Assignment 3

Question 1:

A) For each word, identify the native pronunciation and then provide the Indian non-native pronunciation based on the given rule (format: previous phoneme – target phoneme – next phoneme: Indian specific variation).

Words and Rules:

1. **Word:** "fish", **Rule:** Any - /f/ - Any : Substituted with /p^h/ .
2. **Word:** "film", **Rule:** None - consonant - consonant: vowel is inserted before or within both the consonants.
3. **Word:** "vowel", **Rule:** Any - /v,w/ - Any : Substituted with /b^h/ .
4. **Word:** "measure", **Rule:** Any - /tʃ, dʒ, s, z, ʃ, ʒ / - Any : Substituted with /Es/ or /Ez/ or /əz/.
5. **Word:** "lamp", **Rule:** Nasal – Plosive – Any : Plosive is voiced

B) Write a code (your preferred language) to automate this process for the words in lexicon (link), applying the five rules mentioned above.

Ans: A)

1. Word: "fish"

- **Native Pronunciation**: /fɪʃ/
- **Rule**: Any - /f/ - Any: Substituted with /ph/.
- **Indian Non-native Pronunciation**: /phɪʃ/
- The phoneme /f/ in "fish" is substituted with the aspirated voiceless bilabial plosive /p^h/, resulting in /p^hɪʃ/.

2. Word: "film"

- **Native Pronunciation**: /fɪlm/
- **Rule**: None - consonant - consonant: Insert a vowel before or within the consonants.
- **Indian Non-native Pronunciation**: /fɪləm/ (a schwa is inserted to break up the consonant cluster)
- A vowel (such as /ə/ or /ʊ/) is inserted between the consonant cluster /l/ and /m/, breaking the cluster and making it easier to pronounce, resulting in /fɪləm/ or /fɪlʊm/.

3. Word: "vowel"

- **Native Pronunciation**: /vaʊəl/
- **Rule**: Any - /v, w/ - Any: Substitute with /bh/.
- **Indian Non-native Pronunciation**: /bhaʊəl/
- The phoneme /v/ in "vowel" is substituted with the voiced aspirated bilabial plosive /bʰ/, resulting in /ˈbʰaʊəl/.

4. Word: "measure"

- **Native Pronunciation**: /ˈmɛʒər/
- **Rule**: Any - /tʃ, dʒ, s, z, ʃ, ʒ / - Any: Substitute with /Es/ or /Ez/ or /əz/.
- **Indian Non-native Pronunciation**: /ˈmɛzər/ or /ˈmɛzʊr/ (depending on the choice of substitution)
- The phoneme /ʒ/ in "measure" is substituted with /z/ or /s/, often accompanied by an added vowel sound, resulting in pronunciations like /ˈmɛzər/, /ˈmɛsər/, or /ˈmɛʒəz/.

5. Word: "lamp"

- **Native Pronunciation**: /læmp/
- **Rule**: Nasal - Plosive - Any: Plosive is voiced.
- **Indian Non-native Pronunciation**: /læmb/
- The voiceless plosive /p/ following the nasal /m/ is voiced to /b/, resulting in /læmb/.

Question 2:

Apply the following phonological rules to the given words:

- **Aspiration:** Voiceless stops are aspirated at the beginning of words.
- **Flapping:** /t/ and /d/ between two vowels are realized as a flap.
- **Palatalization:** /k/ and /g/ before /y/ are palatalized.

1. pot
2. butter
3. city
4. kite
5. cute

1. **pot**: /P AA T/
   The voiceless stop /P/ at the beginning of the word is aspirated, producing a puff of air.
2. **butter**: /B AH DX ER/
   The /T/ sound between two vowels is realized as a flap /DX/, a softer sound where the tongue quickly taps the roof of the mouth.
3. **city**: /S IH DX IY/
   Similarly, the /T/ between vowels is flapped, resulting in the sound /DX/.
4. **kite**: /K AY T/
   The voiceless stop /K/ at the start of the word is aspirated, adding a puff of air to its pronunciation.
5. **cute**: /K Y UW T/
   The /K/ is both aspirated and palatalized because it precedes the /Y/ sound, bringing it closer to the palate when pronounced.

Question 3:

Analyze the accentual patterns in the following phrases and sentences. For each example, identify the stressed syllables, describe the overall intonation, and explain how stress placement affects meaning or emotional tone.

1. blackboard
2. She loves to dance.
3. The cat is on the mat.
4. He is a good boy.

1. Blackboard

- **Stressed Syllable**:
  - **Stressed**: *Black*
  - **Unstressed**: *board*
- **Overall Intonation**:
  - **Stress Pattern**: "Blackboard" is a compound noun with primary stress on the first syllable (*black*).
  - **Intonation Pattern**: After the stressed syllable (*black*), the pitch dips slightly and remains flat through the unstressed syllable (*board*).
- **Effect of Stress Placement**:
  - The stress on *black* clarifies that "blackboard" is a unified concept, referring to a specific classroom object. If stress were shifted to *board*, it could imply a general description of a board that is black in color, altering the meaning.

2. She loves to dance

- **Stressed Syllables**:
  - **Stressed**: *Loves*
  - **Unstressed**: *She*, *to*, *dance* (though *dance* may carry slight emphasis due to its role as the object).
- **Overall Intonation**:
  - **Stress Pattern**: The primary stress falls on the verb *loves*.
  - **Intonation Pattern**: The sentence follows a natural rise and fall typical of declarative statements. The pitch rises slightly on *loves* and falls towards the end, especially on *dance*.
- **Effect of Stress Placement**:
  - Emphasizing *loves* conveys strong emotion or affection for dancing. Stressing *she* instead might shift focus to who loves to dance, while stressing *dance* emphasizes the activity itself, potentially changing the emotional tone or meaning.

3. The cat is on the mat

- **Stressed Syllables**:
  - **Stressed**: *Cat*, *mat*
  - **Unstressed**: *The*, *is*, *on*, *the*
- **Overall Intonation**:
  - **Stress Pattern**: Stress typically falls on the key content words, *cat* and *mat*, as they carry the most meaning.
  - **Intonation Pattern**: A falling intonation is common, as this is a simple declarative sentence. The pitch may peak slightly on *cat* and descend steadily through *is on the mat*.
- **Effect of Stress Placement**:
  - Stress on *cat* and *mat* highlights the core elements of the statement. If stress were placed on *on*, it might suggest emphasis on the relationship or location, subtly changing the sentence's focus or intention.

4. He is a good boy

- **Stressed Syllables**:
  - **Stressed**: *Good*, *boy*
  - **Unstressed**: *He*, *is*, *a*
- **Overall Intonation**:
  - **Stress Pattern**: The main stress usually falls on *good* and *boy*, emphasizing the descriptive phrase.
  - **Intonation Pattern**: A falling intonation is typical, with a slight rise on *good* and a final drop on *boy*.
- **Effect of Stress Placement**:

- Stressing *good* and *boy* emphasizes the positive quality and subject of the sentence. Changing stress to *He* might shift focus to the subject, implying contrast (e.g., *He* is a good boy, not someone else). Stress on *is* could add a tone of affirmation or correction.

Question 4:

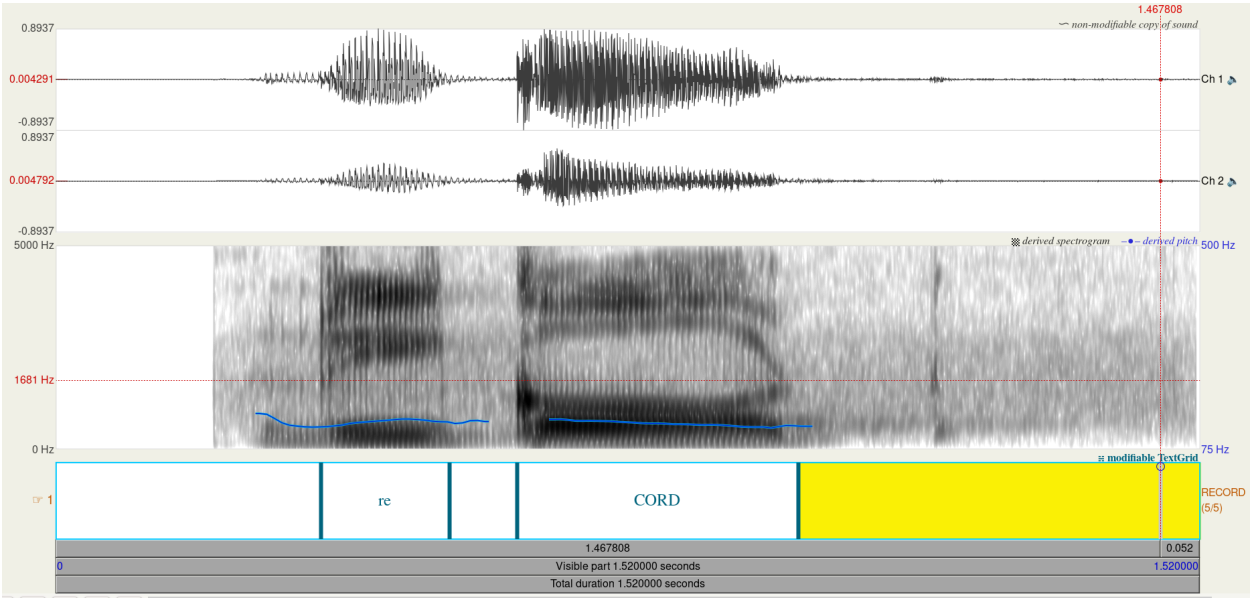
Select a two-syllable word that has different meanings based on stress placement. Record it (atleast 2 variations) with change in stress and report the energy, duration, and pitch at the syllable level. Use any forced alignment tool to determine the syllable boundaries.

I selected the two-syllable word "**RECORD**," which changes meaning depending on stress placement. I recorded two variations, altering the stress, and analyzed the energy, duration, and pitch at the syllable level. Using a forced alignment tool, I identified the syllable boundaries for precise analysis.

- **Variation 1 (re-CORD)**: When stress is on the second syllable (*CORD*), "record" is used as a verb, meaning to capture or document information, sounds, or visuals. For example, one might *record* a video or log data into a system.
- **Variation 2 (RE-cord)**: When stress is on the first syllable (*RE*), "record" functions as a noun, referring to an object or medium that stores information, such as a vinyl *record*, a written account, or documentation of events.

Energy, Duration and Pitch Analysis:-

1) **reCORD (Verb)**

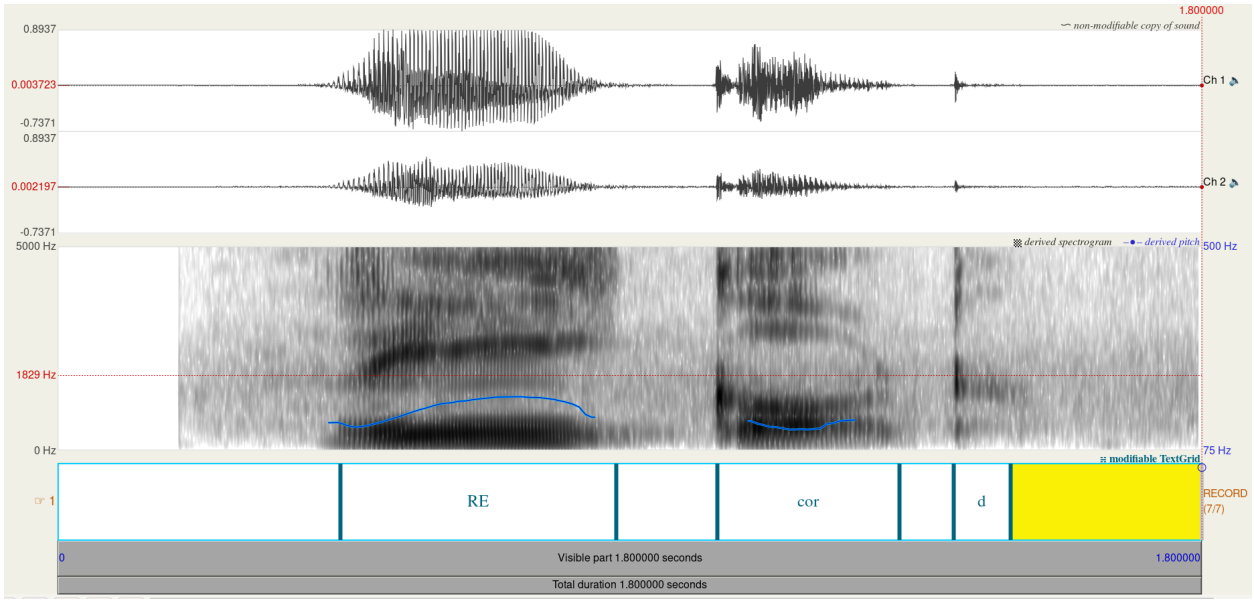


Syllable 1: re
Syllable 2: CORD

```
Syllable 1:
  Start Time: 0.35204 seconds
  End Time: 0.52287 seconds
  Duration: 0.17082999999999993 seconds
  Energy: 95.89808654785156
  Pitch: 366.5113830566406 Hz

Syllable 2:
  Start Time: 0.61274 seconds
  End Time: 0.98644 seconds
  Duration: 0.37370000000000003 seconds
  Energy: 289.0708312988281
  Pitch: 638.5459594726562 Hz
```

2) REcord (Noun)



Syllable 1: RE

Syllable 2: cord

```
Syllable 1:
  Start Time: 0.44428 seconds
  End Time: 0.87839 seconds
  Duration: 0.43411 seconds
  Energy: 440.1026306152344
  Pitch: 351.440185546875 Hz

Syllable 2:
  Start Time: 1.0375 seconds
  End Time: 1.3241 seconds
  Duration: 0.28659999999999997 seconds
  Energy: 59.16623306274414
  Pitch: 674.3868408203125 Hz
```

Question 5:

Using the following pairs of words, analyze the differences in their phonetic transcriptions and derive a general rule for Grapheme-to-Phoneme (G2P) conversion:

Word Pairs:

- "bit" /bɪt/
- "beat" /biːt/
- "kit" /kɪt/
- "key" /kiː/

1) bit (/bɪt/) vs. beat (/biːt/)

- **Grapheme**:
  In *bit*, the vowel is represented by a single "i," while in *beat*, it is represented by the combination "ea."
- **Phoneme**:
  - In *bit*, the vowel sound is /ɪ/, a short vowel.
  - In *beat*, the vowel sound is /iː/, a long vowel.
- **Pattern**:
  A single "i" typically corresponds to the short vowel /ɪ/, while "ea" is associated with the long vowel /iː/.

2) kit (/kɪt/) vs. key (/kiː/)

- **Grapheme**:
  Kit uses a single "i" for the vowel, while *key* uses "ey."
- **Phoneme**:
  - In *kit*, the vowel sound is /ɪ/, a short vowel.
  - In *key*, the vowel sound is /iː/, a long vowel.
- **Pattern**:
  Similar to the first pair, a single "i" corresponds to the short vowel /ɪ/, whereas "ey" represents the long vowel /iː/.

General Rule for Grapheme-to-Phoneme (G2P) Conversion:

From this analysis, we can derive the following rules for vowel-to-sound mapping:

1. **Single "i" in Closed Syllables**:
   - When the letter "i" appears in a closed syllable, it generally corresponds to the short vowel sound /ɪ/.
   - **Examples**: *bit* → /bɪt/, *kit* → /kɪt/.
2. **Vowel Combinations "ea" and "ey"**:

- The combinations "ea" and "ey" are often pronounced with the long vowel sound /iː/.
- **Examples**: *beat* → /biːt/, *key* → /kiː/.

Question 6:

Can we detect syllable stress using Automatic Speech Recognition (ASR)? If so, which components of the ASR system should be modified, and how would these modifications enhance the detection of syllable stress particularly for non-native speakers of English?

Yes, ASR systems can be adapted to detect syllable stress by incorporating specific modifications to their components:

ASR System Modifications

1. **Expand the Phoneme Inventory:**
   - Add distinct phonemes for stressed (e.g., /AA1/) and unstressed (e.g., /AA0/) vowels to differentiate stress patterns in speech.
2. **Enhance the Pronunciation Dictionary:**
   - Include multiple variants of word pronunciations, each reflecting a different stress pattern.
   - **Example:** The word *banana* could have variations such as /B AA1 N AA0 N AA0/, /B AA0 N AA1 N AA0/, etc., to account for possible stress placements.
3. **Refine the Acoustic Model:**
   - Train the acoustic model with annotated datasets that mark stressed and unstressed vowels, enabling it to recognize stress distinctions.
4. **Implement Forced Alignment:**
   - Use forced alignment to map the spoken input to the most probable stress pattern, helping identify stressed syllables accurately.

Advantages for Non-Native Speakers

1. **Enhanced Stress Recognition:**
   - The system can pinpoint incorrect stress patterns, a frequent challenge for non-native speakers, aiding in diagnosis and improvement.
2. **Personalized Pronunciation Feedback:**
   - Provides targeted feedback on stress placement, supporting learners in correcting their speech production effectively.
3. **Efficient Stress Modeling:**
   - Focusing on vowel-level modeling simplifies the detection process and makes stress recognition more efficient.