

- Manipulate a particular aspect of a language's morphology.
- Train on the new language.
- Evaluate on reordered treebanks for the new language.

Evaluating on synthetic versions of a language

The idea:

- Manipulate a particular aspect of a language's morphology.
- Train on the new language.
- Evaluate on reordered treebanks for the new language.

Some properties that could be manipulated (Bickel and Nichols, 2007):

Balthasar Bickel and Johanna Nichols. Inflectional morphology, volume 3, page 169–240. Cambridge University Press, 2 edition, 2007.

Evaluating on synthetic versions of a language

The idea:

- Manipulate a particular aspect of a language's morphology.
- Train on the new language.
- Evaluate on reordered treebanks for the new language.

Some properties that could be manipulated (Bickel and Nichols, 2007):

- Syncretism; are nom and acc forms in a paradigm the same?

Balthasar Bickel and Johanna Nichols. Inflectional morphology, volume 3, page 169–240. Cambridge University Press, 2 edition, 2007.

Evaluating on synthetic versions of a language

The idea:

- Manipulate a particular aspect of a language's morphology.
- Train on the new language.
- Evaluate on reordered treebanks for the new language.

Some properties that could be manipulated (Bickel and Nichols, 2007):

- Syncretism; are nom and acc forms in a paradigm the same?
- Flexivity; do morphemes exhibit high degree of allomorphy?

Balthasar Bickel and Johanna Nichols. Inflectional morphology, volume 3, page 169–240. Cambridge University Press, 2 edition, 2007.

Evaluating on synthetic versions of a language

The idea:

- Manipulate a particular aspect of a language's morphology.
- Train on the new language.
- Evaluate on reordered treebanks for the new language.

Some properties that could be manipulated (Bickel and Nichols, 2007):

- Syncretism; are nom and acc forms in a paradigm the same?
- Flexivity; do morphemes exhibit high degree of allomorphy?
- Exponence; how much grammatical information in a morph/formative?

Balthasar Bickel and Johanna Nichols. Inflectional morphology, volume 3, page 169–240. Cambridge University Press, 2 edition, 2007.

Evaluating on synthetic versions of a language

The idea:

- Manipulate a particular aspect of a language's morphology.
- Train on the new language.
- Evaluate on reordered treebanks for the new language.

Some properties that could be manipulated (Bickel and Nichols, 2007):

- Syncretism; are nom and acc forms in a paradigm the same?
- Flexivity; do morphemes exhibit high degree of allomorphy?
- Exponence; how much grammatical information in a morph/formative?
- Synthesis; how many morphs/formatives per word?

Balthasar Bickel and Johanna Nichols. Inflectional morphology, volume 3, page 169–240. Cambridge University Press, 2 edition, 2007.

Preliminary experiments (Polish base)

1. Making the case marking (very) explicit:

Zielon**ą**	łąk**ę**	→	Zielona**acc**	łąka**acc**
Green.fem.acc	meadow.acc		Green.fem.acc	meadow.acc

Preliminary experiments (Polish base)

1. Making the case marking (very) explicit:

Zielon**ą**	łąk**ę**	→	Zielona**acc**	łąka**acc**
Green.fem.acc	meadow.acc		Green.fem.acc	meadow.acc

What changes?

Preliminary experiments (Polish base)

1. Making the case marking (very) explicit:

Zielon**ą**	łąk**ę**	Zielona**acc**	łąka**acc**
Green.fem.acc	meadow.acc	Green.fem.acc	meadow.acc

What changes?

- no syncretism
- no allomorphy
- less exponence for adjectives
  - previously two categories in one morph (gender, case), now two morphs

Preliminary experiments (Polish base)

2. Making the case marking (very) explicit and unbound:

Zielon**ą**	łąk**ę**	**Acc**	zielona	łąka
Green.fem.acc	meadow.acc	→ acc	Green.fem	meadow

Preliminary experiments (Polish base)

2. Making the case marking (very) explicit and unbound:

Zielon**ą**	łąk**ę**		**Acc**	zielona	łąka
Green.fem.acc	meadow.acc	→	acc	Green.fem	meadow

What changes?

Preliminary experiments (Polish base)

2. Making the case marking (very) explicit and unbound:

Zielon**ą**	łąk**ę**		**Acc**	zielona łąka
Green.fem.acc	meadow.acc	→	acc	Green.fem meadow

What changes?

- no syncretism
- no allomorphy
- less exponence for adjectives
- less synthesis

Zielon**ą** łąk**ę**

Green.fem.acc meadow.acc

	BERT					CNN				
	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	
svo	94.1	100	98.3	86.9	*295*	88.0	93.6	91.5	77.0	*295*
ovs	91.0	94.2	89.8	77.0	*295*	69.9	35.6	43.4	54.1	*295*
sov	90.5	97.3	82.1	68.9	*295*	74.2	84.7	34.0	44.3	*295*
osv	91.0	92.9	89.8	72.1	*295*	67.7	43.7	37.9	47.5	*295*
vso	84.8	78.0	71.9	68.9	*295*	66.6	32.2	43.4	44.3	*295*
vos	89.1	82.0	94.0	85.2	*295*	69.7	18.0	79.1	72.1	*295*

Zielon**ą** łąk**ę**

Green.fem.acc meadow.acc

	BERT					CNN				
	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	
svo	94.1	100	98.3	86.9	*295*	88.0	93.6	91.5	77.0	*295*
ovs	91.0	94.2	89.8	77.0	*295*	69.9	35.6	43.4	54.1	*295*
sov	90.5	97.3	82.1	68.9	*295*	74.2	84.7	34.0	44.3	*295*
osv	91.0	92.9	89.8	72.1	*295*	67.7	43.7	37.9	47.5	*295*
vso	84.8	78.0	71.9	68.9	*295*	66.6	32.2	43.4	44.3	*295*
vos	89.1	82.0	94.0	85.2	*295*	69.7	18.0	79.1	72.1	*295*

Zielona**acc** łąka**acc**

Green.fem.acc meadow.acc

	BERT					CNN				
	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	
svo	94.2	100	97.9	91.8	*295*	89.4	98.6	95.3	85.2	*295*
ovs	93.5	98.6	97.9	86.9	*295*	87.0	95.9	93.2	77.0	*295*
sov	91.7	99.7	87.7	68.9	*295*	83.5	97.3	72.8	50.8	*295*
osv	92.7	98.6	95.7	82.0	*295*	85.8	89.5	87.2	68.9	*295*
vso	90.1	92.9	88.5	68.9	*295*	79.6	96.3	58.7	54.1	*295*
vos	91.4	90.8	97.9	82.0	*295*	84.5	79.0	92.3	86.9	*295*

Zielon**ą** łąk**ę**

Green.fem.acc meadow.acc

	BERT					CNN				
	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	
svo	94.1	100	98.3	86.9	*295*	88.0	93.6	91.5	77.0	*295*
ovs	91.0	94.2	89.8	77.0	*295*	69.9	35.6	43.4	54.1	*295*
sov	90.5	97.3	82.1	68.9	*295*	74.2	84.7	34.0	44.3	*295*
osv	91.0	92.9	89.8	72.1	*295*	67.7	43.7	37.9	47.5	*295*
vso	84.8	78.0	71.9	68.9	*295*	66.6	32.2	43.4	44.3	*295*
vos	89.1	82.0	94.0	85.2	*295*	69.7	18.0	79.1	72.1	*295*

Zielona**acc** łąka**acc**

Green.fem.acc meadow.acc

	BERT					CNN				
	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	
svo	94.2	100	97.9	91.8	*295*	89.4	98.6	95.3	85.2	*295*
ovs	93.5	98.6	97.9	86.9	*295*	87.0	95.9	93.2	77.0	*295*
sov	91.7	99.7	87.7	68.9	*295*	83.5	97.3	72.8	50.8	*295*
osv	92.7	98.6	95.7	82.0	*295*	85.8	89.5	87.2	68.9	*295*
vso	90.1	92.9	88.5	68.9	*295*	79.6	96.3	58.7	54.1	*295*
vos	91.4	90.8	97.9	82.0	*295*	84.5	79.0	92.3	86.9	*295*

Zielon**ą** łąk**ę**

Green.fem.acc meadow.acc

	BERT					CNN				
	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	
svo	94.1	100	98.3	86.9	*295*	88.0	93.6	91.5	77.0	*295*
ovs	91.0	94.2	89.8	77.0	*295*	69.9	35.6	43.4	54.1	*295*
sov	90.5	97.3	82.1	68.9	*295*	74.2	84.7	34.0	44.3	*295*
osv	91.0	92.9	89.8	72.1	*295*	67.7	43.7	37.9	47.5	*295*
vso	84.8	78.0	71.9	68.9	*295*	66.6	32.2	43.4	44.3	*295*
vos	89.1	82.0	94.0	85.2	*295*	69.7	18.0	79.1	72.1	*295*

Zielona**acc** łąka**acc**

Green.fem.acc meadow.acc

	BERT					CNN				
	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	
svo	94.2	100	97.9	91.8	*295*	89.4	98.6	95.3	85.2	*295*
ovs	93.5	98.6	97.9	86.9	*295*	87.0	95.9	93.2	77.0	*295*
sov	91.7	99.7	87.7	68.9	*295*	83.5	97.3	72.8	50.8	*295*
osv	92.7	98.6	95.7	82.0	*295*	85.8	89.5	87.2	68.9	*295*
vso	90.1	92.9	88.5	68.9	*295*	79.6	96.3	58.7	54.1	*295*
vos	91.4	90.8	97.9	82.0	*295*	84.5	79.0	92.3	86.9	*295*

Zielon**ą** łąk**ę**

Green.fem.acc meadow.acc

Zielona**acc** łąka**acc**

Green.fem.acc meadow.acc

Acc zielona łąka

acc Green.fem meadow

	BERT					CNN				
	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	
svo	94.1	100	98.3	86.9	295	88.0	93.6	91.5	77.0	295
ovs	91.0	94.2	89.8	77.0	295	69.9	35.6	43.4	54.1	295
sov	90.5	97.3	82.1	68.9	295	74.2	84.7	34.0	44.3	295
osv	91.0	92.9	89.8	72.1	295	67.7	43.7	37.9	47.5	295
vso	84.8	78.0	71.9	68.9	295	66.6	32.2	43.4	44.3	295
vos	89.1	82.0	94.0	85.2	295	69.7	18.0	79.1	72.1	295

	BERT					CNN				
	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	
svo	94.2	100	97.9	91.8	295	89.4	98.6	95.3	85.2	295
ovs	93.5	98.6	97.9	86.9	295	87.0	95.9	93.2	77.0	295
sov	91.7	99.7	87.7	68.9	295	83.5	97.3	72.8	50.8	295
osv	92.7	98.6	95.7	82.0	295	85.8	89.5	87.2	68.9	295
vso	90.1	92.9	88.5	68.9	295	79.6	96.3	58.7	54.1	295
vos	91.4	90.8	97.9	82.0	295	84.5	79.0	92.3	86.9	295

	BERT					CNN				
	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	
svo	96.0	99.7	98.3	93.4	295	93.2	99.7	96.6	90.2	295
ovs	95.9	100	98.3	90.2	295	92.4	99.0	97.0	80.3	295
sov	94.9	99.7	93.2	78.7	295	90.4	97.3	84.3	65.6	295
osv	95.1	99.0	97.0	80.3	295	92.1	96.6	97.0	72.1	295
vso	93.8	98.6	89.8	67.2	295	90.0	98.6	83.8	67.2	295
vos	94.8	94.6	97.4	91.8	295	91.2	90.8	95.3	85.2	295

Zielon**ą** łąk**ę**

Green.fem.acc meadow.acc

Zielona**acc** łąka**acc**

Green.fem.acc meadow.acc

Acc zielona łąka

acc Green.fem meadow

	BERT					CNN				
	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	
svo	94.1	100	98.3	86.9	295	88.0	93.6	91.5	77.0	295
ovs	91.0	94.2	89.8	77.0	295	69.9	35.6	43.4	54.1	295
sov	90.5	97.3	82.1	68.9	295	74.2	84.7	34.0	44.3	295
osv	91.0	92.9	89.8	72.1	295	67.7	43.7	37.9	47.5	295
vso	84.8	78.0	71.9	68.9	295	66.6	32.2	43.4	44.3	295
vos	89.1	82.0	94.0	85.2	295	69.7	18.0	79.1	72.1	295

	BERT					CNN				
	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	
svo	94.2	100	97.9	91.8	295	89.4	98.6	95.3	85.2	295
ovs	93.5	98.6	97.9	86.9	295	87.0	95.9	93.2	77.0	295
sov	91.7	99.7	87.7	68.9	295	83.5	97.3	72.8	50.8	295
osv	92.7	98.6	95.7	82.0	295	85.8	89.5	87.2	68.9	295
vso	90.1	92.9	88.5	68.9	295	79.6	96.3	58.7	54.1	295
vos	91.4	90.8	97.9	82.0	295	84.5	79.0	92.3	86.9	295

	BERT					CNN				
	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	
svo	96.0	99.7	98.3	93.4	295	93.2	99.7	96.6	90.2	295
ovs	95.9	100	98.3	90.2	295	92.4	99.0	97.0	80.3	295
sov	94.9	99.7	93.2	78.7	295	90.4	97.3	84.3	65.6	295
osv	95.1	99.0	97.0	80.3	295	92.1	96.6	97.0	72.1	295
vso	93.8	98.6	89.8	67.2	295	90.0	98.6	83.8	67.2	295
vos	94.8	94.6	97.4	91.8	295	91.2	90.8	95.3	85.2	295

Zielon**ą** łąk**ę**

Green.fem.acc meadow.acc

Zielona**acc** łąka**acc**

Green.fem.acc meadow.acc

Acc zielona łąka

acc Green.fem meadow

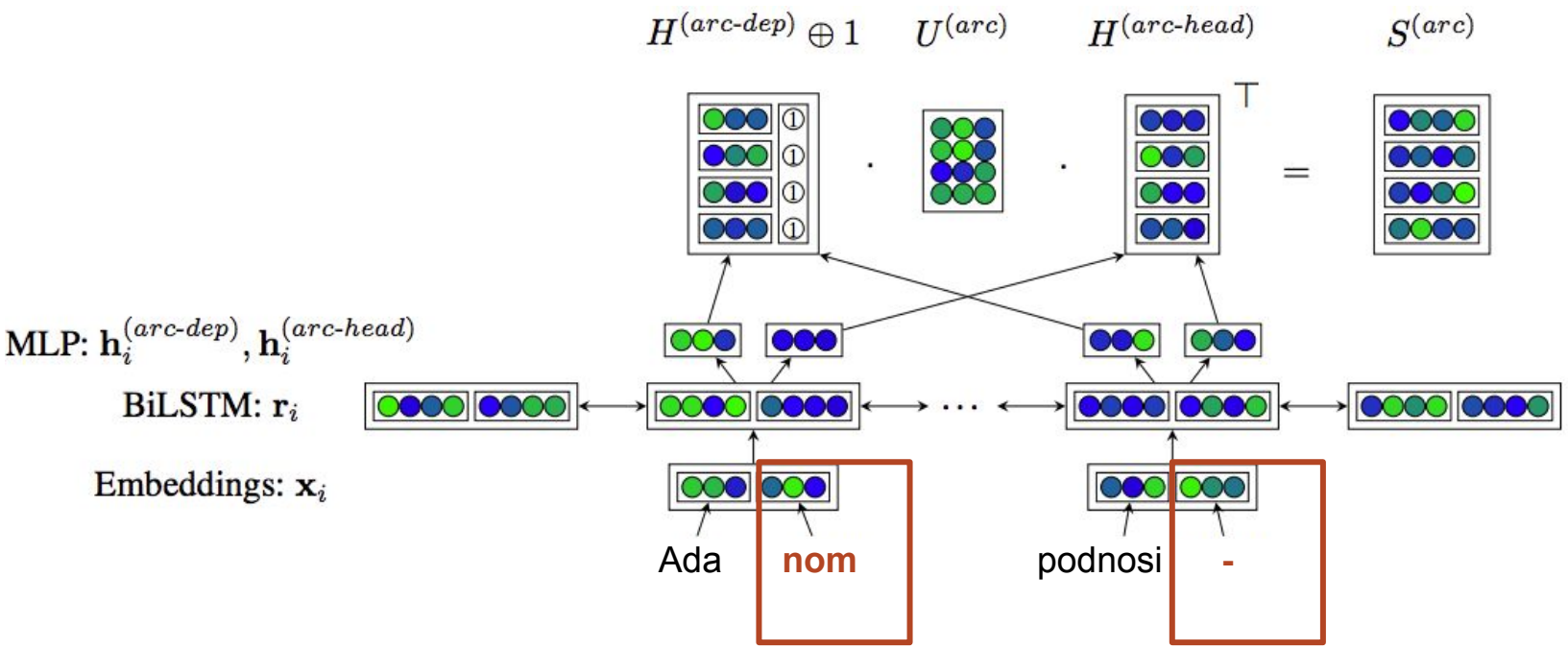
	BERT					CNN				
	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	
svo	94.1	100	98.3	86.9	295	88.0	93.6	91.5	77.0	295
ovs	91.0	94.2	89.8	77.0	295	69.9	35.6	43.4	54.1	295
sov	90.5	97.3	82.1	68.9	295	74.2	84.7	34.0	44.3	295
osv	91.0	92.9	89.8	72.1	295	67.7	43.7	37.9	47.5	295
vso	84.8	78.0	71.9	68.9	295	66.6	32.2	43.4	44.3	295
vos	89.1	82.0	94.0	85.2	295	69.7	18.0	79.1	72.1	295

	BERT					CNN				
	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	
svo	94.2	100	97.9	91.8	295	89.4	98.6	95.3	85.2	295
ovs	93.5	98.6	97.9	86.9	295	87.0	95.9	93.2	77.0	295
sov	91.7	99.7	87.7	68.9	295	83.5	97.3	72.8	50.8	295
osv	92.7	98.6	95.7	82.0	295	85.8	89.5	87.2	68.9	295
vso	90.1	92.9	88.5	68.9	295	79.6	96.3	58.7	54.1	295
vos	91.4	90.8	97.9	82.0	295	84.5	79.0	92.3	86.9	295

	BERT					CNN				
	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	
svo	96.0	99.7	98.3	93.4	295	93.2	99.7	96.6	90.2	295
ovs	95.9	100	98.3	90.2	295	92.4	99.0	97.0	80.3	295
sov	94.9	99.7	93.2	78.7	295	90.4	97.3	84.3	65.6	295
osv	95.1	99.0	97.0	80.3	295	92.1	96.6	97.0	72.1	295
vso	93.8	98.6	89.8	67.2	295	90.0	98.6	83.8	67.2	295
vos	94.8	94.6	97.4	91.8	295	91.2	90.8	95.3	85.2	295

How does this compare to feeding case as a feature?

- Leave the language unchanged.
- Feed an embedding of a gold case as an input.



Zielon**ą** łąk**ę**

Green.fem.acc meadow.acc

Zielona**acc** łąka**acc**

Green.fem.acc meadow.acc

Acc zielona łąka

acc Green.fem meadow

	BERT					CNN				
	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	
svo	94.1	100	98.3	86.9	295	88.0	93.6	91.5	77.0	295
ovs	91.0	94.2	89.8	77.0	295	69.9	35.6	43.4	54.1	295
sov	90.5	97.3	82.1	68.9	295	74.2	84.7	34.0	44.3	295
osv	91.0	92.9	89.8	72.1	295	67.7	43.7	37.9	47.5	295
vso	84.8	78.0	71.9	68.9	295	66.6	32.2	43.4	44.3	295
vos	89.1	82.0	94.0	85.2	295	69.7	18.0	79.1	72.1	295

	BERT					CNN				
	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	
svo	94.2	100	97.9	91.8	295	89.4	98.6	95.3	85.2	295
ovs	93.5	98.6	97.9	86.9	295	87.0	95.9	93.2	77.0	295
sov	91.7	99.7	87.7	68.9	295	83.5	97.3	72.8	50.8	295
osv	92.7	98.6	95.7	82.0	295	85.8	89.5	87.2	68.9	295
vso	90.1	92.9	88.5	68.9	295	79.6	96.3	58.7	54.1	295
vos	91.4	90.8	97.9	82.0	295	84.5	79.0	92.3	86.9	295

	BERT					CNN				
	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	
svo	96.0	99.7	98.3	93.4	295	93.2	99.7	96.6	90.2	295
ovs	95.9	100	98.3	90.2	295	92.4	99.0	97.0	80.3	295
sov	94.9	99.7	93.2	78.7	295	90.4	97.3	84.3	65.6	295
osv	95.1	99.0	97.0	80.3	295	92.1	96.6	97.0	72.1	295
vso	93.8	98.6	89.8	67.2	295	90.0	98.6	83.8	67.2	295
vos	94.8	94.6	97.4	91.8	295	91.2	90.8	95.3	85.2	295

Zielon**ą** łąk**ę**

Green.fem.acc meadow.acc

Zielona**acc** łąka**acc**

Green.fem.acc meadow.acc

Acc zielona łąka

acc Green.fem meadow

Gold case as input

	BERT					CNN				
	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	
svo	94.1	100	98.3	86.9	295	88.0	93.6	91.5	77.0	295
ovs	91.0	94.2	89.8	77.0	295	69.9	35.6	43.4	54.1	295
sov	90.5	97.3	82.1	68.9	295	74.2	84.7	34.0	44.3	295
osv	91.0	92.9	89.8	72.1	295	67.7	43.7	37.9	47.5	295
vso	84.8	78.0	71.9	68.9	295	66.6	32.2	43.4	44.3	295
vos	89.1	82.0	94.0	85.2	295	69.7	18.0	79.1	72.1	295

	BERT					CNN				
	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	
svo	94.2	100	97.9	91.8	295	89.4	98.6	95.3	85.2	295
ovs	93.5	98.6	97.9	86.9	295	87.0	95.9	93.2	77.0	295
sov	91.7	99.7	87.7	68.9	295	83.5	97.3	72.8	50.8	295
osv	92.7	98.6	95.7	82.0	295	85.8	89.5	87.2	68.9	295
vso	90.1	92.9	88.5	68.9	295	79.6	96.3	58.7	54.1	295
vos	91.4	90.8	97.9	82.0	295	84.5	79.0	92.3	86.9	295

	BERT					CNN				
	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	
svo	96.0	99.7	98.3	93.4	295	93.2	99.7	96.6	90.2	295
ovs	95.9	100	98.3	90.2	295	92.4	99.0	97.0	80.3	295
sov	94.9	99.7	93.2	78.7	295	90.4	97.3	84.3	65.6	295
osv	95.1	99.0	97.0	80.3	295	92.1	96.6	97.0	72.1	295
vso	93.8	98.6	89.8	67.2	295	90.0	98.6	83.8	67.2	295
vos	94.8	94.6	97.4	91.8	295	91.2	90.8	95.3	85.2	295

	BERT					CNN				
	all	nsubj	obj	iobj	#	all	nsubj	obj	iobj	
svo	95.2	100	99.1	93.4	295	91.6	99.0	97.4	86.9	295
ovs	94.2	98.6	97.0	86.9	295	90.3	98.6	92.8	78.7	295
sov	93.0	99.3	86.0	78.7	295	87.7	95.9	82.1	62.3	295
osv	93.8	96.9	94.5	85.2	295	88.8	95.3	90.2	80.3	295
vso	90.6	96.3	77.4	75.4	295	84.5	93.2	65.5	65.6	295
vos	92.3	90.5	97.0	90.2	295	86.6	80.3	92.8	83.6	295

Insights

- NNs can rely on morphological features over the word order.
- The current architectures are just not suited to handle some types of morphology.
- Word order provides more straightforward signal.

Insights

- NNs can rely on morphological features over the word order.
- The current architectures are just not suited to handle some types of morphology.
- Word order provides more straightforward signal.

What is hard?

- bound morphemes
- high degrees of syncretism?
- allomorphy?
- morpheme ambiguity?

Insights

- NNs can rely on morphological features over the word order.
- The current architectures are just not suited to handle some types of morphology.
- Word order provides more straightforward signal.

What is hard?

- bound morphemes
- high degrees of syncretism?
- allomorphy?
- morpheme ambiguity?

Thank you!

Evaluating on synthetic versions of Polish

- **Primary focus**: case marking
- **Language base**: Polish

Nominative-accusative syncretism

	n-stem declension		t-stem declension	
	Singular	**Plural**	**Singular**	**Plural**
Nominative	imię	imiona	cielę	cielęta
Accusative	imię	imiona	cielę	cielęta
Genitive	imienia	imion	cielęcia	cieląt
Locative	imieniu	imionach	cielęciu	cielętach
Dative	imieniu	imionom	cielęciu	cielętom
Instrumental	imieniem	imionami	cielęciem	cielętami
Vocative	imię	imiona	cielę	cielęta

Neuter

		singular	plural
nominative		-	*
genitive	**personal**	-a	*
	animate		
	inanimate	-a / -u	
dative		*	-om
accusative	**personal**	*same as the genitive*	*same as the genitive*
	animate		*same as the nominative*
	inanimate	*same as the nominative*	
instrumental		-em	-ami[1]
locative		*	-ach
vocative		*same as the locative*[2]	*same as the nominative*

Masculine

	Hard declension		Soft declension	
	Singular	**Plural**	**Singular**	**Plural**
Nominative	mapa	mapy	granica	granice
Accusative	mapę	mapy	granicę	granice
Genitive	mapy	map	granicy	granic
Locative	mapie	mapach	granicy	granicach
Dative	mapie	mapom	granicy	granicom
Instrumental	mapą	mapami	granicą	granicami
Vocative	mapo	mapy	granico	granice

Feminine

Evaluating on synthetic versions of Polish

- **Primary focus**: case marking
- **Language base**: Polish

Allomorphy

		n-stem declension		t-stem declension	
		Singular	Plural	Singular	Plural
Nominative		imię	imiona	cielę	cielęta
Accusative		imię	imiona	cielę	cielęta
Genitive		imienia	imion	cielęcia	cieląt
Locative		imieniu	imionach	cielęciu	cielętach
Dative		imieniu	imionom	cielęciu	cielętom
Instrumental		imieniem	imionami	cielęciem	cielętami
Vocative		imię	imiona	cielę	cielęta

Neuter

		singular	plural
nominative		-	*
genitive	personal	-a	*
	animate		
	inanimate	-a / -u	
dative		*	-om
accusative	personal	*same as the genitive*	*same as the genitive*
	animate		*same as the nominative*
	inanimate	*same as the nominative*	
instrumental		-em	-ami[1]
locative		*	-ach
vocative		*same as the locative*[2]	*same as the nominative*

Masculine

	Hard declension		Soft declension	
	Singular	Plural	Singular	Plural
Nominative	mapa	mapy	granica	granice
Accusative	mapę	mapy	granicę	granice
Genitive	mapy	map	granicy	granic
Locative	mapie	mapach	granicy	granicach
Dative	mapie	mapom	granicy	granicom
Instrumental	mapą	mapami	granicą	granicami
Vocative	mapo	mapy	granico	granice

Feminine

Evaluating on synthetic versions of Polish

- **Primary focus**: case marking
- **Language base**: Polish

Morph ambiguity (-a)

		singular	plural
nominative		-	*
genitive	personal	-a	*
	animate		
	inanimate	-a / -u	
dative		*	-om
accusative	personal	*same as the genitive*	*same as the genitive*
	animate		*same as the nominative*
	inanimate	*same as the nominative*	
instrumental		-em	-ami[1]
locative		*	-ach
vocative		*same as the locative*[2]	*same as the nominative*

Masculine

	n-stem declension		t-stem declension	
	Singular	Plural	Singular	Plural
Nominative	imię	imiona	cielę	cielęta
Accusative	imię	imiona	cielę	cielęta
Genitive	imienia	imion	cielęcia	cieląt
Locative	imieniu	imionach	cielęciu	cielętach
Dative	imieniu	imionom	cielęciu	cielętom
Instrumental	imieniem	imionami	cielęciem	cielętami
Vocative	imię	imiona	cielę	cielęta

Neuter

	Hard declension		Soft declension	
	Singular	Plural	Singular	Plural
Nominative	mapa	mapy	granica	granice
Accusative	mapę	mapy	granicę	granice
Genitive	mapy	map	granicy	granic
Locative	mapie	mapach	granicy	granicach
Dative	mapie	mapom	granicy	granicom
Instrumental	mapą	mapami	granicą	granicami
Vocative	mapo	mapy	granico	granice

Feminine

Evaluating on synthetic versions of Polish

- **Primary focus**: case marking
- **Language base**: Polish

\+ making the morphs more easily segmentable or unbound

		singular	plural
nominative		-	*
genitive	personal	-a	*
	animate		
	inanimate	-a / -u	
dative		*	-om
accusative	personal	*same as the genitive*	*same as the genitive*
	animate		*same as the nominative*
	inanimate	*same as the nominative*	
instrumental		-em	-ami[1]
locative		*	-ach
vocative		*same as the locative*[2]	*same as the nominative*

Masculine

	n-stem declension		t-stem declension	
	Singular	Plural	Singular	Plural
Nominative	imi**ę**	imion**a**	ciel**ę**	cielęt**a**
Accusative	imi**ę**	imion**a**	ciel**ę**	cielęt**a**
Genitive	imieni**a**	imion	cielęci**a**	cieląt
Locative	imieni**u**	imion**ach**	cielęci**u**	cielęt**ach**
Dative	imieni**u**	imion**om**	cielęci**u**	cielęt**om**
Instrumental	imieni**em**	imion**ami**	cielęci**em**	cielęt**ami**
Vocative	imi**ę**	imion**a**	ciel**ę**	cielęt**a**

Neuter

	Hard declension		Soft declension	
	Singular	Plural	Singular	Plural
Nominative	mapa	map**y**	granic**a**	granic**e**
Accusative	map**ę**	map**y**	granic**ę**	granic**e**
Genitive	map**y**	map	granic**y**	granic
Locative	map**ie**	map**ach**	granic**y**	granic**ach**
Dative	map**ie**	map**om**	granic**y**	granic**om**
Instrumental	map**ą**	map**ami**	granic**ą**	granic**ami**
Vocative	map**o**	map**y**	granic**o**	granic**e**

Feminine