OVERT SPEECH RETRIEVAL FROM NEUROMAGNETIC SIGNALS USING WAVELETS AND ARTIFICIAL NEURAL NETWORKS

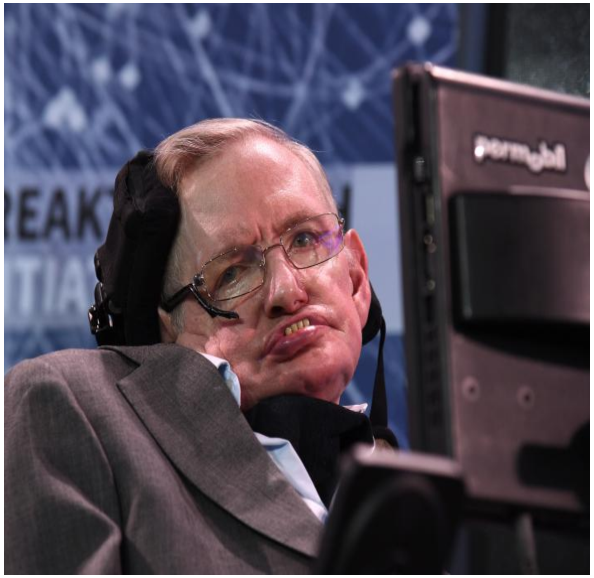
Debadatta Dash[1], Paul Ferrari[2], Saleem Malik[3], and Jun Wang[1, 4]

[1]Department of Bioengineering, The University of Texas at Dallas, TX, USA
[2]Department of Psychology, The University of Texas at Austin, TX, USA
[3]MEG Lab, Cook Children's Hospital, Fort Worth, TX, USA
[4]Callier Center for Communication Disorders, University of Texas at Dallas, TX, USA

11.28.18

Motivation

- **Speech impairment** in **Locked-in Syndrome** (**ALS** or severe brain damage) patients.
- These patients need effective Brain-Communication Interfaces (BCIs).
- Current standard BCI: Electroencephalography (EEG) – Slow (1 word/min.)
- Need for next generation, faster BCI for real-time communication

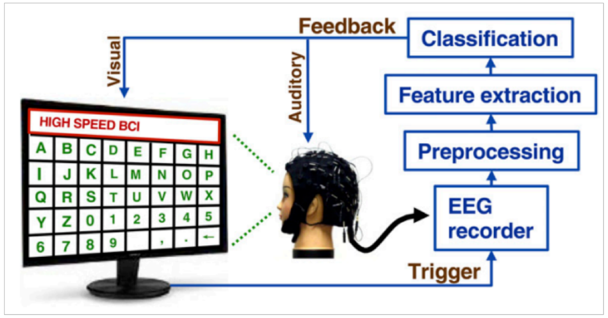


Stephen Hawking;
*Photo Courtesy: Bryan Bedder (**Medical Daily**)*

Current Research

➢ **Electroencephalography (EEG)**

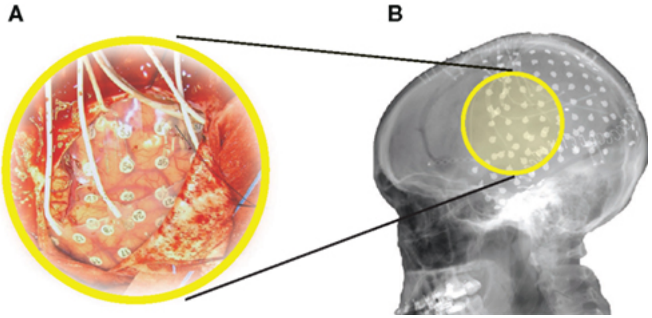
- Requires the subjects to select control characters
- Decoding visual/attention cue for word synthesis.
- Slow: Average synthesis rate: 1word/minute
- Electric field distorted at skull and scalp
- Low spatial resolution



Source: Christopher Packham, Medical Xpress

➢ **Electrocorticography (ECoG)**

- For the first time showed the possibility of direct **brain to text mapping**
- **Invasive**
- Not practical for healthy people



Source: Brunner et al., Front. NeuroSci.

Our Approach

➢ **Magnetoencephalography (MEG)**

- Measures the post-synaptic potentials induced magnetic fields during synaptic transmission.
- Optimal spatio-temporal resolution
- Non-invasive
- Reference free
- Quieter
- Undistorted magnetic field at skull and scalp

Elekta Neuromag TRIUX System

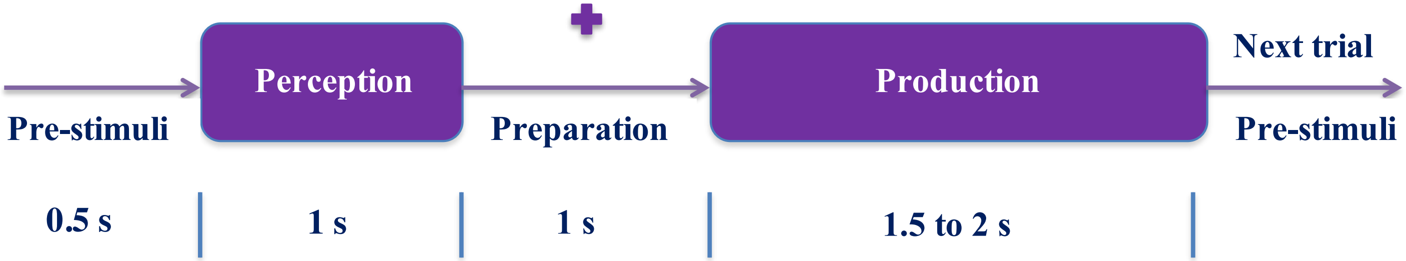
Data Collection

➢ **MEG Unit**:
- Neuromag Triux (Elekta, Ltd.)
- 306 channel
  (204 gradiometers and 102 magnetometers)

➢ **Subjects:**
- 4 (2 males + 2 females)
- Healthy with normal vision; no speech or cognitive disorders.

➢ **Stimuli**
- 100 repetitions of 5 short phrases (e.g. *I need help*.)

➢ **Protocol**



Data Preprocessing

- **Frequency:**

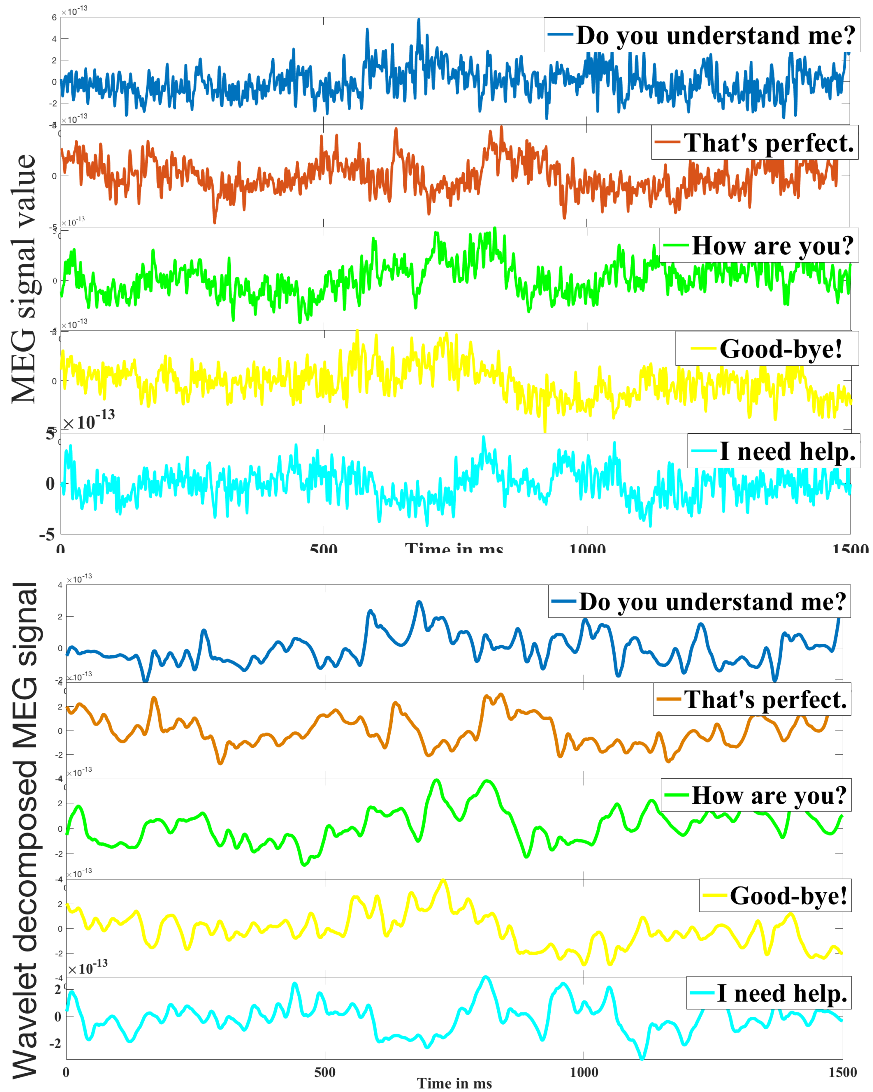
- Acquisition **–** 4 kHz; Band-pass filter - 0.1-1.3 kHz; down-sample - 1 kHz

- **Noise removal:**

1. **ECG-** Cardiac signal recording
2. **EOG-** Eye-Blink recording
3. **Background Noise**- Empty room MEG signal acquisition
4. **Head motion** – Tracking with continuous localization technique
5. **Manual Checking** – Erroneous samples and high motion artifacts

Feature Selection

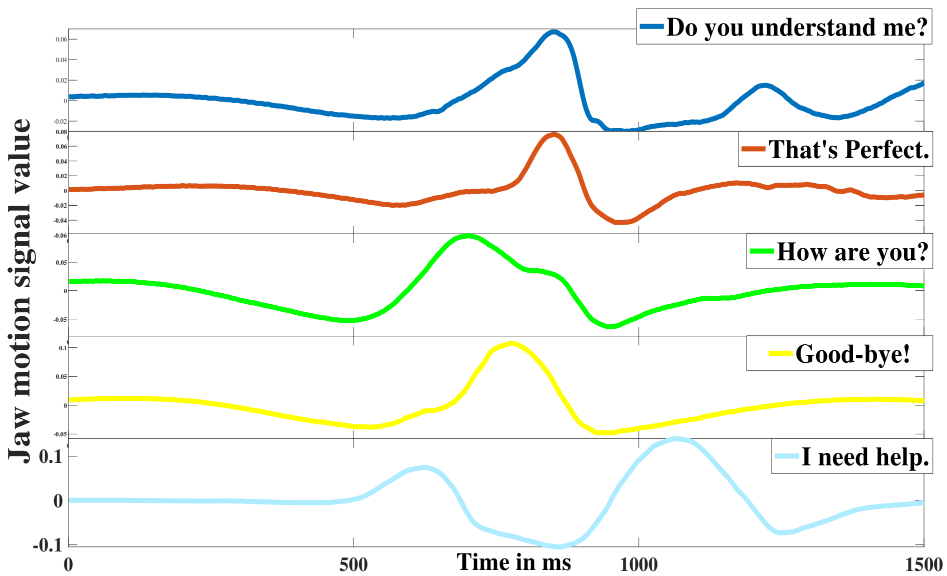
- **Wavelet Analysis**
  - Daubechies (db)-4 wavelet
  - Denoising and feature band selection
  - 3 level decomposition
  - Signal restricted to Gamma band
  - Feature – RMS values



Feature Selection

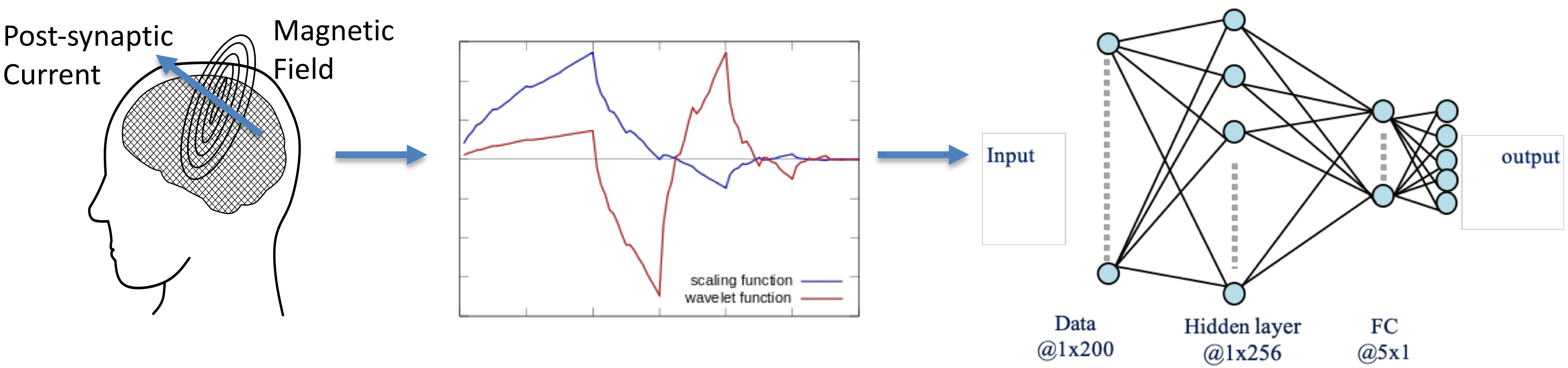
➢ **Jaw Signal**

- A custom air-pressure sensor
- Connected to an air-filled bladder
- Adhered below the jaw of the subjects
- Record the jaw motion during articulation
- Digitized in real-time as separate MEG channels.



Proposed Approach:

MEG signals → Wavelet Denoising → Feature Extraction → Decoding (ANN)



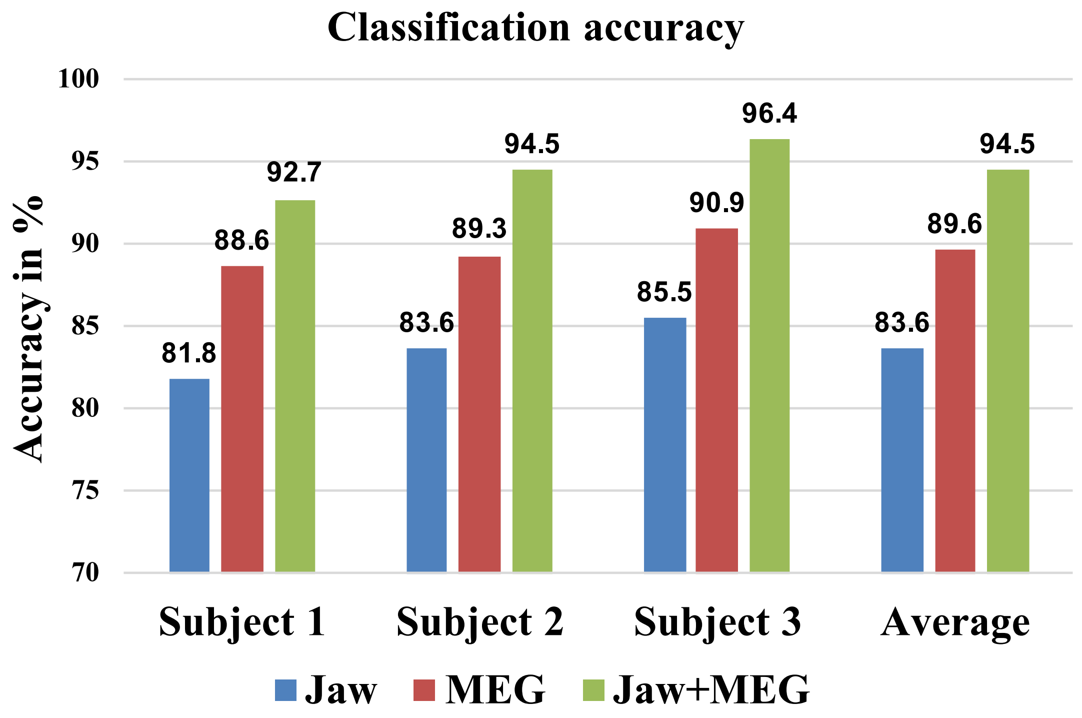
Results

- ➢ **Accuracy:**
- **Jaw < MEG < Jaw + MEG**
- Non-redundancy

  across the two signal sources
- Existence of

  complementary information

  In jaw and brain activity.



Results

- Confusion Matrices:

Test Confusion Matrix

Output Class \ Target Class	1	2	3	4	5	
1	9	0	0	0	0	100%
2	0	7	0	0	0	100%
3	0	1	10	0	0	90.9%
4	4	1	0	13	1	68.4%
5	0	1	0	0	12	92.3%
	55.6%	70%	100%	100%	92.3%	85.5%

Jaw

Test Confusion Matrix

Output Class \ Target Class	1	2	3	4	5	
1	8	0	0	2	0	80.0%
2	0	15	0	0	0	100%
3	0	1	12	0	0	92.3%
4	0	0	0	11	0	100%
5	0	1	0	1	4	66.7%
	100%	88.2%	100%	78.6%	100%	90.9%

MEG

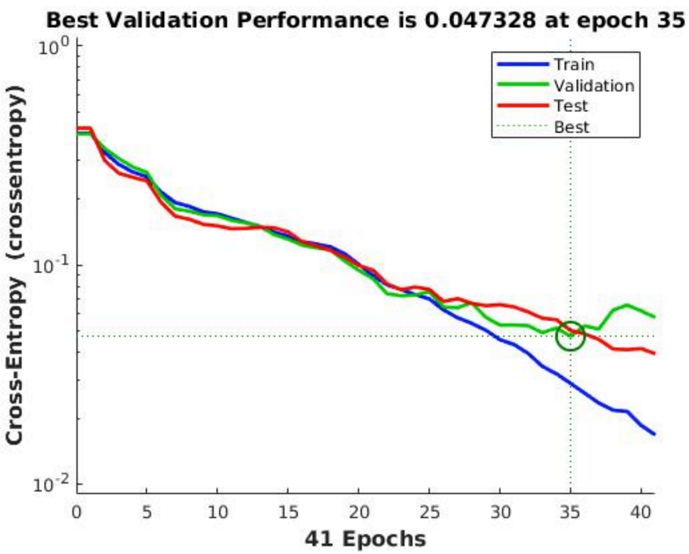
Test Confusion Matrix

Output Class \ Target Class	1	2	3	4	5	
1	5	1	0	0	0	83.3%
2	0	11	0	0	0	100%
3	0	0	13	0	0	100%
4	1	0	0	11	0	91.7%
5	0	0	0	0	13	100%
	83.3%	91.7%	100%	100%	100%	96.4%

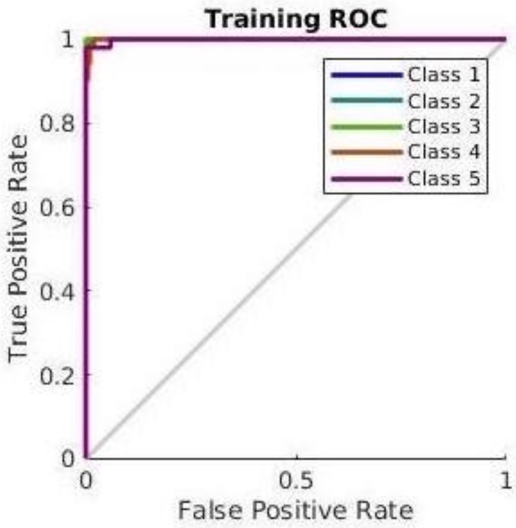
Jaw + MEG

Results

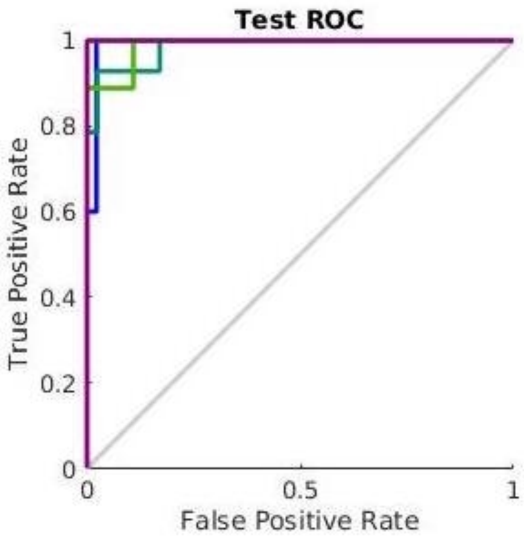
➢ Performance of ANN:



Performance curve



Training ROC



Test ROC

Conclusions

- Possibility of direct, non-invasive neural speech decoding
- Jaw and MEG contains complementary information
- Results validated through ROC and performance curves
- 90% average speech decoding accuracy

Limitations

- MEG is costly (2M $) ~ 8K-10K $
- Prune to artifacts and head motion ~ High SNR.
- Large size ~ Helmet size
- Fixed in MSR ~ Fixed to head

Current MEG scanner based on SQUID sensors

New MEG scanner with OPM sensors

Boto et al. (2018) *Nature*

Source: Boto et al. 2018, Nature.

Acknowledgments

- NIH-R03 DC013990 (PI: Wang)
- American Speech-Language-Hearing Foundation via a New Century Scholar Research Grant (PI: Wang)
- Dr. Ted Mau, Dr. Myungjong Kim, Dr. Angel W. Hernandez-Mulero
- IEEE SPS travel grant

Debadatta Dash

Dr. Paul Ferrari

Dr. Saleem Malik

Dr. Jun Wang

Thank you