Decoding Speech from ECoG with Machine Translation Models

Roman Burakov
Supervised by: Nadiya Shvai, Bo Wang

Background - challenges in decoding speech

There has been a significant amount of work on decoding speech from BCI. However, key challenges remain unaddressed:

1. Data collection is extremely difficult.
2. Poor generalization due to variability in brain data.
3. Natural language is complex, and brain is even more complex.

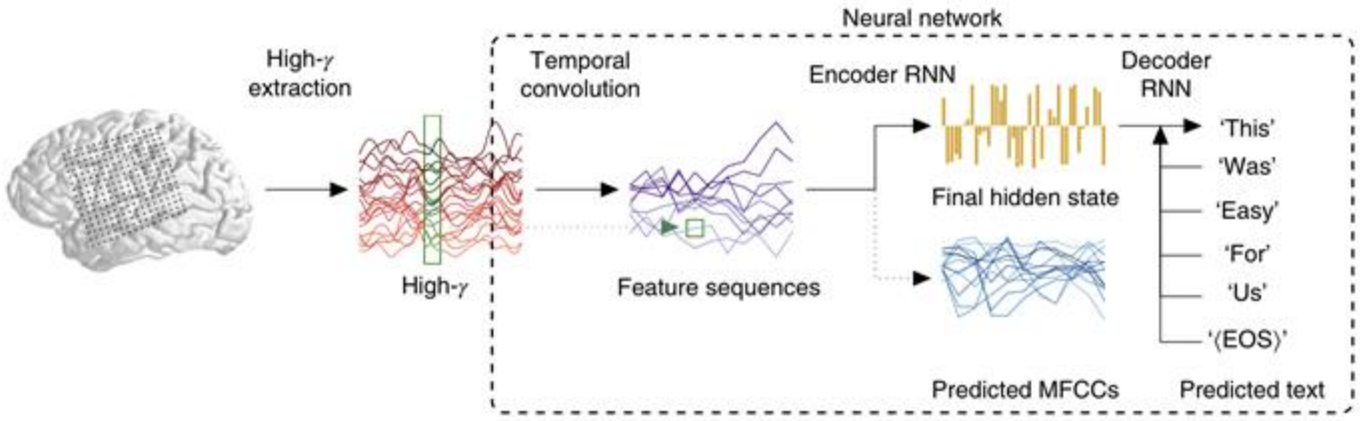
Possible addressings of these issues:

1. Transfer-learning from subject to subject.
2. Leveraging large language models trained on massive datasets.

Background - related work

Makin et al., *Machine translation of cortical activity to text with an encoder–decoder framework*, Nature Neuroscience 2020

- 4 participants with < 40 minutes of speech production each, with 30-50 unique sentences
- Transfer learning from one participant to another, with swapping the encoder



Data - 2020 Makin et al.

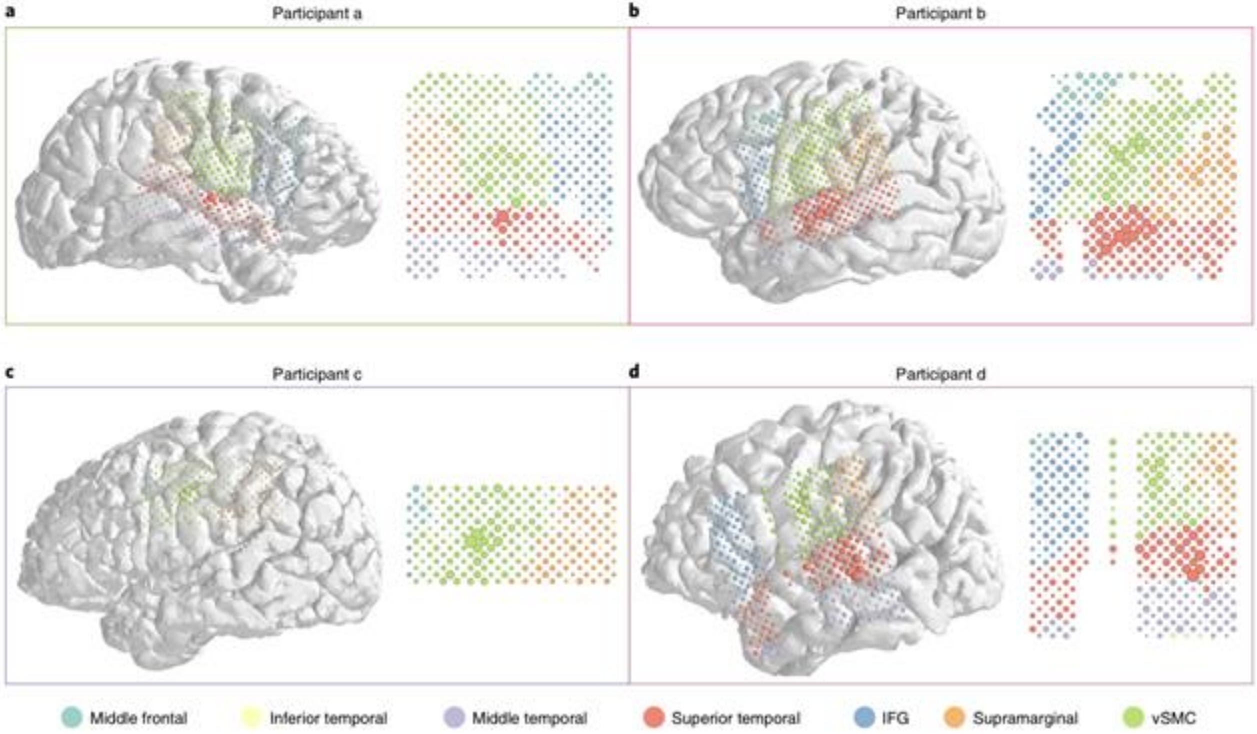
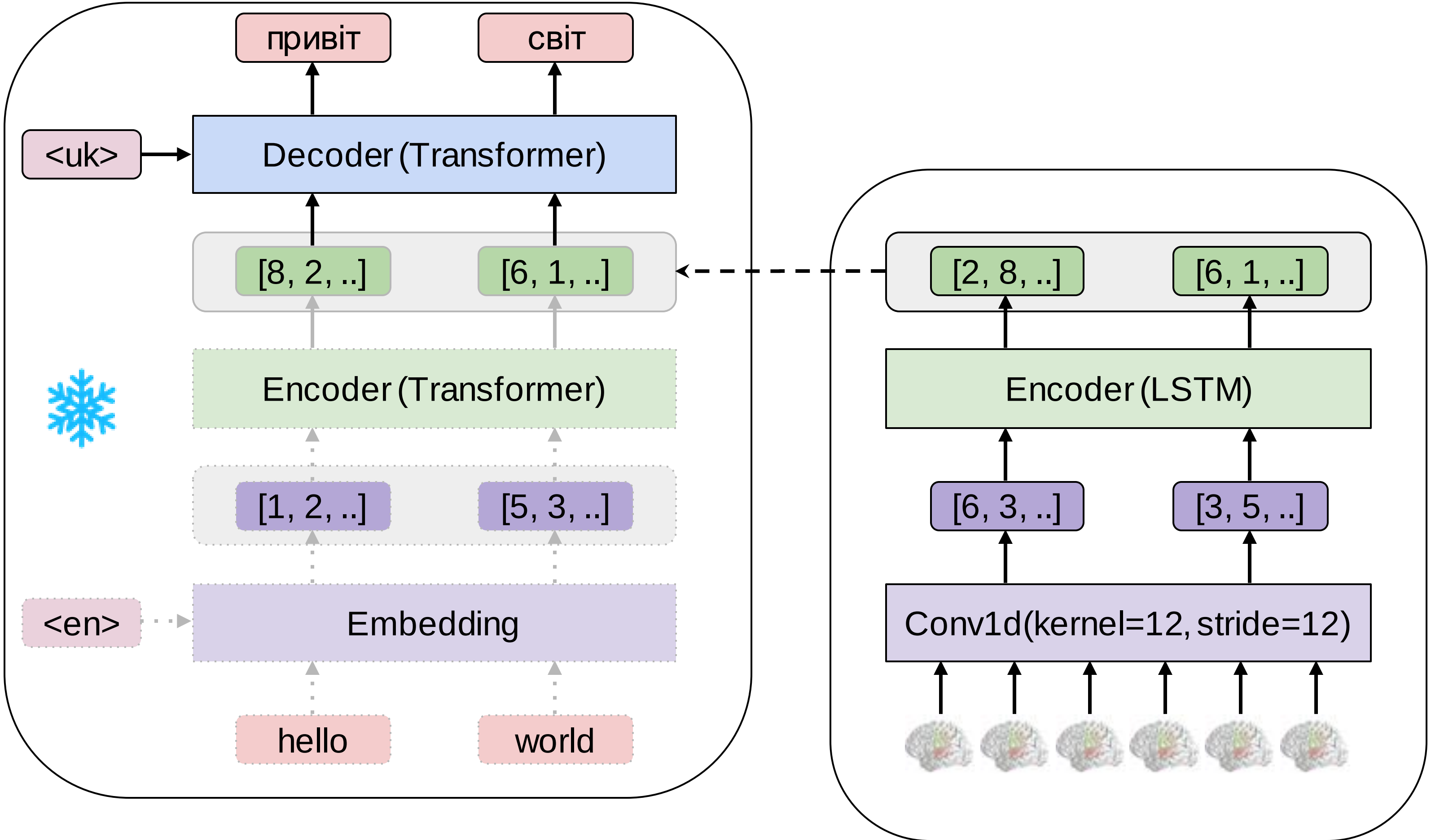


Fig. 5 | Electrode coverage and contributions. a–d, Anatomical reconstructions of the four participants (colored frames indicating participant identity according to the color scheme used throughout), with the location of the ECoG electrodes indicated with colored discs. For each disc, the area indicates the electrode's contribution to decoding (see the Methods), and the color indicates the anatomical region (see key).

Model architecture



Results - qualitative

The trained model has several significant advantages:

- Decoded sentences are generally grammatically correct.
- Decoded sentences resemble true sentences semantically, not structurally.
    - I love my mum -> I honour my family
- Native support for decoding into different languages.

However, there are also downsides:

- Semantic similarity depends on the perspective.
- Resulting method is more computationally demanding.

Results - quantitative

subject	model	WER (% ↓)	BLEU (% ↑)	BERTScore (% ↑)
a	Makin et al.	**53.5**	**12.3**	51.2
a	Ours	56.8	9.5	**71.8**
b	Makin et al.	3.4	56.1	85.1
b	Ours	**3.1**	**61.9**	**93.4**
c	Makin et al.	19.3	26.2	77.7
c	Ours	**15.0**	**29.8**	**82.0**
d	Makin et al.	**10.9**	33.7	82.5
d	Ours	11.3	**36.9**	**89.6**

Table 2: Evaluation results for each of the four subjects. Makin et al. is our implementation of [6]. While WER is not always better, we show significant improvement in BERTScore across all subjects, and in BLEU across three out of four subjects.

Thanks for your attention!