**3. Transformare foneme-text**

Modulul **fonemeText.py** este un script Python ce transformă fonemele extrase de modulul transformare **audio-foneme** în cuvinte.

**Documentație legată de tema proiectului:**

Definiția termenului “fonem”: fonem sn [At: DA ms / Pl: ~e / E: fr phonème] (Lin) 1 Cea mai mică unitate sonoră a limbii, care are funcțiunea de a diferenția cuvintele între ele, precum și formele gramaticale ale aceluiași cuvânt. 2 (Înv) Sunet.

Fonologia limbii române stabilește inventarul de unități segmentale (vocale și consoane) și unități suprasegmentale (accent, intonație) cu valoare funcțională, caracteristic limbii române. Totodată, ca în fonologia oricărei alte limbi, sunt determinate seriile de sunete echivalente (alofone) prin care se realizează fiecare fonem în parte.

Prin analiza valorii funcționale a diferitelor aspecte fonetice se ajunge la concluzia că limba română standard folosește următoarele unități segmentale:

• șapte vocale,

• patru semivocale, care în unele analize sunt echivalate cu vocalele corespunzătoare,

• o vocală asilabică și devocalizată /ʲ/, care uneori este echivalată cu vocala /i/,

• douăzeci de consoane

și următoarele unități suprasegmentale:

• accent,

• intonație,

fiecare dintre acestea putându-se realiza fonetic în mai multe moduri echivalente fonologic.

Lipsesc așadar din inventarul fonematic al limbii române numeroase vocale și consoane, clicuri, croneme, tonuri etc. Astfel de aspecte fonetice pot apărea totuși de exemplu în interjecții, pronunții idiolectice, regionale sau arhaice, ori în împrumuturi neadaptate.

Unitățile segmentale sunt sunetele simple (vocale și consoane) din care se alcătuiesc silabele și apoi cuvintele. Din punct de vedere funcțional, existența a două categorii distincte de sunete --- vocale și consoane --- și diferența dintre ele se poate demonstra astfel:

1. Se constată că seria [a, e, i, ...] și seria [b, k, d, ...] sunt incompatibile, în sensul că dacă în cadrul fiecărei serii în parte prin înlocuirea unui sunet cu altul din aceeași serie se obține fie un cuvânt cu un alt sens fie un cuvânt perceput ca posibil, dar fără sens (de exemplu par-por sau par-car), înlocuirea unui sunet dintr-o serie cu un sunet din cealaltă serie nu este posibilă fără schimbarea structurii cuvântului (de exemplu aer-cer, unde numărul de silabe se modifică).

2. Nu este posibilă eliminarea dintr-un cuvânt a unui sunet din seria [a, e, i, ...] (înlocuirea lui cu zero) fără ca structura cuvântului să se modifice. În schimb sunetele din seria [b, k, d, ...] pot lipsi din cuvânt.

Astfel se face distincție între vocale, ca sunete care pot alcătui singure o silabă și care pot purta accentul, și consoane, care nu au această calitate. Există desigur și alte definiții pentru vocale și consoane, care au în vedere modul de articulare (deci aspectul fonetic) sau proprietățile acustice ale acestora, dar în subiectul de față interesează cu precădere relațiile pe care le contractează diversele sunete. Trebuie remarcat de asemenea că o astfel de analiză funcțională mai are avantajul de a releva rolul de vocală sau consoană al unui sunet în particular pentru limba română; în alte limbi aceleași sunete pot juca alte roluri, de exemplu în limba engleză sunetele [l] și [m] pot funcționa în anumite condiții ca vocale.

Construirea cuvintelor se face pe baza unui dicționar în care este descrisă scrierea fonetică a cuvintelor. Rezultatul va fi salvat într-un fișier text folosind scriptul Python textul\_final.py, pentru a putea fi preluat de modulul aliniere texte.

Observații:

Folosind funcția de generare text din audio din librăria PocketSphinx, am obținut cuvinte mult mai apropiate de textul real decât încercând să generam foneme din audio, și apoi transformând fonemele în text.

Textul obținut:

“to the end of world war two world was split into two are

east and west

this marked the beginning of the air are cold the cold war”

Textul original:

“After the end of the world war two, the world was split into two: east and west.

This marked the beginning of the era called the cold war”

După cum se observă, cuvintele pot fi mult mai ușor aproximate, algoritmul de generare audio-text fiind mult mai eficient decât audio-foneme.