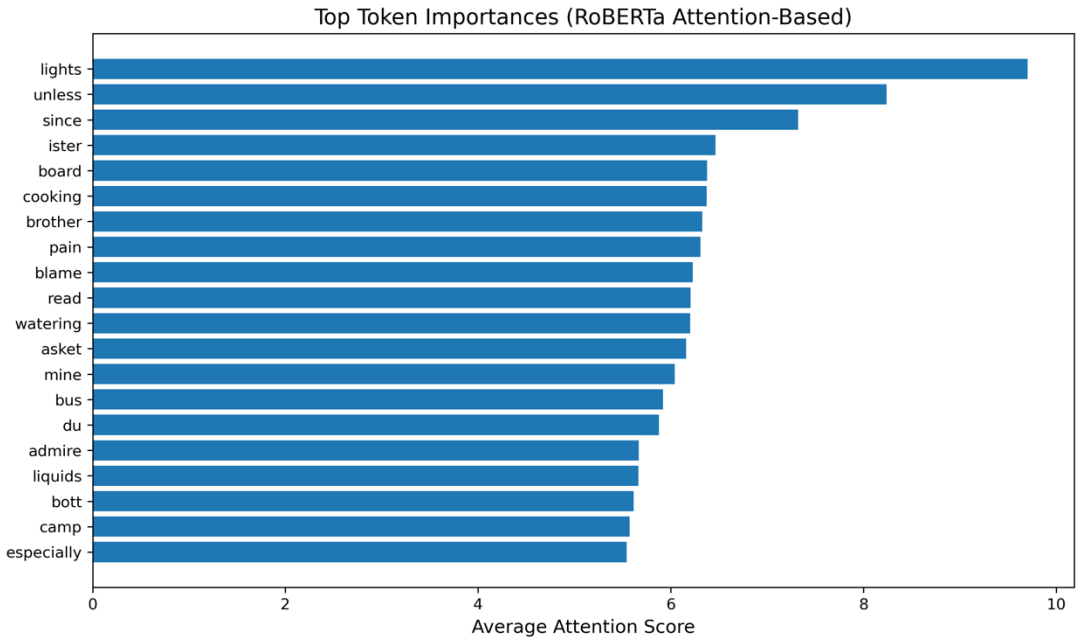
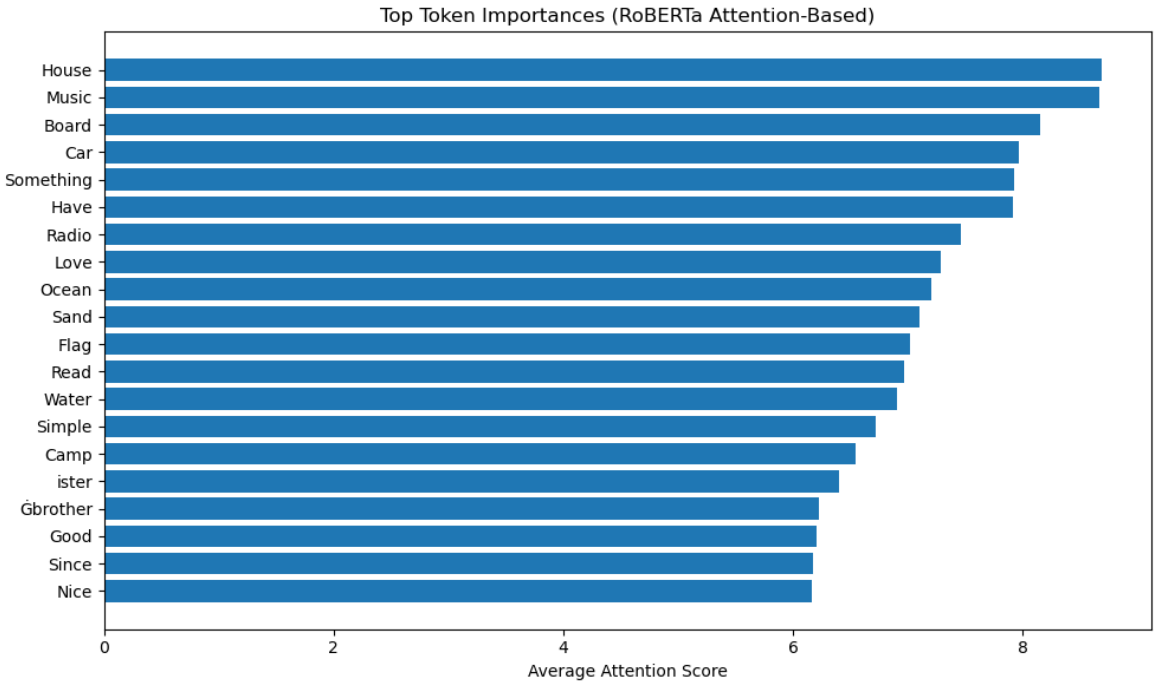
Feature importance analysis on best performing models (RoBERTa, BERT, MLP+MentalBERT embeddings

- Transformer based models: **RoBERTa**

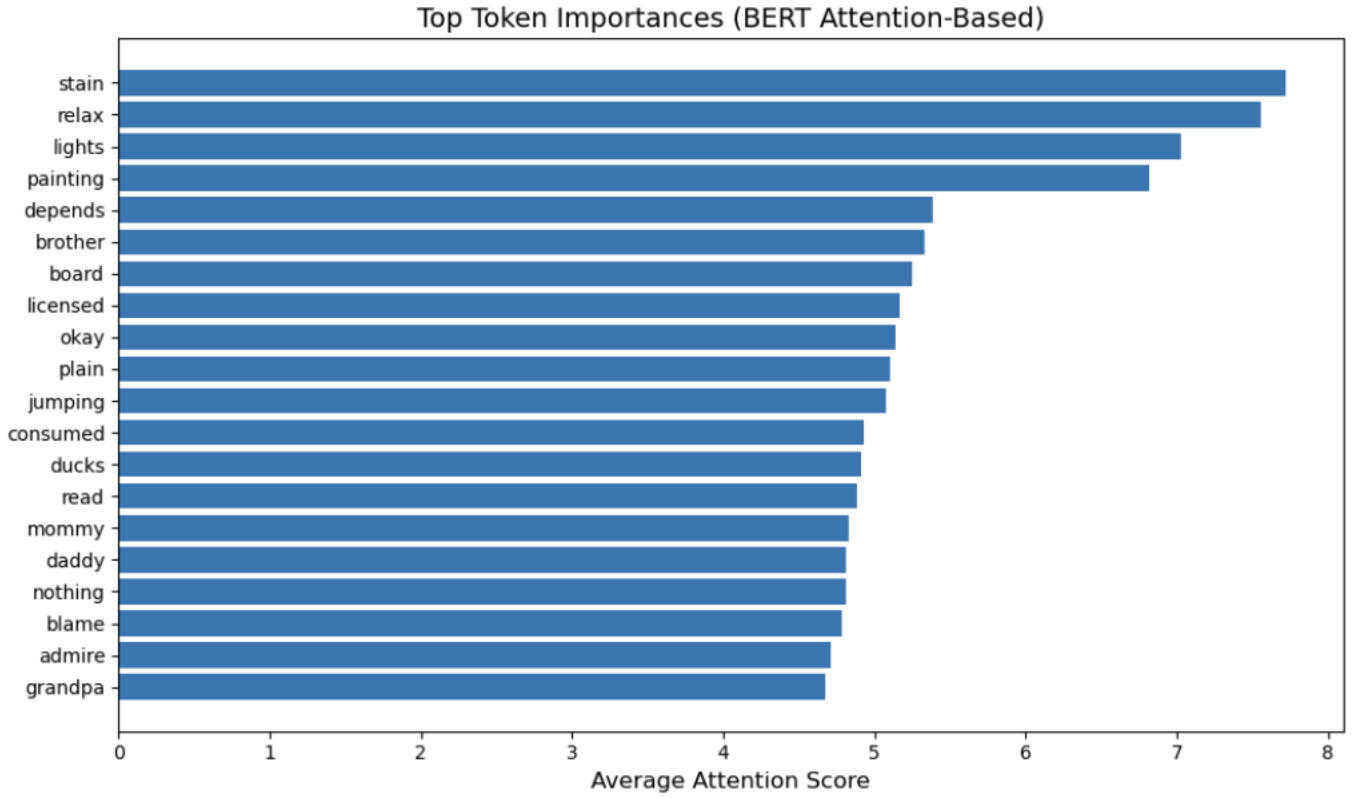


Top features when applying the model on lower-cased data



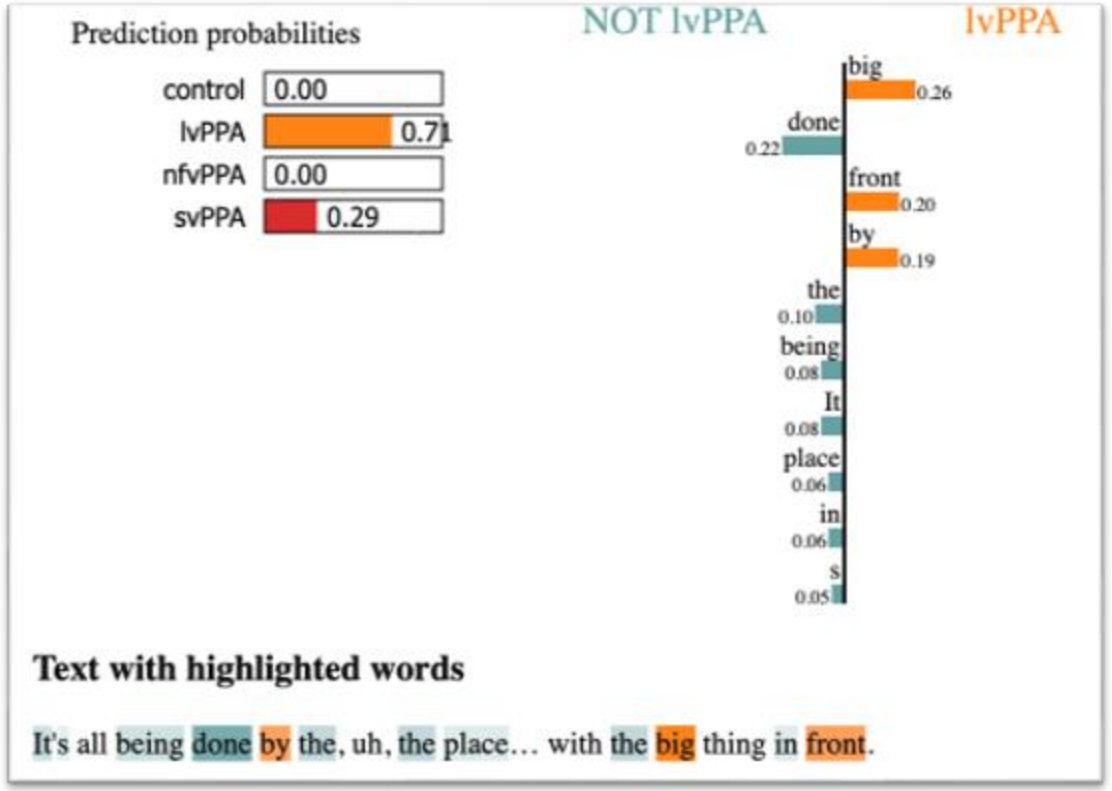
Top features when applying the model on unprocessed data (no lowercasing)

- Transformer based models: **BERT**

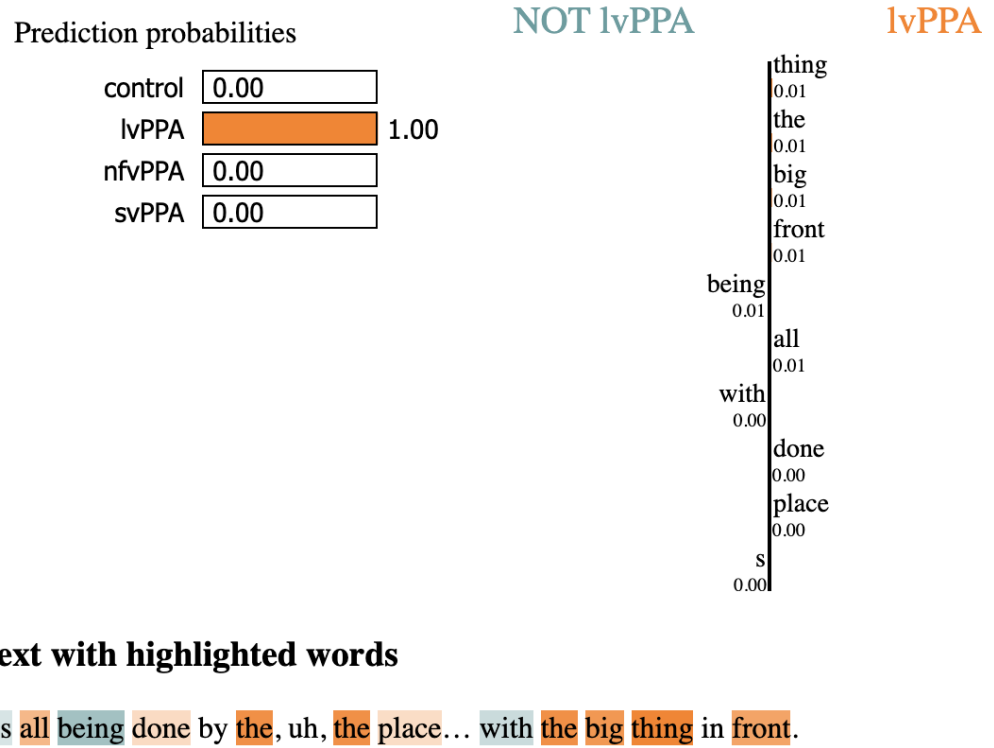


Using LIME (Local Interpretable Model-agnostic Explanations)

- Explains **why** a model made a prediction for **one specific case** Like a “local diagnostic”
- shows which features mattered most
- Works with any model (black box or not)



Synthetic sentence: manually generated
RoBERTa’s explanation



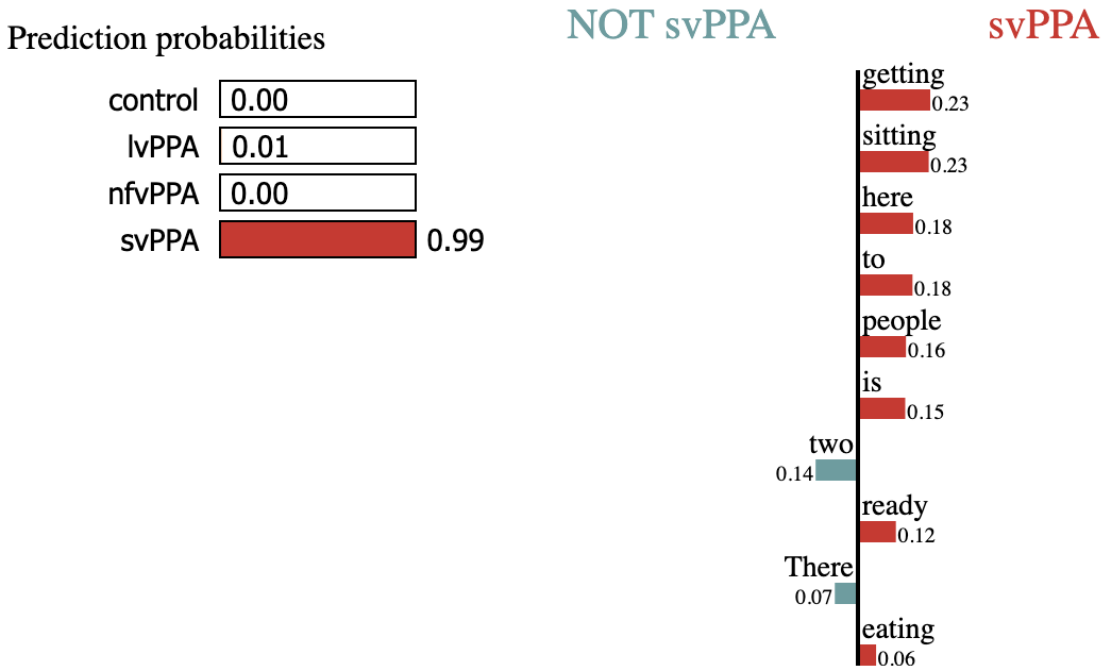
Synthetic sentence: manually generated
BERT’s explanation

- LIME on a sample from the dataset

Text with highlighted words

There is two people sitting here getting ready to do some eating, reading.

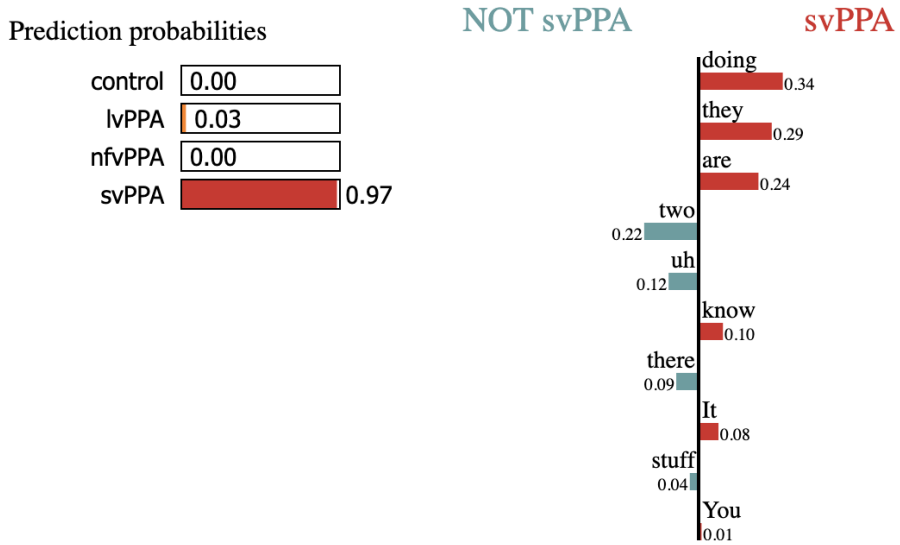
Real data sample (which we can't share in the paper)
RoBERTa's explanation



Text with highlighted words

There is two people sitting here getting ready to do some eating, reading.

Real data sample (which we can't share in the paper)
BERT's explanation



Text with highlighted words

It's two, they are there doing the, uh, stuff. You know.

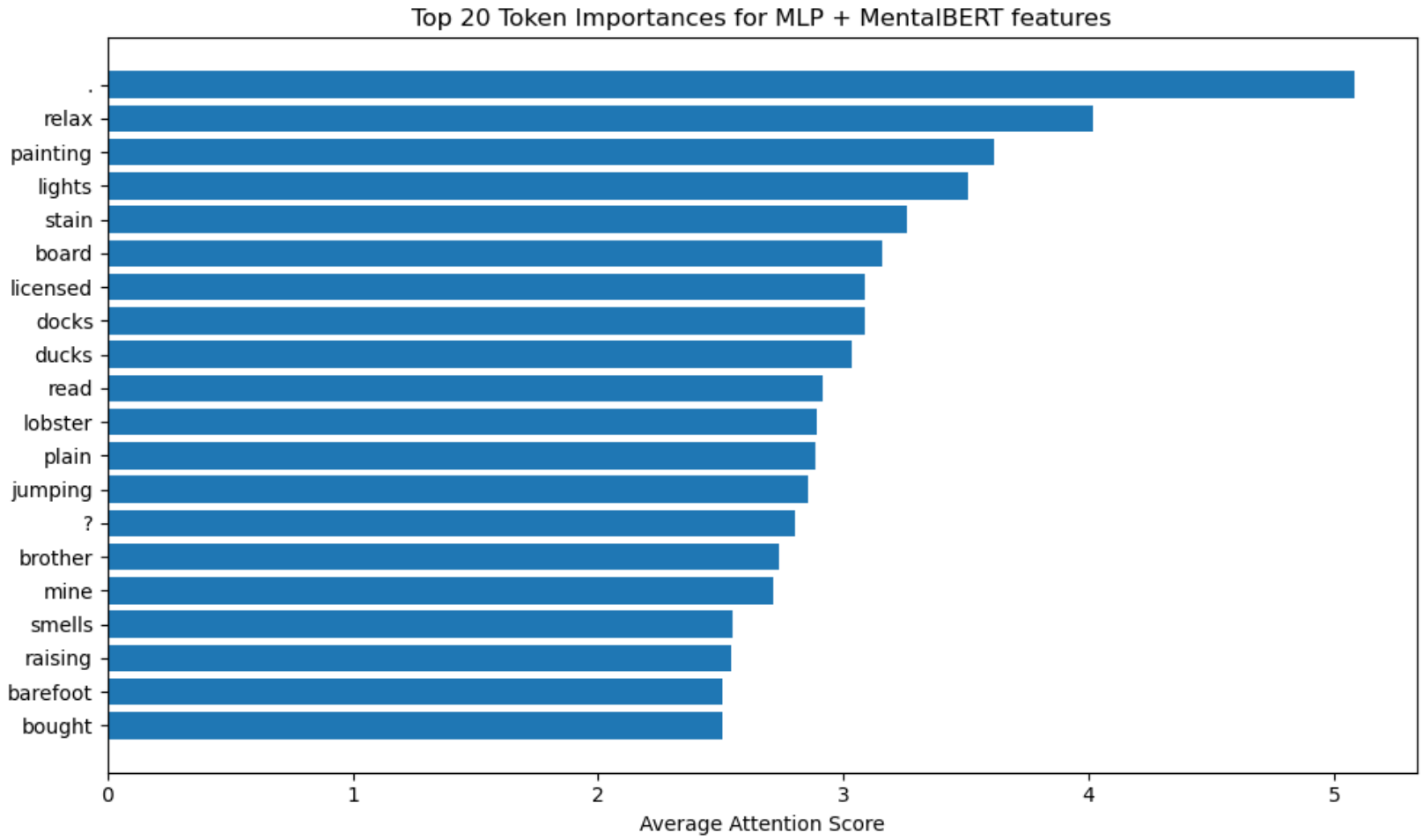
Synthetic sentence: manually generated (roBERTa)

Text with highlighted words

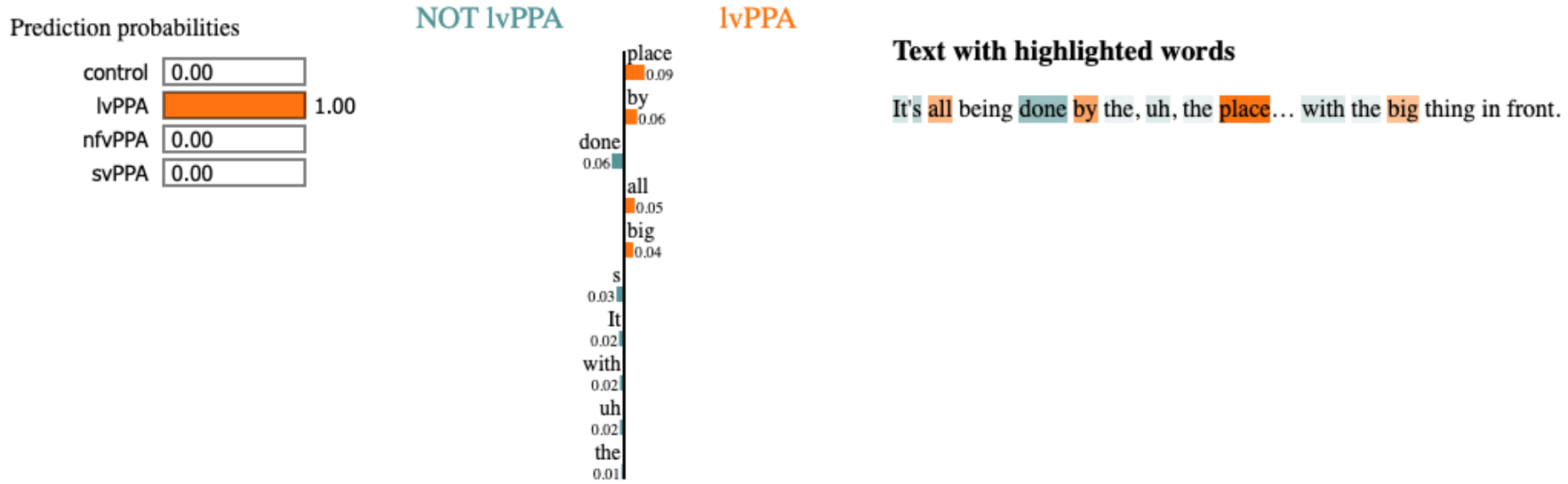
It's two, they are there doing the, uh, stuff. You know.

Synthetic sentence: manually generated (BERT)

MLP + MentalBERT’s embeddings

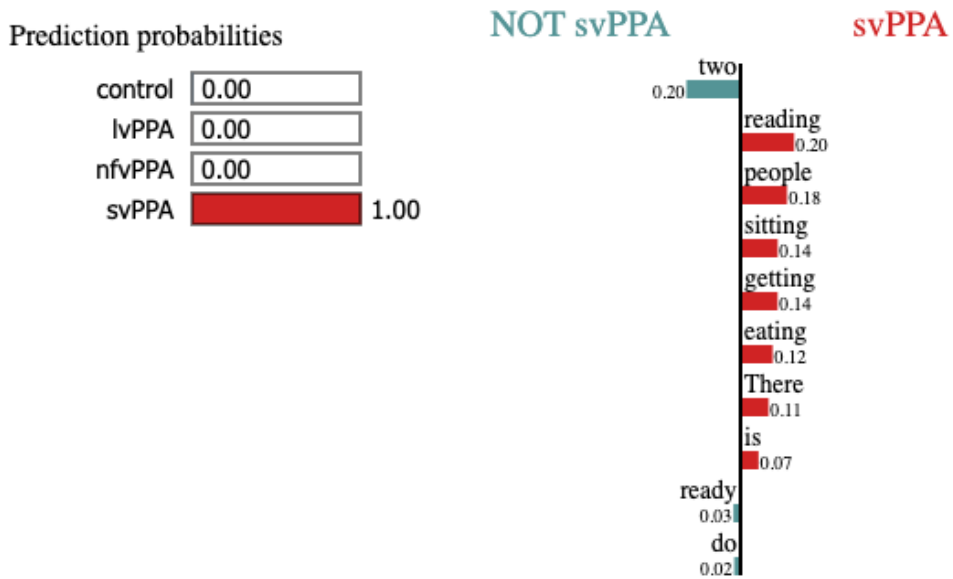


```
Prediction probabilities:

control    : 0.0000
lvPPA      : 1.0000
nfvPPA     : 0.0000
svPPA      : 0.0000
```



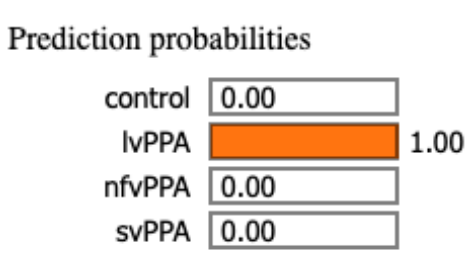
Synthetic sentence: manually generated

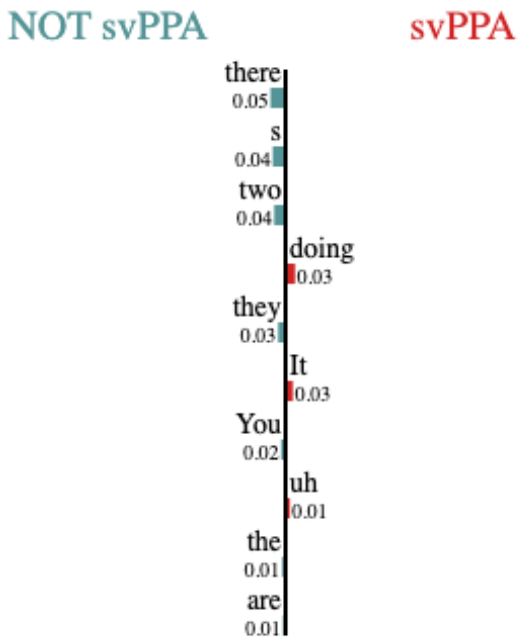
```
Prediction probabilities:

control    : 0.0034
lvPPA      : 0.0001
nfvPPA     : 0.0000
svPPA      : 0.9965
```

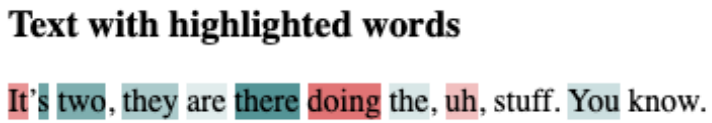


Real data sample (which we can't share in the paper)

```
Prediction probabilities:

control    : 0.0000
lvPPA      : 1.0000
nfvPPA     : 0.0000
svPPA      : 0.0000
```







Synthetic sentence: manually generated