A Neural Model of Adaptation in Reading

Marten van Schijndel and Tal Linzen

November 4, 2018

Department of Cognitive Science, Johns Hopkins University

Thought experiment

Thought experiment

…cassowary …

Thought experiment

…cassoway …

…cassoway? …

Thought experiment

...cassowary ...

...cassowary? ...

...cassowary ...

Thought experiment

...cassowary ...

...cassowary? ...

...cassowary ...

...cassowary! ...

Thought experiment

…cassowary …

…cassowary? …

…cassowary …

…cassowary! …

You are now less surprised when this person says ‘Cassowary’

A psycholinguistic example of adaptation (Fine & Jaeger, 2016)

The

A psycholinguistic example of adaptation (Fine & Jaeger, 2016)

The soldiers

A psycholinguistic example of adaptation (Fine & Jaeger, 2016)

The soldiers warned

A psycholinguistic example of adaptation (Fine & Jaeger, 2016)

The soldiers warned about

A psycholinguistic example of adaptation (Fine & Jaeger, 2016)

The soldiers warned about the

A psycholinguistic example of adaptation (Fine & Jaeger, 2016)

The soldiers warned about the dangers

A psycholinguistic example of adaptation (Fine & Jaeger, 2016)

The soldiers warned about the dangers conducted

A psycholinguistic example of adaptation (Fine & Jaeger, 2016)

The soldiers warned about the dangers conducted the

A psycholinguistic example of adaptation (Fine & Jaeger, 2016)

The soldiers warned about the dangers conducted the raid.

A psycholinguistic example of adaptation (Fine & Jaeger, 2016)

The soldiers warned about the dangers conducted the raid.

Unreduced: The soldiers (who were) warned about the dangers …

A psycholinguistic example of adaptation (Fine & Jaeger, 2016)

The soldiers warned about the dangers conducted the raid.

Unreduced: The soldiers (who were) warned about the dangers …

By end of experiment, subjects expected RRC more than at beginning

A psycholinguistic example of adaptation (Fine & Jaeger, 2016)

The soldiers warned about the dangers conducted the raid.

Unreduced: The soldiers (who were) warned about the dangers ...

By end of experiment, subjects expected RRC more than at beginning

- Humans adapt to syntactic structures

Adaptation is studied in NLP

- Domain adaptation (Kuhn & de Mori, 1990; McClosky, 2010)
  News Model $\rightarrow$ Biomedical Text
- Handling unknown words (Grave et al., 2015)
  Learn new words from context
- Style adaptation (Jaech & Ostendorf, 2017)
  Lawyer A $\rightarrow$ Lawyer B

Adaptation is studied in NLP

- Domain adaptation (Kuhn & de Mori, 1990; McClosky, 2010)
  News Model $\rightarrow$ Biomedical Text
- Handling unknown words (Grave et al., 2015)
  Learn new words from context
- Style adaptation (Jaech & Ostendorf, 2017)
  Lawyer A $\rightarrow$ Lawyer B

But can we model human adaptation?

Our proposed model

LSTM language model (Gives prob of next word in sequence)

Base Model: Trained on Wikipedia (90M words)
(Gulordava et al., 2018)

Our proposed model

LSTM language model (Gives prob of next word in sequence)

Base Model: Trained on Wikipedia (90M words)
(Gulordava et al., 2018)

Adaptation algorithm:

LSTM language model (Gives prob of next word in sequence)

Base Model: Trained on Wikipedia (90M words)
(Gulordava et al., 2018)

Adaptation algorithm:

1. Test on a sentence

Our proposed model

LSTM language model (Gives prob of next word in sequence)

Base Model: Trained on Wikipedia (90M words)
(Gulordava et al., 2018)

Adaptation algorithm:

1. Test on a sentence
2. Update weights based on that sentence

Our proposed model

LSTM language model (Gives prob of next word in sequence)

Base Model: Trained on Wikipedia (90M words)
(Gulordava et al., 2018)

Adaptation algorithm:

1. Test on a sentence
2. Update weights based on that sentence
3. Repeat on remaining sentences

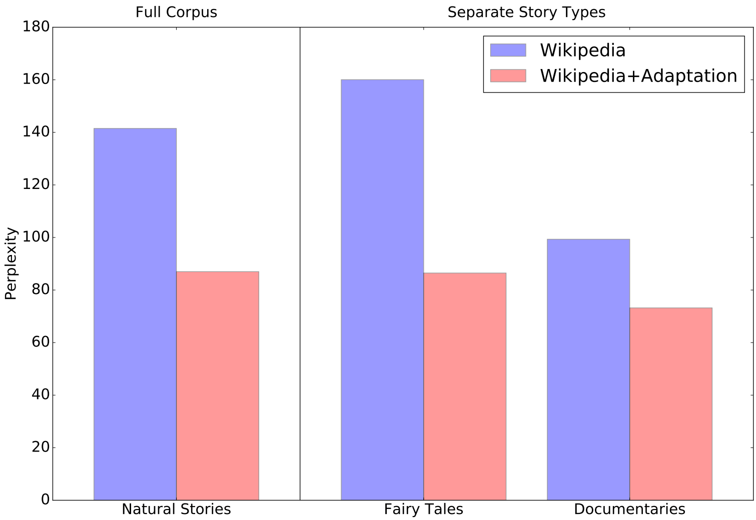
Experiment 1 (standard):
Does adaptation improve model accuracy?

Accuracy Evaluation Data

Test data: Natural Stories Corpus (Futrell et al., 2017)

- 10 texts (485 sentences)
  - 7 Fairy Tales
  - 3 Documentaries

Accuracy Results



Experiment 2:
Evaluate model against human adaptation

Evaluation Measure: Surprisal

Reading times can be predicted with surprisal (Smith and Levy, 2013)

$$\text{Surprisal}(w_i) = -\log \text{P}(w_i \mid w_{1..i-1})$$

- Timeline of adaptation is similar to human adaptation
- Adaptive surprisal predicts reading times better than non-adaptive surprisal

Evaluation: Reading Times

The soldiers (who were) warned about the dangers conducted the raid.

The soldiers (who were) warned about the dangers conducted the raid.

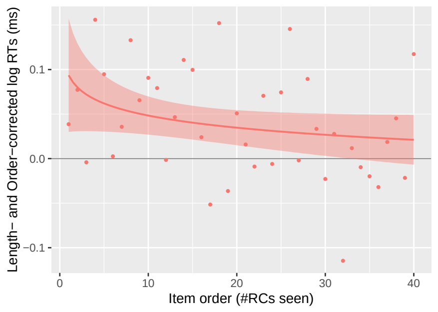


Figure 1: Human reading times

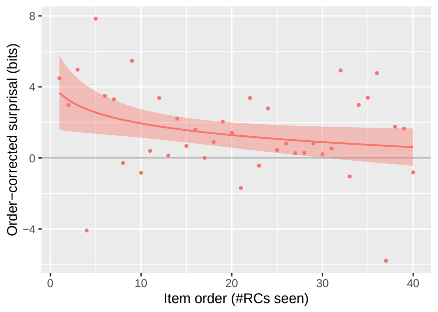
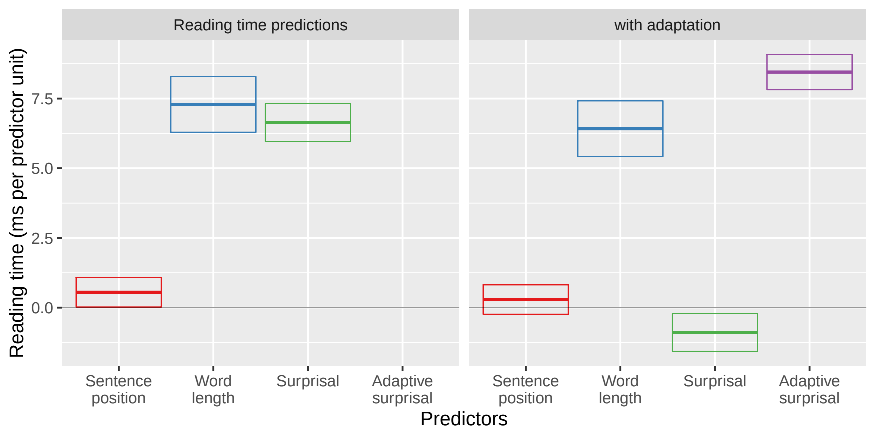


Figure 2: Adaptive model surprisal

Evaluation: Reading Times



Experiment 3:
How sensitive is adaptation to different signals?
Vocabulary? Syntax?

Generated 200 dative sentence pairs

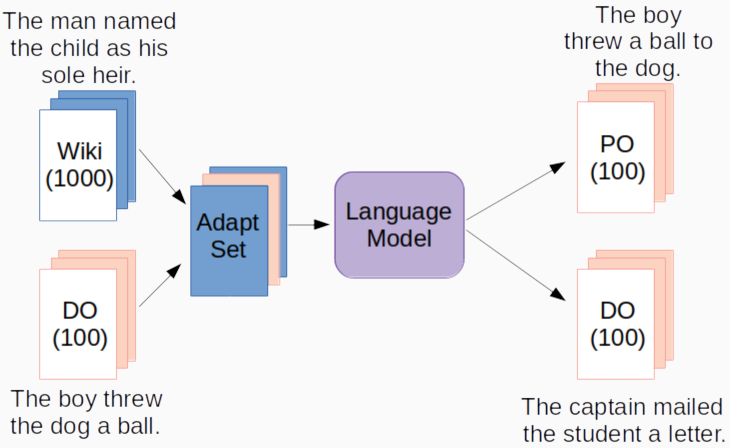
Prepositional Object (PO):

The boy threw the ball to the dog.

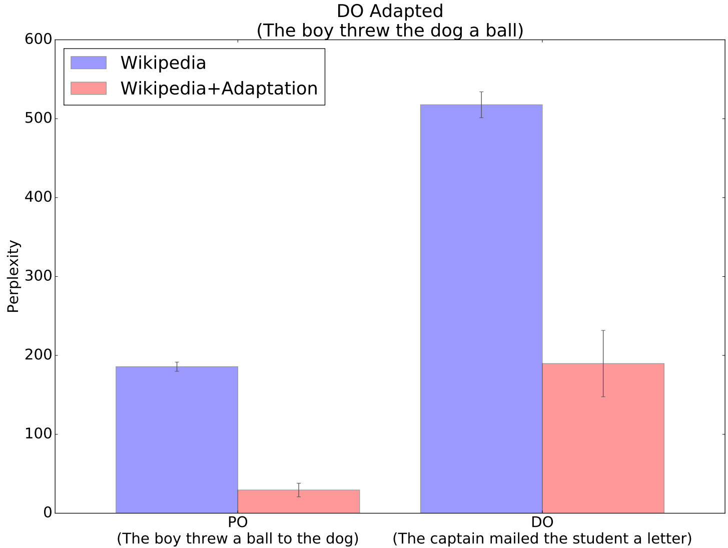
Double Object (DO):

The boy threw the dog the ball.

Dative evaluation paradigm



Model adapts to vocabulary *and* syntax



Conclusion

- Proposed a simple adaptation mechanism which
  - Is more accurate than a non-adaptive model
  - Makes more human-like predictions than a non-adaptive model

Conclusion

- Proposed a simple adaptation mechanism which
  - Is more accurate than a non-adaptive model
  - Makes more human-like predictions than a non-adaptive model
  - Is not dependent on learning rate (see paper)
  - Does not seem to suffer from catastrophic forgetting (see paper)

Conclusion

- Proposed a simple adaptation mechanism which
  - Is more accurate than a non-adaptive model
  - Makes more human-like predictions than a non-adaptive model
  - Is not dependent on learning rate (see paper)
  - Does not seem to suffer from catastrophic forgetting (see paper)
- Proposed new ways of evaluating adaptation:

Conclusion

- Proposed a simple adaptation mechanism which
  - Is more accurate than a non-adaptive model
  - Makes more human-like predictions than a non-adaptive model
  - Is not dependent on learning rate (see paper)
  - Does not seem to suffer from catastrophic forgetting (see paper)
- Proposed new ways of evaluating adaptation:
  - Human adaptive behavior

Conclusion

- Proposed a simple adaptation mechanism which
  - Is more accurate than a non-adaptive model
  - Makes more human-like predictions than a non-adaptive model
  - Is not dependent on learning rate (see paper)
  - Does not seem to suffer from catastrophic forgetting (see paper)
- Proposed new ways of evaluating adaptation:
  - Human adaptive behavior
  - Psycholinguistic experiments to probe signal sensitivity:
    Adaptation is sensitive to both vocabulary and syntax

Thanks!

MultiNLI has 10 domains (Williams, et al. 2018)

Split each domain into training and testing sets (1000 sentences each)

1. Adapt to a training domain

Catastrophic Forgetting Test

MultiNLI has 10 domains (Williams, et al. 2018)

Split each domain into training and testing sets (1000 sentences each)

1. Adapt to a training domain
2. Adapt to a second training domain

Catastrophic Forgetting Test

MultiNLI has 10 domains (Williams, et al. 2018)

Split each domain into training and testing sets (1000 sentences each)

1. Adapt to a training domain
2. Adapt to a second training domain

Does the model forget the first adaptive training domain?

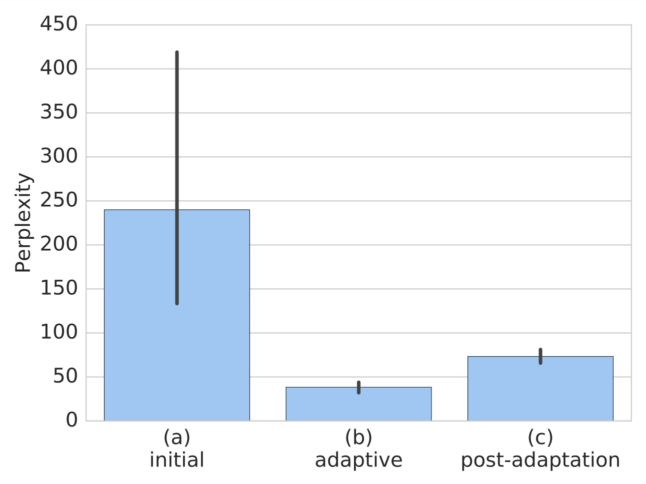


Figure 3: Perplexity on the held-out set of G_1 (a) before adaptation, (b) after adaptation to G_1, (c) after adapting to G_1 then adapting to G_2.

Model adapts to vocabulary *and* syntax

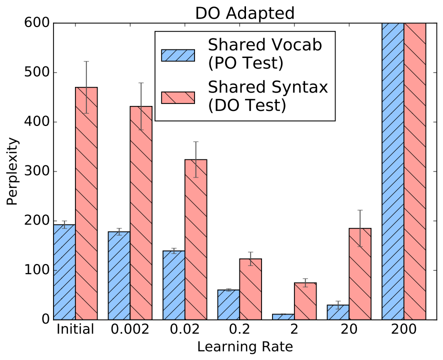


Figure 4: Learning rate influence over syntactic and lexical adaptation.

	$\hat{\beta}$	$\hat{\sigma}$	t
Without adaptive surprisal:			
Sentence position	0.55	0.53	1.03
Word length	**7.29**	**1.00**	**7.26**
Non-adaptive Surprisal	**6.64**	**0.68**	**9.79**
With adaptive surprisal:			
Sentence position	0.29	0.53	0.55
Word length	**6.42**	**1.00**	**6.40**
Non-adaptive Surprisal	-0.89	0.68	-1.31
Adaptive Surprisal	**8.45**	**0.63**	**13.42**

Fixed effects of linear mixed regression

Qualitative Adaptation Timeline

The soldiers warned about the dangers conducted the raid.

The soldiers warned about the dangers conducted the raid.

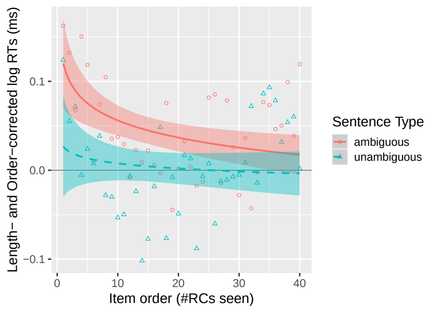


FIGURE 5: Human reading times

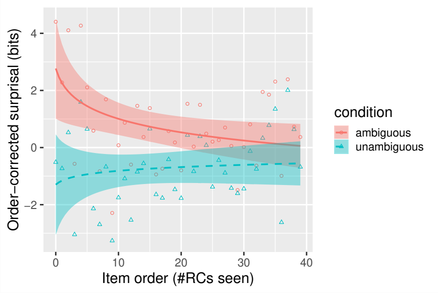


FIGURE 6: Adaptive model surprisal