

# Annotation

Issues and Strategies in Annotation at Phoneme Level

Presenter: Mahwish Farooq

1st July, 2014

Center for Language Engineering (CLE)

# Basic Rules of Annotation

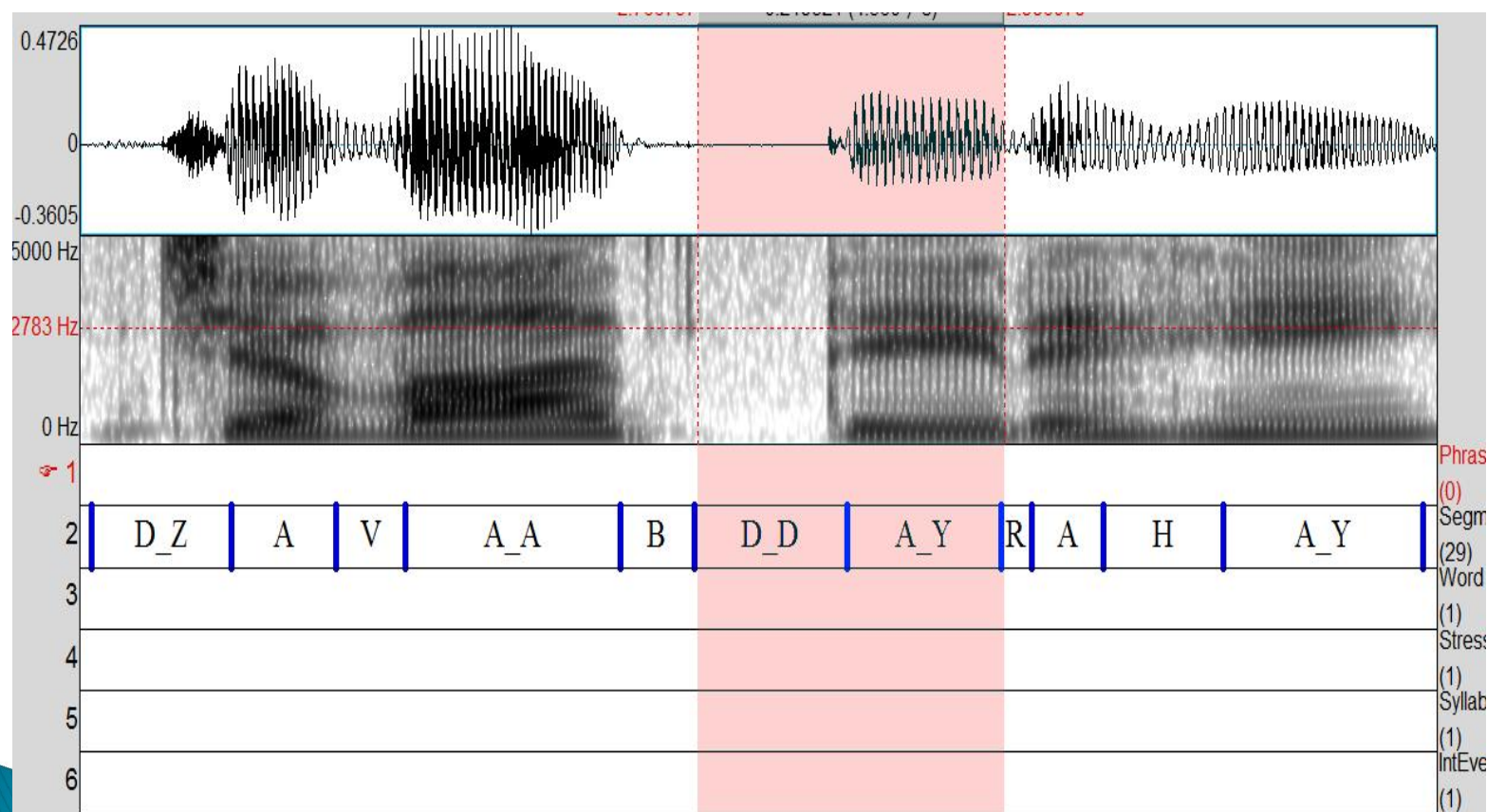
- ▶ Phonological labeling
- ▶ Mark silence and pauses
- ▶ Each consonant and vowel is distinctly marked in the Text Grid file.
- ▶ Mark at the zero crossing line going from negative to positive value
- ▶ Marking **vowel + consonant** combination
- ▶ Marking **vowel + vowel** combination
- ▶ Marking **consonant + consonant** combination

# Issues Faced In Annotation

- Co-articulation Effect
- Glottalization
- Gemination
- Nasalization
- Deletion of sounds
- Issue of /J/ “يا” sound
- Diphthongs

# 1. Co-articulation Effect

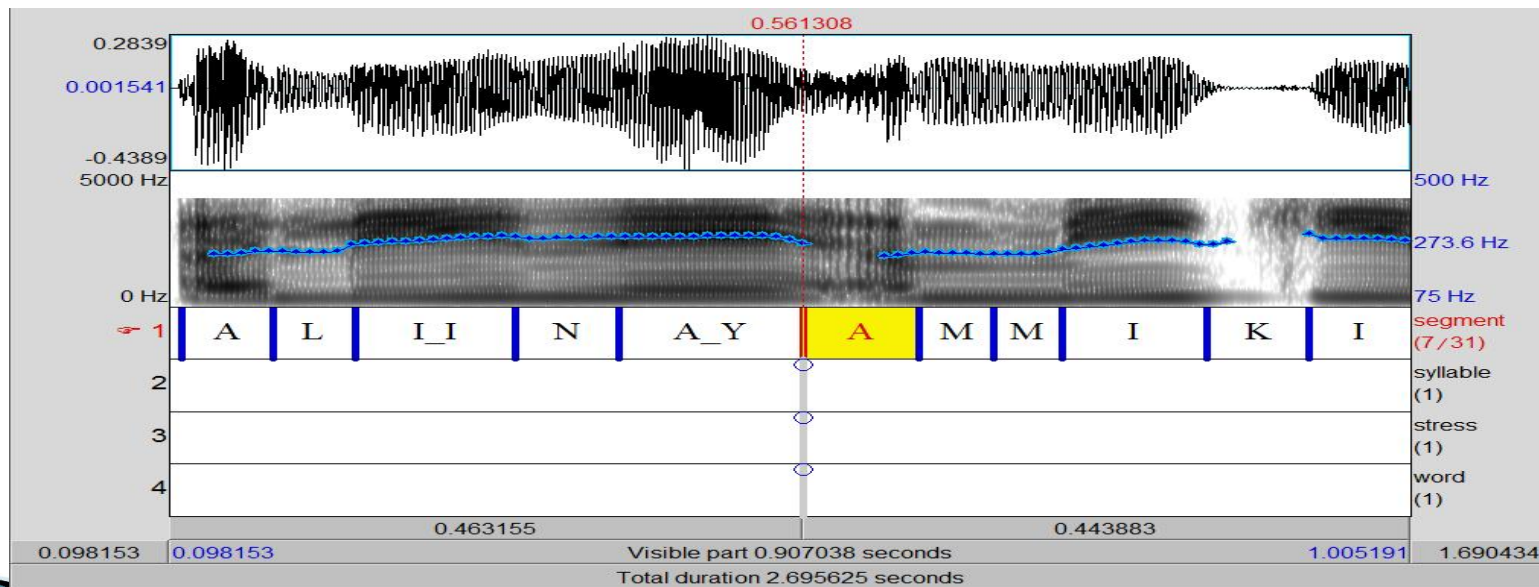
Phonemic pairs lose their characteristics i.e. lose voicing, aspiration, frication and vice versa.



## 2. Glottalization

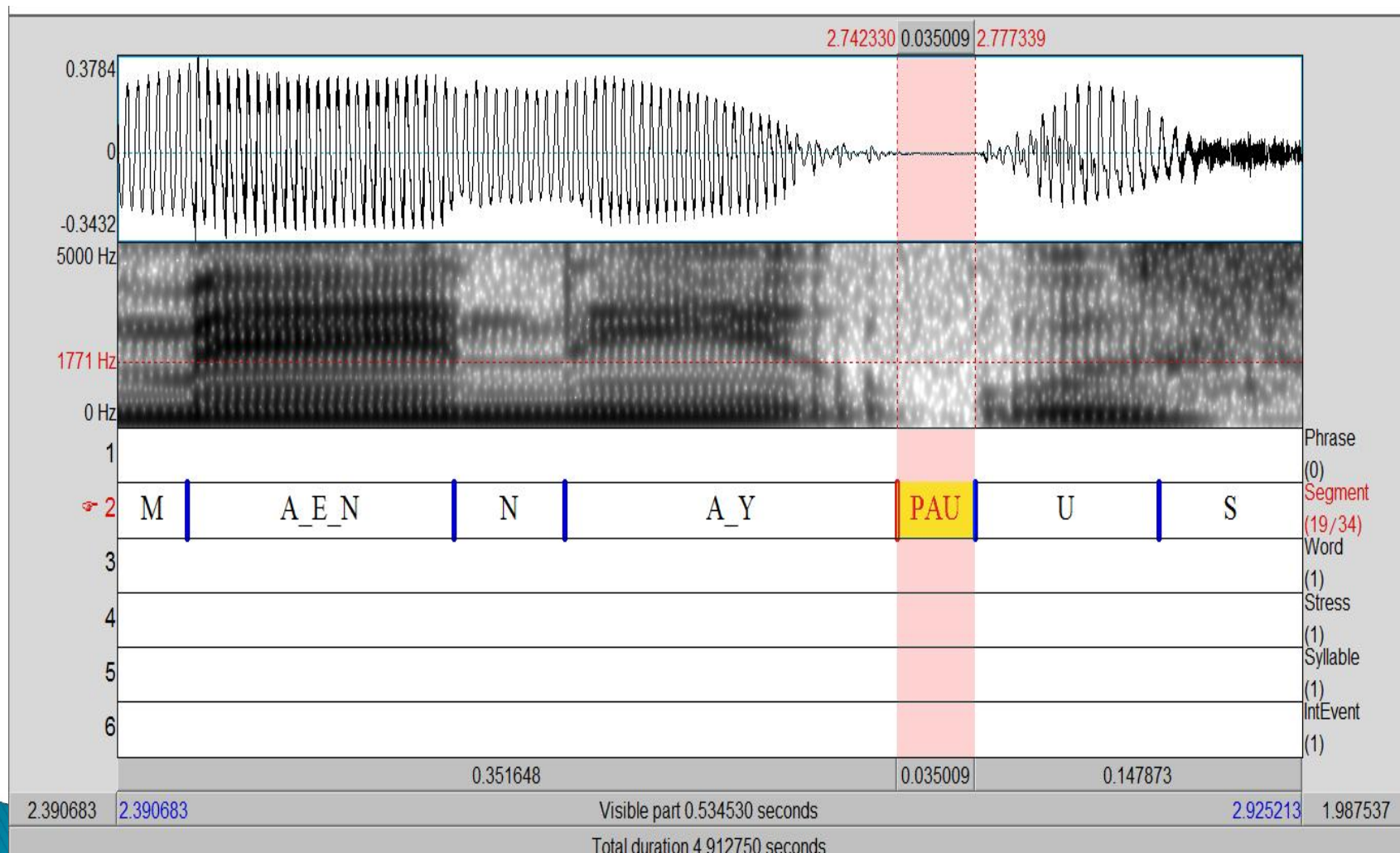
During the articulation of voiced sounds, the partial or complete closure of glottis is called glottalization.

- ▶ In a sentence, if the click or glottal stop occurs between two words these should be added to the respective sound which it is a part of.





# Glottalization: Complete closure



## 3. Gemination

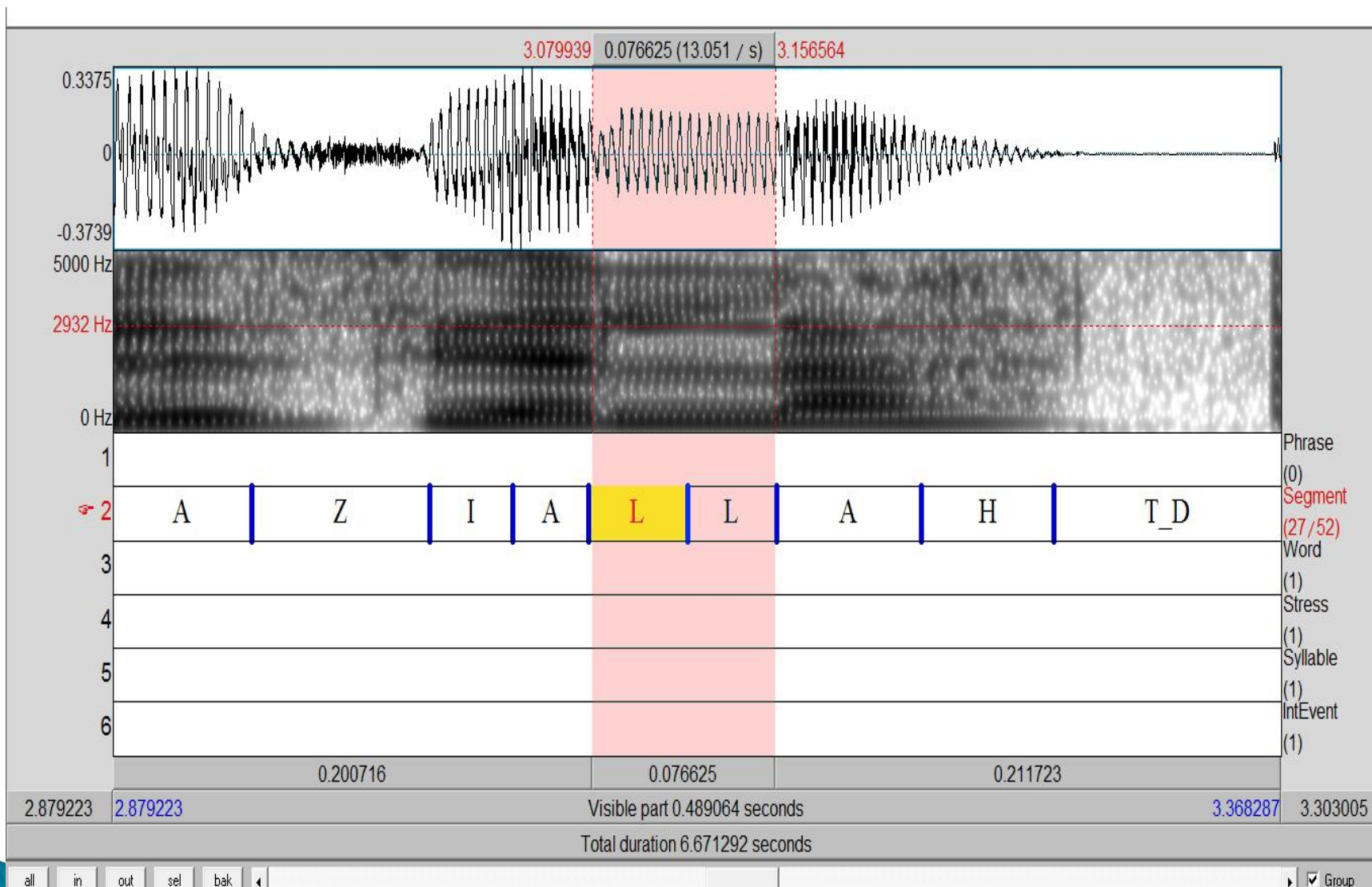
### **Within words and across the words**

In case of gemination across the words or within the word, the sound is divided in two equal halves.

- US+ SA\_Y
- AM+MA\_A\_N

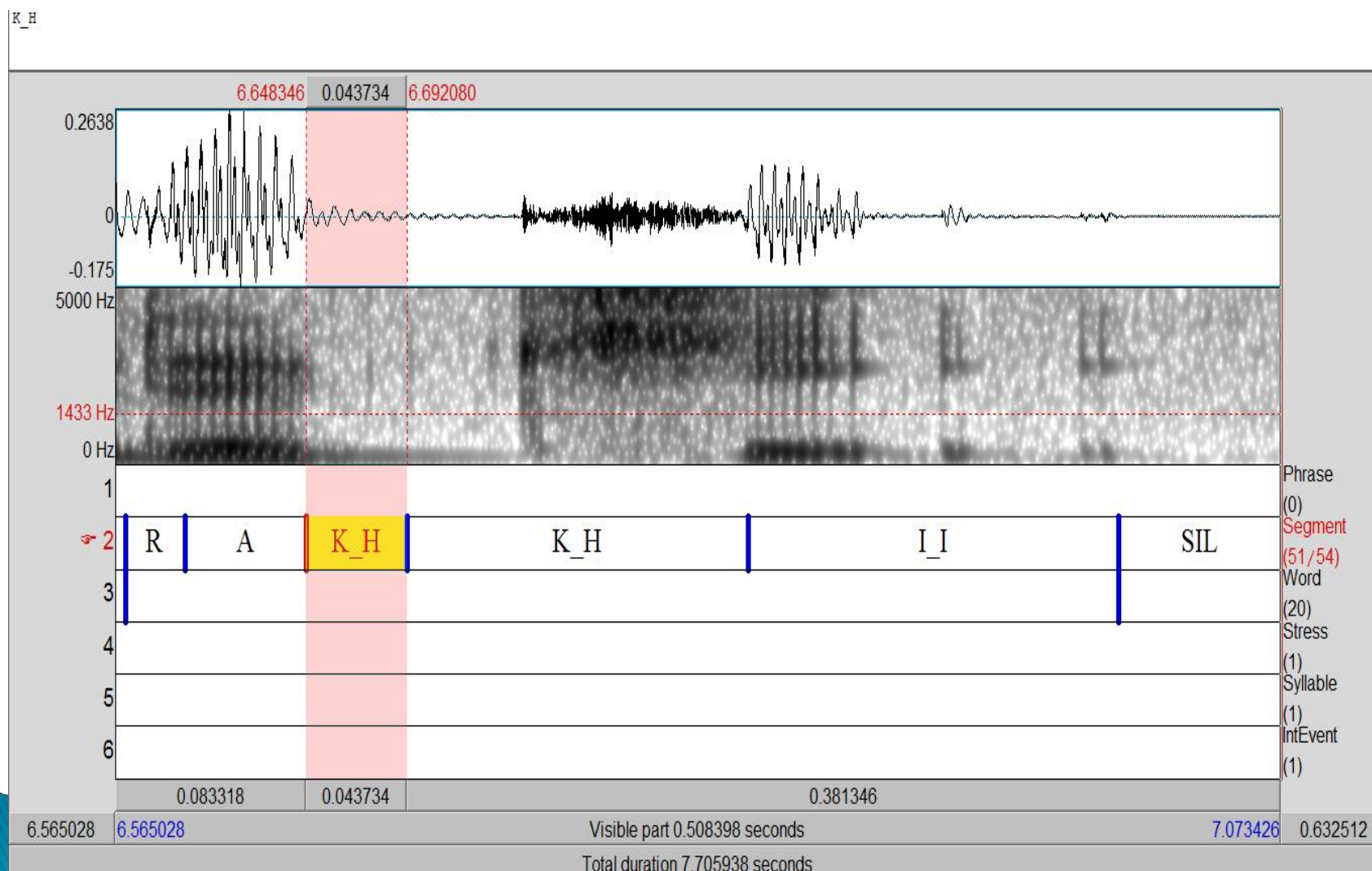
- ▶ In stops
- ▶ In fricatives
- ▶ In affricates etc.

# Gemination in laterals

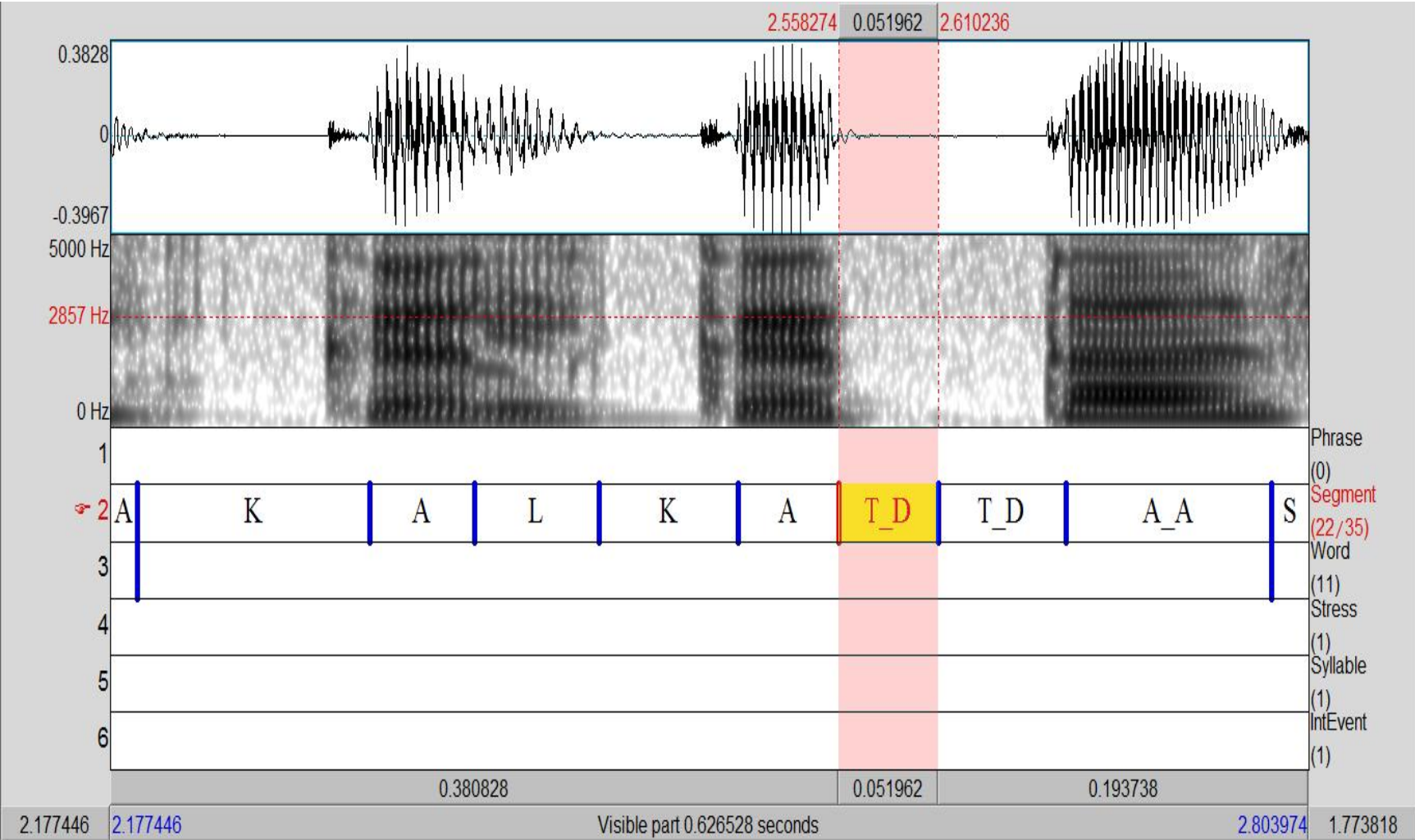




# Gemination in Aspirated Stop



# Gemination in Stops



## 4. Nasalization

### Contextual nasalization:

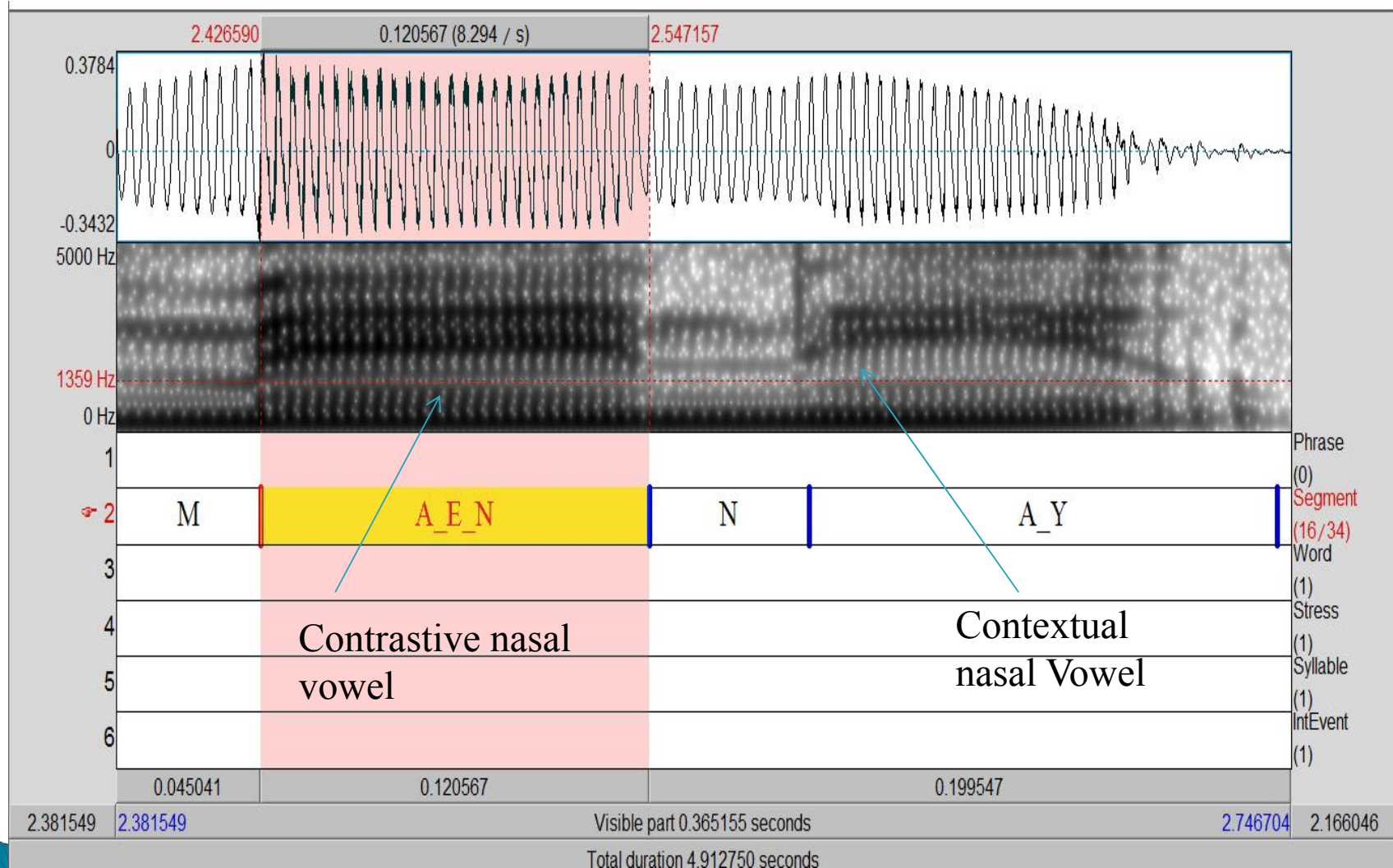
- انتظار
- کان

### Contrastive nasalization:

- کہا vs کہاں
- ہے vs ہیں

A vowel should be labeled as a nasal vowel only if it is contrastively nasalized, if a vowel is contextually nasalized, it should be labeled as an oral vowel.

# Contrastive vs Contextual Nasals

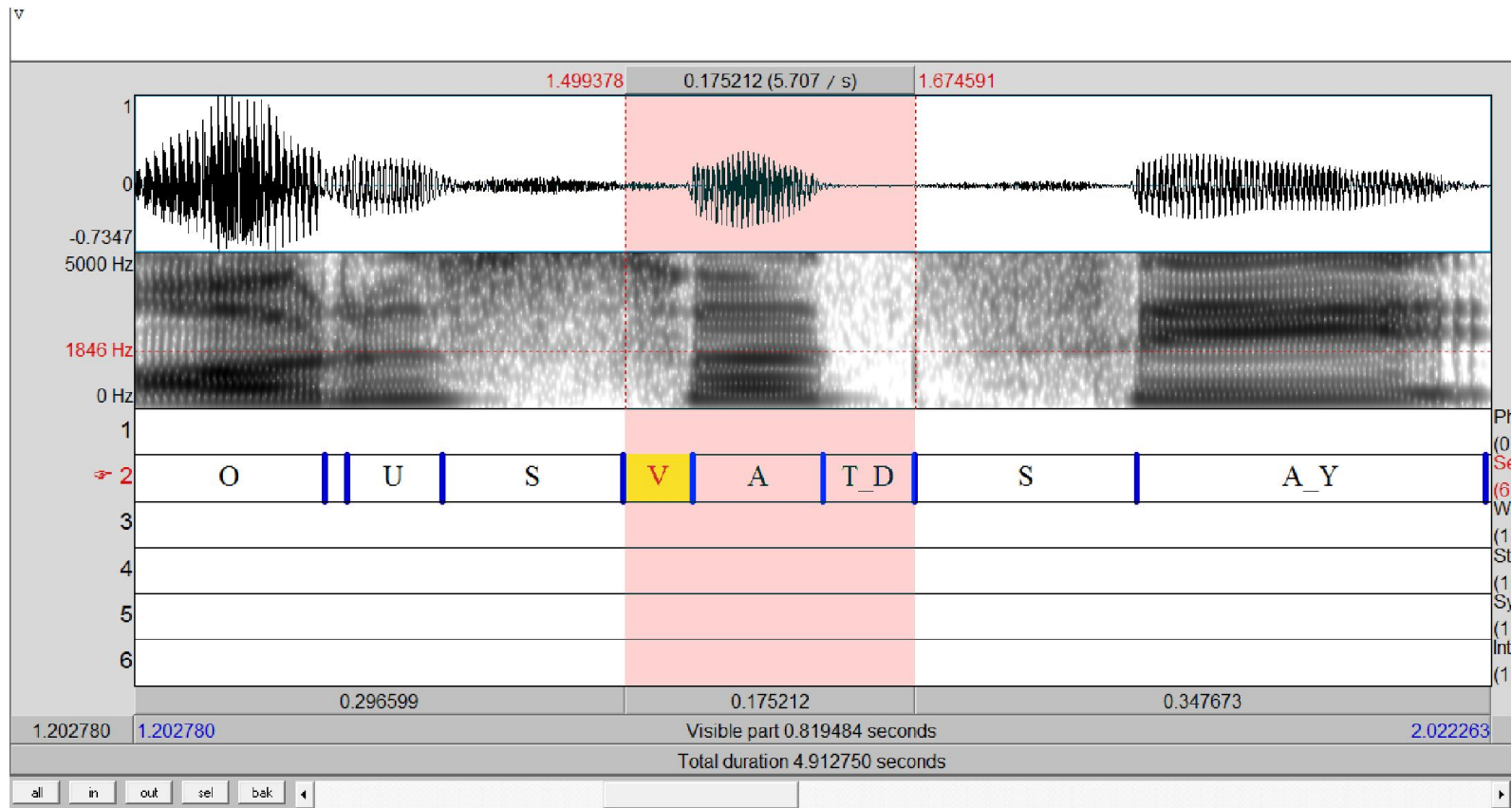


## 5. Deletion of Sound

- Each label should be carefully marked after analyzing the wave form and the spectrum of the sound. If a sound is not visible it should not be marked.
  - Deletion of /r/ in KAR
  - Deletion of /q/



# Sound deletion /Q/



## 6. /R/ Sound

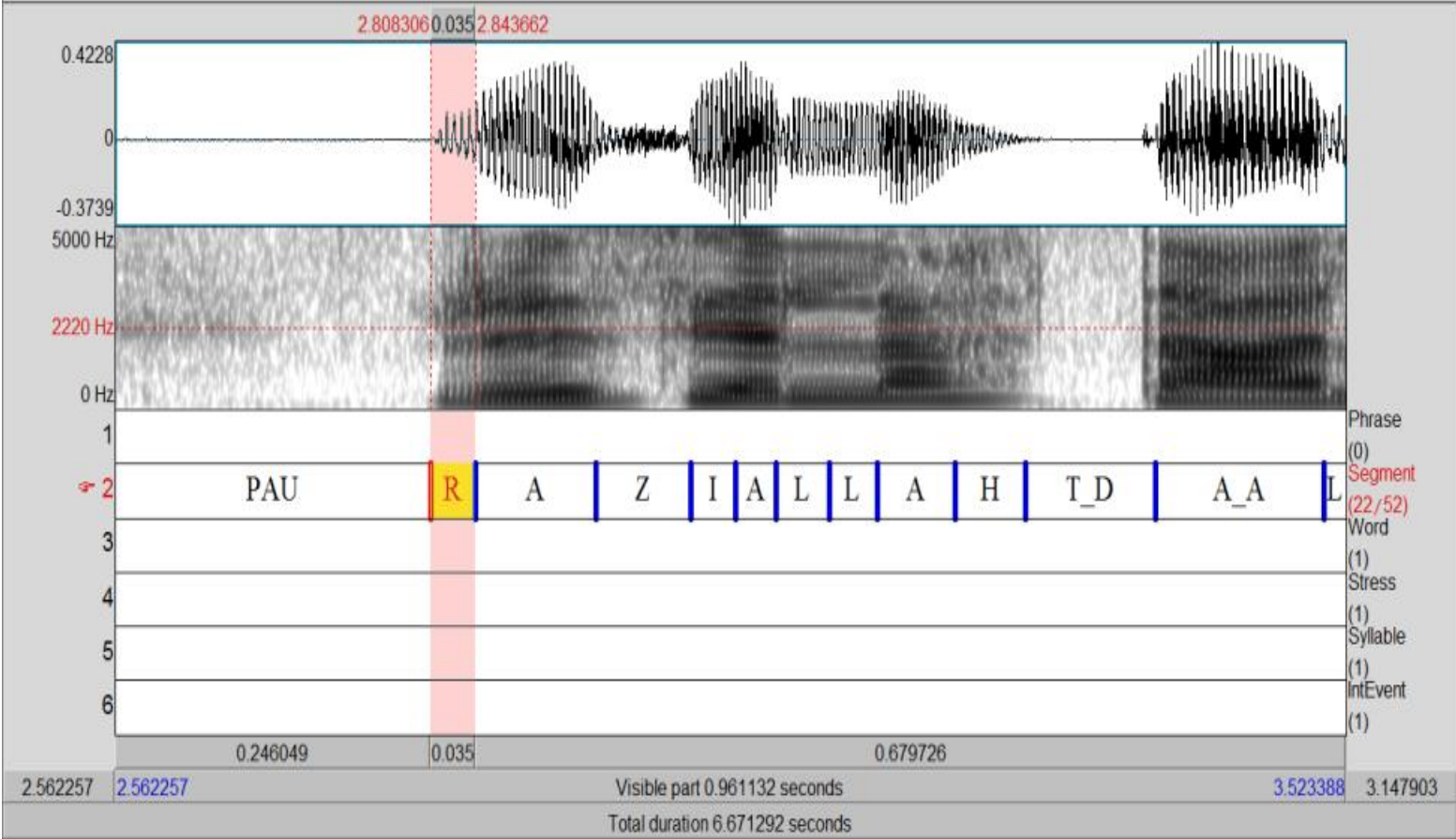
/r/ consonant occurs as **post alveolar, trill, flap** etc.

/r/ sound shows variable properties at different positions

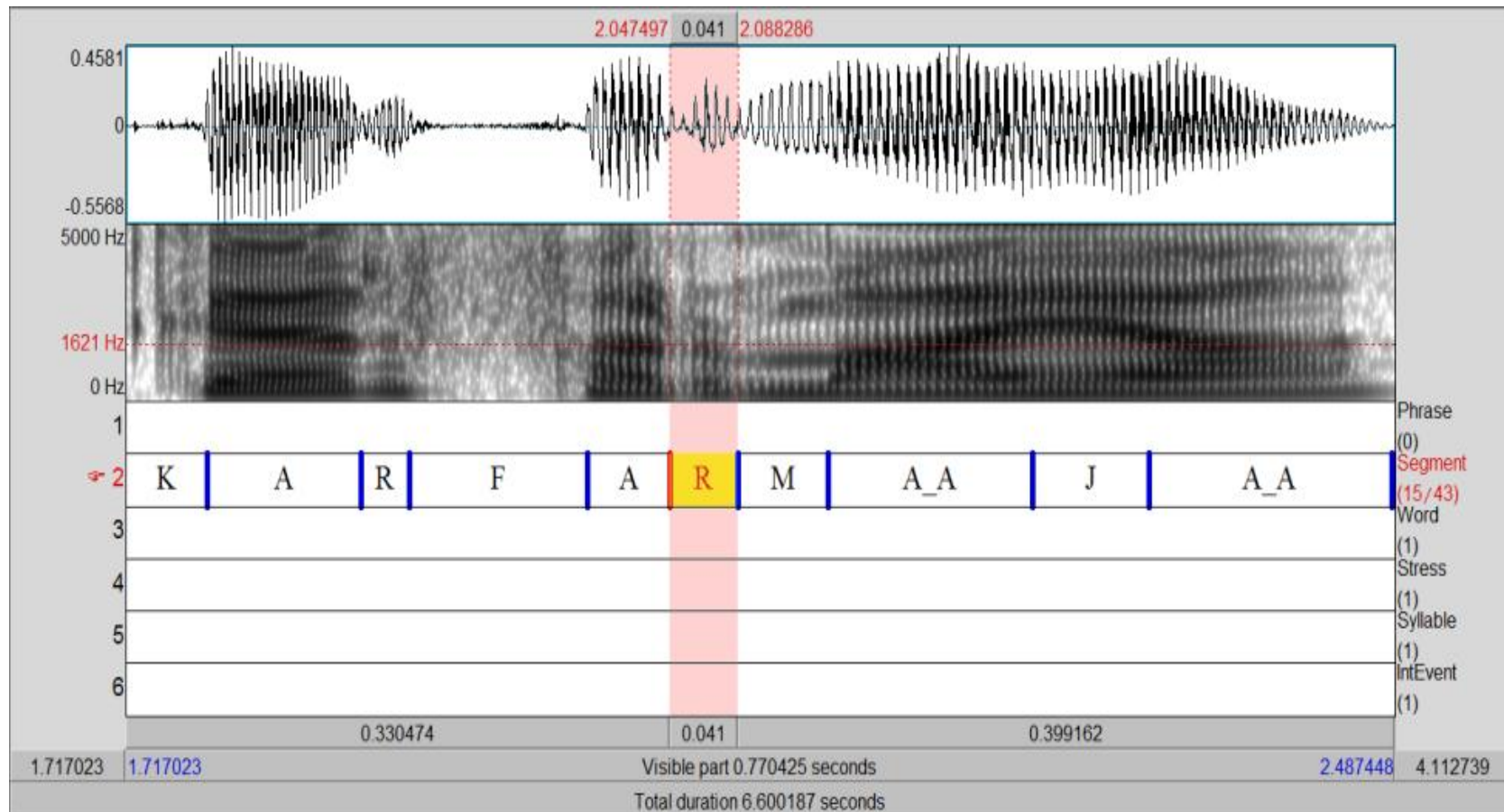
- ▶ at initial position
- ▶ at medial position
- ▶ at final position

But /r/ final is difficult to find out

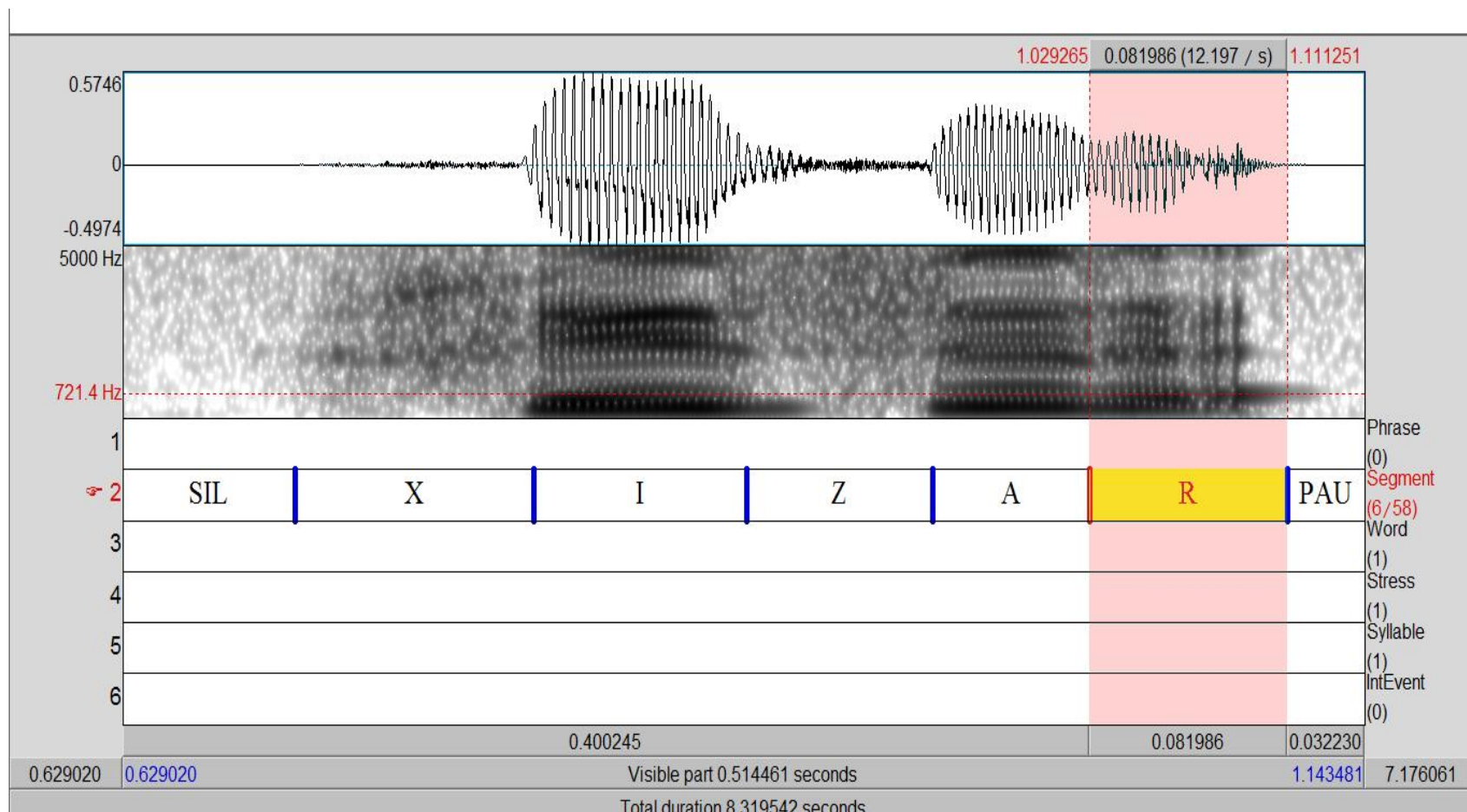
# /R/ at initial position



# /R/ at medial position



# /R/ at final position





## 7. /J/ sound

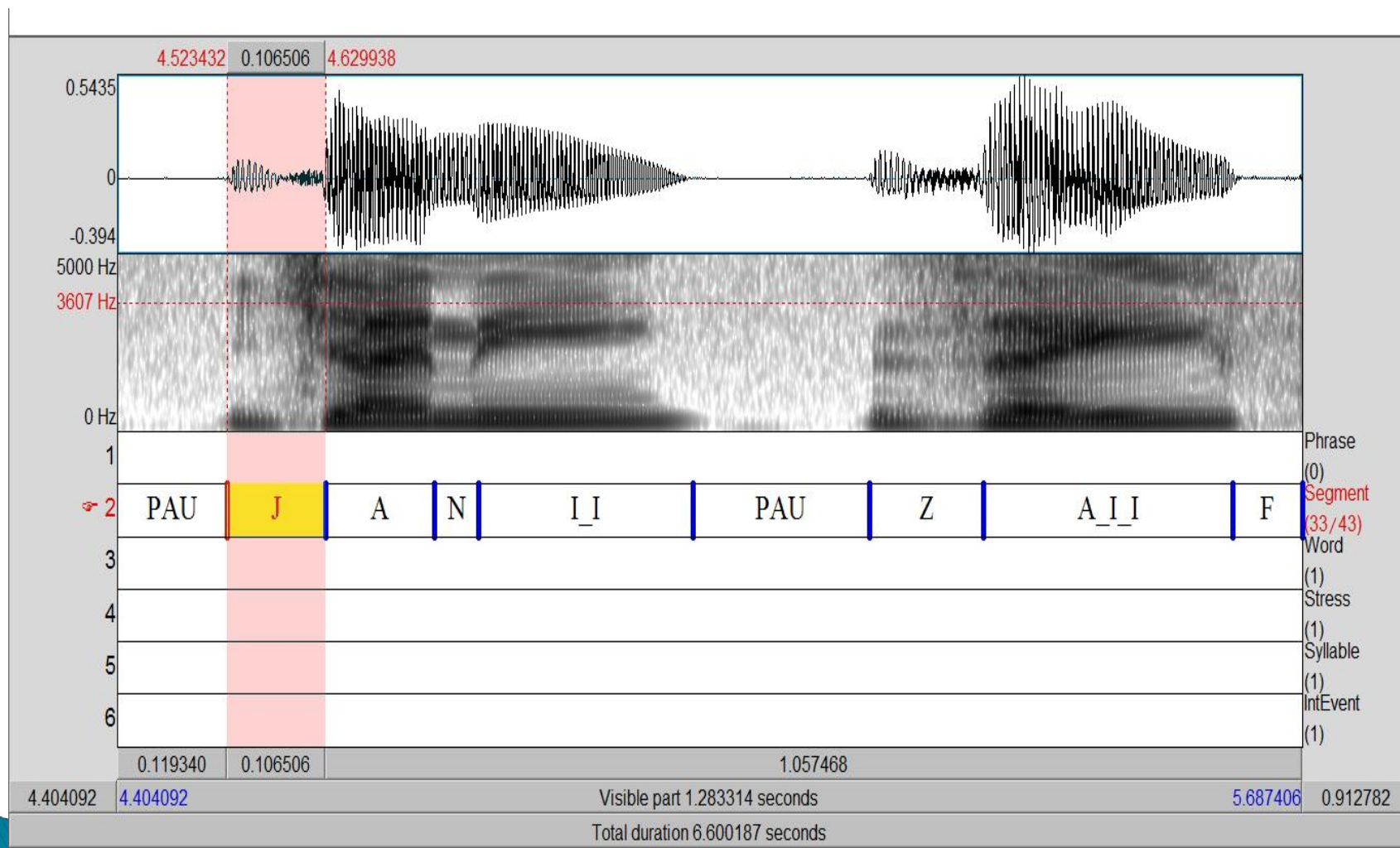
/J/ sound creates problem in identification at different positions.

- ▶ at initial position
- ▶ at medial position

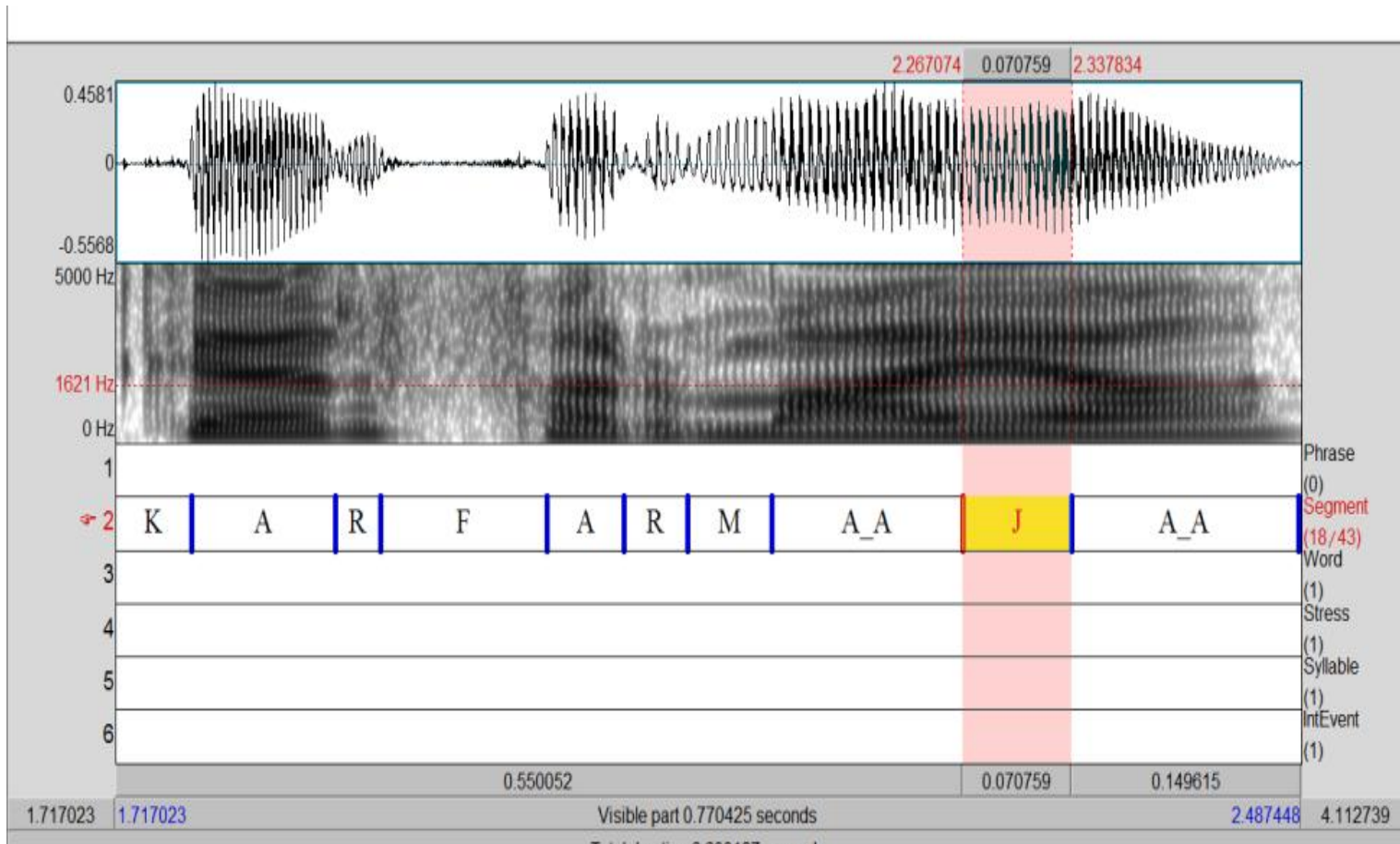
/J/ sound changes its sound when is uttered inter vocally; either it is deleted and replaced with a vowel sound by making a diphthong i.e.

- KA\_E\_H\_A\_A
- KI\_U\_U\_N

# /J/ at initial position



# /J/ medial



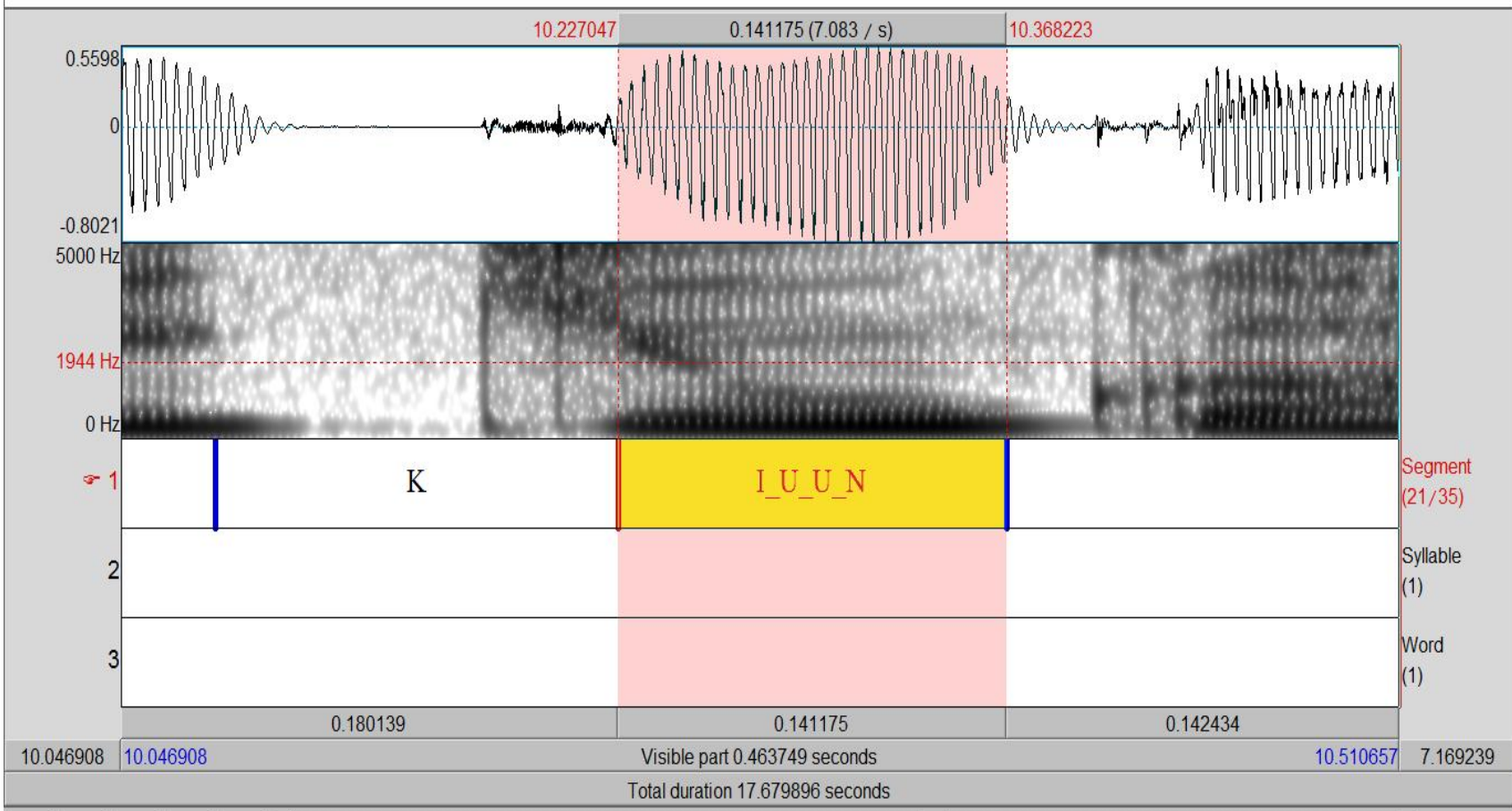
## 8. Diphthongs

“Diphthong is a single sound produced when two vowels are paired together in a specific sequence”

(Linda I. House)

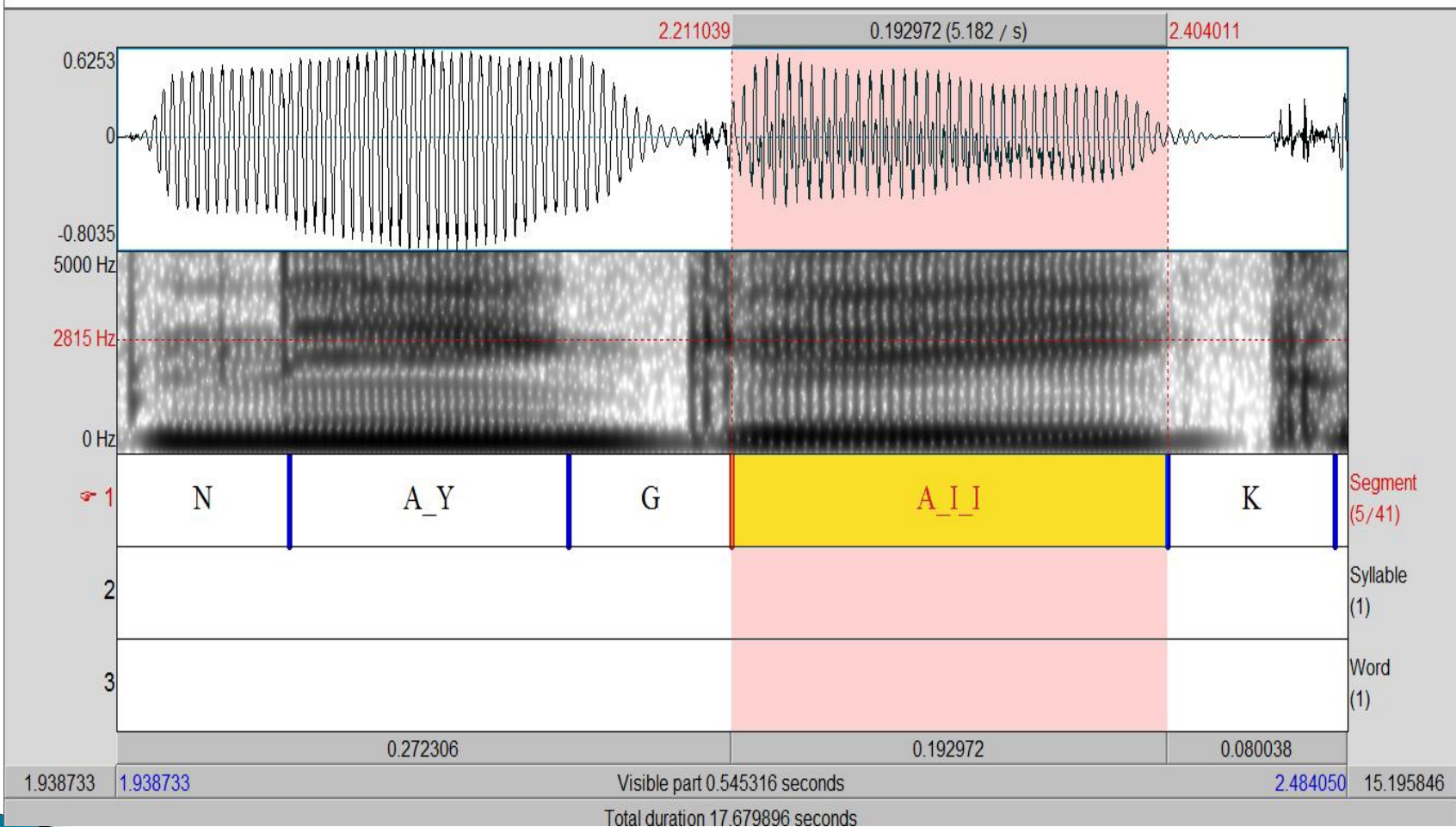
- ▶ On glide of a diphthong
- ▶ Transition phase
- ▶ Off glide
- ▶ vowels lose their pure quality
- ▶ labeled with single phonological symbol

# I\_U\_U\_N Diphthong





# A\_I\_I dipthong



# Suggested Diphthongs

7 diphthongs of Urdu are under discussion, formed by the deletion of /J/

Sr.#	Urdu Letter	IPA	CISAMPA
1	ی، ے، ِ	əi:	A_I_I
2	ے، ے، ِ	æe:	A_A_Y
3	آ	ɑ:e:	A_A_A_Y
4	آی	ɑ:i:	A_A_I_I
5	ی، ُوں	iū:	I_U_U_N
6	آ	æɑ:	A_E_H_A_A
7	و، ی	u:i:	U_U_I_I

# Questions