

Informe inicial - Classificació de senyals EEG del cervell humà mitjançant tècniques d'aprenentatge automàtic

Informació preliminar sobre la qüestió a tractar o el problema a resoldre, especificant i comentant les fonts d'informació utilitzades

Aquesta proposta de TFG busca estudiar diferències en la percepció del ritme lingüístic. Treballa sobre una recerca anterior on l'idea era veure si la llengua materna tenia alguna influència en la percepció de les paraules segons el ritme de parla de l'idioma. S'utilitza un conjunt de dades EEG, ja processades i convertides en valors Time-Frequency, i un conjunt de dades d'estímul.

Aquestes dades venen d'un context de recerca previ [1] [2]. On es tracta investigar si és el ritme inherent d'una llengua o la llengua materna d'una persona, el que més influeix en aquesta sincronització cerebral. Per això es diferencien en l'estímul 3 grups diferenciats per el ritme de parla. Aquests són els idiomes stressed-timed, syllable-timed i mora-timed, de més a menys segons la variabilitat en la duració de les síl·labes.

L'estudi es va realitzar sobre un grup de persones de parla materna anglesa (stress-timed) i un altre grup de parla materna espanyola (syllable-timed). Segons els resultats en mètriques PLV (Phase Locking Value), que ens permet mesurar la variació en la diferència de fases. Tots els participants, independentment de la seva llengua materna, mostraven una major sincronització amb el japonès (mora-timed), una sincronització intermitja amb l'espanyol (syllable-timed) i una menor sincronització amb l'anglès (stress-timed).

El primer conjunt de dades esmentat consta de dades EEG dels participants, dividits en dos grups segons idioma natiu, gravades quan escoltaven estímuls des de diferents idiomes, cadascun amb un ritme lingüístic. Per tant, per cadascun dels àudios d'estímul tenim 20 audios d'estímul per a cadascuna de les condicions (syllable, stress i mora). Es grava la reacció de cada participant a cadascuna d'elles dues vegades i això forma un total de 120 gravacions per participant. De cadascuna guardem 64 canals entre les freqüències 3hz i 8hz. D'aquests, 4 canals són usats com a referència, i 13 són descartats ja que es troben en una posició massa exterior, tot i així es poden usar tots per a la visualització. En cada canal es registra cada 1 mil·lisegon la intensitat de la reacció durant entre 3 i 5 segons. Després de l'anàlisi TF, es guarda un valor per a cada 50 ms. Això resulta en una gran quantitat de dades que cal gestionar com s'explica més endavant.

L'anàlisi TF (Time-Frequency) sobre aquestes dades, ens ajuda a veure com la potència canvia al llarg del temps per a cada freqüència. Ens centrem en una banda entre 5 i 7Hz i visualitzarem la potència per a cada canal en una matriu. S'utilitzen els anomenats Morlet Wavelets, un tipus especial d'ona que combina la sinusoidal amb una funció gaussiana que la suavitza. S'han recollit freqüències al voltant dels 5hz, ja que s'ha comprovat que el procés pel qual l'activitat neuronal al còrtex auditiu se sincronitza amb l'envolupant d'amplitud del senyal de parla captura característiques acústiques i lingüístiques, com la síl·laba, en una freqüència al voltant de 5 Hz —una síl·laba dura aproximadament 200 ms [3] — i juga un paper important en la comprensió i intel·ligibilitat [4]

Les dades estímulo són el resultat de diferents petites frases després de passar per un procés anomenat resíntesis de Saltanaj. Aquest procés canvia algunes lletres deixant només alguns sons de manera que la paraula perd el significat però conserva la sonoritat i ritme de síl·labes.

Tenint en compte que l'estudi anterior conclueix que el ritme de la parla té un paper més important que la llengua materna en el "neural entrainment" en l'àmbit de les ones theta, en aquest treball es tractarà d'avançar en l'exploració de la classificació dels àudios segons el ritme.

Una proposta de l'objectiu del TFG i/o de fins on es vol arribar en el desenvolupament de la qüestió o problema proposat

La proposta inicial per a l'objectiu d'aquest treball es basa en la classificació dels senyals segons el ritme de la seva llengua. Altres objectius són el coneixement de l'estat de l'art i conceptes necessaris, la reducció de dimensionalitat, degut a la gran redundància entre canals, sense perdre significat i el treball de classificació lineal per a trobar diferències significatives en la freqüència de pics per a cada tipus de estímulo. És previst usar diferents tècniques de classificació per a concloure si existeixen o no aquestes diferències.

Explicació general de la metodologia que es seguirà per aconseguir els objectius proposats.

El treball és pensat per ser realitzat en python (ipynb notebooks), pròpies llibreries de python també seran usades per a la visualització de les dades o resultats.

Primerament, es farà un treball de exploració de l'estat de l'art en l'estudi de les dades de reacció a estímulo semblants [3] [4], així com de casos de reducció de dimensionalitat per a audio [5]. També es treballarà en el reconeixement de les dades, proves de visualització i domini de les llibreries a utilitzar.

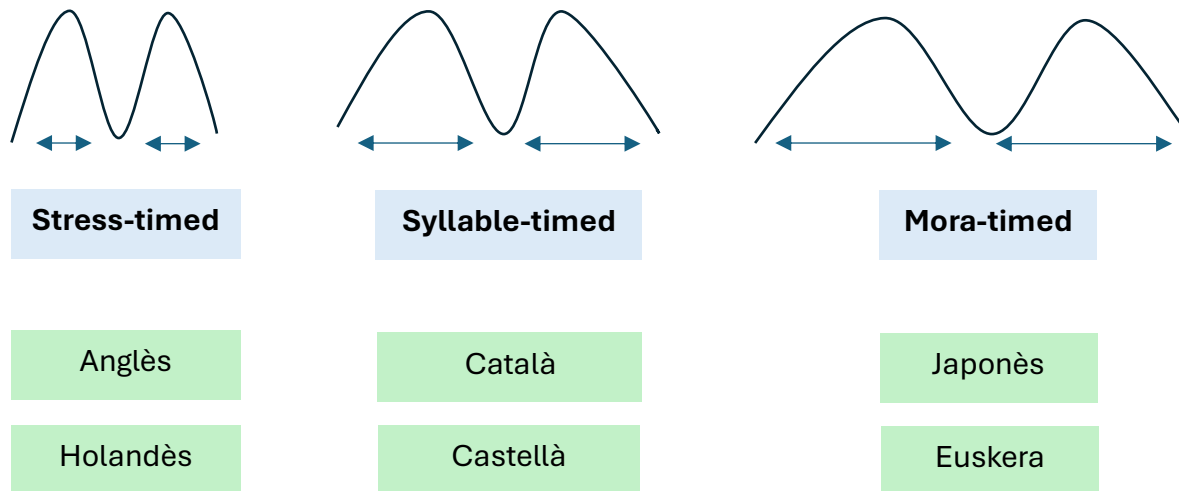
Per a aconseguir una reducció de dimensionalitat de les dades i fer-les més útils i manejables, es buscarà redundància entre els 60 canals, o en cas de que s'en descartin alguns, els restants. És previst reduir a entre 2 i 4 senyals aquests canals, guardant la informació que aporten i respectant el màxim les possibles diferències entre ells. És previst col·lapsar primerament els canals amb dades semblants, o propers físicament. Més endavant, s'usaran tècniques com PCA, SVD o altres tipus de manifold learning.

Per a la classificació, es busca trobar diferències entre els tipus de llengües, tot col·lapsant els grups de llengua materna. És previst que aquesta fase es faci repetidament ja que dependrà de la qualitat de la reducció de dimensionalitat. Altres objectius paral·lels són trobar diferències en la freqüència de pics per a cada tipus de estímulo, així com relacionar els talls de síl·laba amb la resposta EEG que obtenim. Per a aquesta part del projecte, s'utilitzaran diferents tècniques a definir com podrien ser el Machine learning simple, i els seus mètodes de classificació. En aquesta part del treball, començarem usant models més simples, i poc a poc farem servir tècniques més sofisticades.

A mesura que s'obtenen els resultats, és previst realitzar la visualització dels mateixos. Per a la reducció de canals es mostrarà gràficament quins canals són els més redundants i es

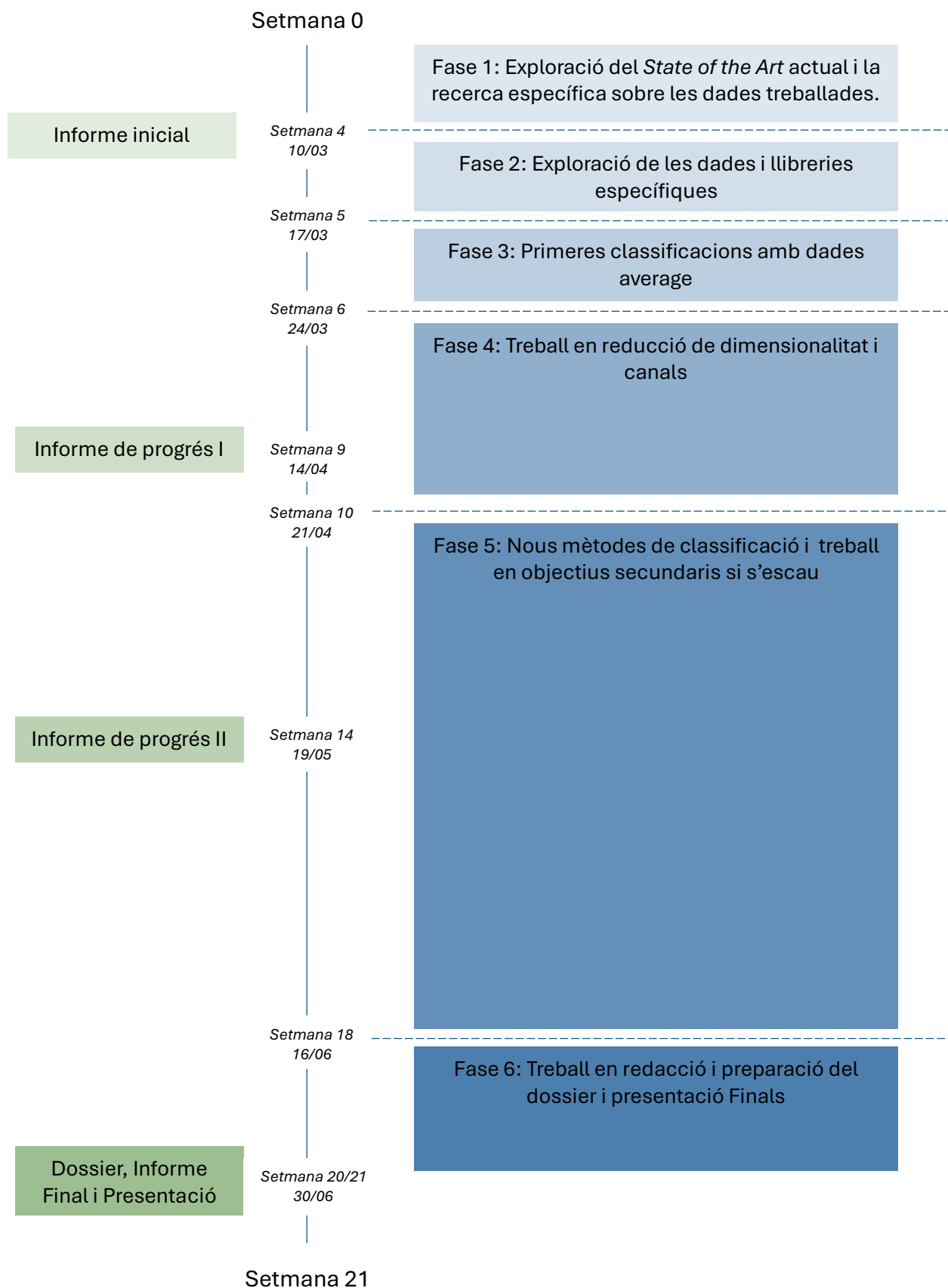
podrà pintar en diferents colors els grups de canals. I per a la resta els gràfics depenen de les tècniques decidides.

La imatge següent mostra els diferents grups en els que classifiquem les llengües segons la durada de les síl·labes. Cal dir, que tot i que es classifiquen els idiomes en grups, cal imaginar un espectre continu, que tallem per a poder treballar i comparar el llenguatge.



Il·lustració 1

Identificació dels passos a seguir per al desenvolupament del projecte proposat, establint una planificació de treball per dur-lo a terme



Bibliografia

- [1] S. Silva, «Complexity of STG signals and linguistic rhythm: a methodological study for EEG data,» *Cerebral Cortex*, vol. 34, pp. 1-13, 2024.
- [2] E. Ekin, «Neural entrainment in theta range is affected by speech properties but not by the native language of the listeners». *CBC*.
- [3] A. Perez, «Differential oscillatory encoding of foreign speech,» *Brain & Language*, vol. 147, pp. 51-57, 2015.
- [4] V. Peter, «Language specificity in cortical tracking of speech rhythm at the mora, syllable, and foot levels,» *Nature*, 2022.
- [5] T. Pál, «Comparison of Dimensionality Reduction Techniques on Audio Signals». *Eötvös Loránd University*.