Natalia Sempere 2025

Week 26

Tongue EMG XR Project

Monday to Wednesday

From Monday to Wednesday, the university was closed due to local holidays. I worked from home with the samples I recorded the previous Friday with Noraxon. However, I soon had to dismiss those recordings as I discovered I had not placed the sensors correctly.

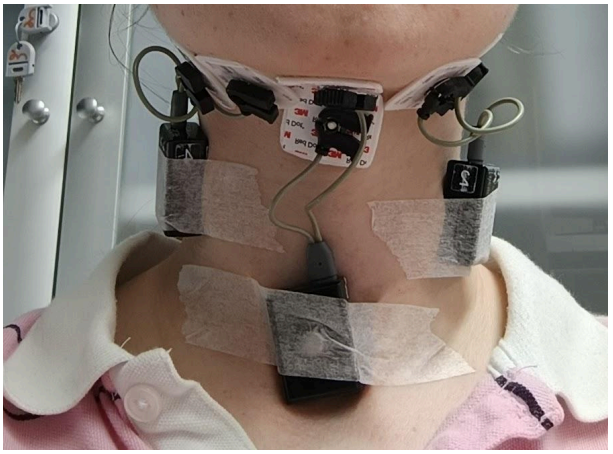

Image of the correct placement.

Thursday

I repeated the recordings with myself and tested the offline classification. The logging was done with 2 different approaches:

- **Duration-based logging:**
    a. First, start pressing the tooth with the tongue.
    b. Press the key to log the class.
    c. Wait 10 seconds while maintaining the pressure on the tongue. During this time, record 5 samples of 2 seconds each.
    d. Stop, rest and repeat.

    Offline accuracy: **0.921 (only 1 user)**
    Best results with:
    - Window size 1.5s

- Bandpass 5-90 Hz
- Notch 50 Hz
- Features: RMS, RMS_SD, ZC, WL, MAV, STD, VAR, IAV, MF
- Feature scaling: Yes
- Feature selection: No
- Classifier: estimated with grid search.
- TKEO: No
- Hilbert: Yes
- Normalization: Yes

- **Onset-based logging:**
    a. First, press the key to log the class.
    b. Move the tongue to press the tooth once and come back to the rest position.
    c. Rest and repeat.

    Offline accuracy: **0.651 (only 1 user)**
    Best results with:
    - Window size 1.5s
    - Bandpass 5-90 Hz
    - Notch 50 Hz
    - Features: RMS, RMS_SD, ZC, WL, MAV, STD, VAR, IAV, MF
    - Feature scaling: Yes
    - Feature selection: No
    - Classifier: estimated with grid search.
    - TKEO: Yes
    - Hilbert: No
    - Normalization: Yes

The most relevant findings were:

- The **Kalman filter** is computationally expensive and does not increase accuracy.
- The **Hilbert** filter works better than **TKEO** in some cases.
- Classifying according to thresholds seems risky, as they change across sessions and users.
- Duration-based logging provides better results than onset-based logging.

Natalia Sempere 2025

Friday

More data was recorded from a second user, both with duration-based and onset-based logging. This user was labelled as MAR. To ease the data capture, a very basic UI has been implemented, which shows a green light to indicate the user when to perform the tongue movement.
Then, a new classifier was trained with data from MAR and myself, obtaining the following results:

- **Duration-based logging:**
  Offline accuracy: **0.863 (2 users)**
  Best results with:
  - Window size **1.75s**
  - Bandpass **5-120 Hz**
  - Notch 50 Hz
  - Features: RMS, RMS_SD, WL, MAV, STD, VAR, IAV, MF
  - Feature scaling: Yes
  - Feature selection: No
  - Classifier: estimated with grid search.
  - TKEO: **Yes**
  - Hilbert: **No**
  - Normalization: **No**

- **Onset-based logging:**
  Offline accuracy: **0.735 (2 users)**
  Best results with:
  - Window size **1.75s**
  - Bandpass **5-120 Hz**
  - Notch 50 Hz
  - Features: RMS, RMS_SD, WL, MAV, STD, VAR, IAV, MF
  - Feature scaling: Yes
  - Feature selection: No
  - Classifier: estimated with grid search.
  - TKEO: **Yes**
  - Hilbert: **No**
  - Normalization: **No**

Summary & Conclusions

	Offline accuracy (1 user)	**Offline accuracy (2 users)**
Duration-based logging	0.921	0.863
Onset-based logging	0.651	0.735

- It seems that **duration-based** logging provides better results than **onset-based** logging.
    - However, it is still to be studied if the onset-based logging can be automatized with an onset detection instead of pressing a key from the keyboard.

Next week

- Clean the code and change variable's names to make it more understandable.
- Add Pooria's onset detection and train the classifier with two users again.
- Gather samples from 2 additional users and train the classifier again.
- Test the classifier in real-time.