Linguistic Portfolios

Volume 3 Article 11

2014

An Acoustic Phonetic Portfolio of a Chinese-Accented English Idiolect

Borui Zhang
St. Cloud State University

AN ACOUSTIC PHONETIC PORTFOLIO OF A CHINESE-ACCENTED ENGLISH IDIOLECT

BORUI ZHANG

ABSTRACT

This acoustic portfolio contains four sections, including nine voice data analysis projects. The first section represents my pronunciation of English using the International Phonetic Alphabet (IPA). The second section describes the spectrogram data parsing of the vowels and consonants as I pronounce them. The third section focuses on acoustic correlates that I use to express lexical stress on homographic and multi-syllabic words. The fourth and final section investigates various phonological rules that apply in my pronunciation of the word <tempest >. Praat[1] and NORM[2] are the two acoustic computer software programs used in this study.

1.0 IPA transcription

The International Phonetic Alphabet (IPA) transcription model has been used for presenting the narrow allophonetic transcription. The transcript below shows my individual pronunciation pattern. I analyzed the data in the following section to verify that my impressions are correct. The existence of the / ḷ / (dark /l/) sound becomes apparent in the sonorant portion of the analysis. I hypothesized that I would produce the diphthongs [eɪ] and [əʊ], but in fact I only produce the diphthong [eɪ] and the monophthong [o] instead of the diphthong [əʊ].

[eɪ stóɹɪ əv ǽʃli æ̃nd̥ mi][3]
A Story of Ashlee and Me

[aɪ wʊd̥ láík tu ʃέɚ æ̃n ĩntɹəstĩŋ stóɹɪ əbaʊt maɪ kʌ́ɫtʃəɹəɫ ɪkspíəɹĩə̃ns əv ðə fɚ́st
I would like to share an interesting story about my cultural experience of the first
tʰãĩm aɪ kéɪm tu ðə jʊnáítɪd̥ stéɪts. wʌ̃n əv maɪ kɚ́ə̃nt ɹúmeɪts wəz ðə fɚ́st fɹέnd̥ aɪ
time I came to the United States. One of my current roommates was the first friend I
méɪd̥ ín ðɪs kʌ̃ntɹɪ. wi mέt ĩn aʊɚ dɔ́ɹmətɔɹɪ. wi wɚ lívĩŋ ʌ̃n ðə séɪm flɔ́ɹ bʌt
made in this country. We met in our dormitory. We were living on the same floor but
dífɹə̃nt ɹúmz. wɛ̃n aɪ dʒʌst əɹáívd aɪ fέɫt só tʰáíɚd̥, so aɪ dɪsáídɪd̥ tu tʰék ə ʃaʊɚ. maɪ ɔ́ɹəɫ
different rooms. When I just arrived I felt so tired, so I decided to take a shower. My oral

[1] Praat is scientific computer software for the analysis of speech in phonetics.
[2] http://ncslaap.lib.ncsu.edu/tools/norm/about_norm1.php NORM is a web-based software package to help linguists in normalizing vowel data.
[3] http://ipa.typeit.org/ is an online phonetic-symbol keyboard which provides various allophone symbols. It also supports copy and paste.

ĩŋglɪʃ wəz nɑt véɹɪ gʊd̥, ǣnd̥ aɪ wəz ʃáí tu spík tu pʰípl̥ æt ðæt tʰáĩm tʰú. ʃi stáɹtɪd̥
English was not very good and I was shy to speak to people at that time too. She started
tu tʰók tu mi wɛ̃n wí kéɪm əkɹós ĩn ðə bǽθɹum. "dɪd jʊ dʒʌst mʊv ĩn tədéɪ?" ʃi
to talk to me *when we came across in the bathroom. "Did you just move in today?" She*
ǽskd̥ mi." jɛ́s" aɪ sǽd̥. "maɪ néɪm ɪz ǽʃli. Náís tu sí jə! wɛ́ɚ ɑɹ jʊ fɹɒm?"
asked me. "Yes" I said. "My name is Ashlee. Nice to see you! Where are you from?"
"ãĩm fɹɒm tʃãĩnə". ʃi lɛ̃nt mi hɚ táʊəɫ fɔɹ bǽθĩŋ, wɛ̃n ʃi nótɪsd̥ ðæt aɪ dɪdnt hǽv
"I'm from China". She lent me her towel for bathing, when she noticed that I didn't have
ə táʊəɫ. wi bɪkéɪm gʊd̥ fɹéndz ǽftɚ ðæt. ʃi hɛ́ɫpd̥ mi wɪð maɪ óɹəl̥ ĩŋglɪʃ, wáíɫ aɪ
a towel. We became good friends after that. She helped me with my oral English, while I
tʰóɫd̥ hɚ sʌ̃mθĩŋ əbaʊt tʃãĩnə ǣnd̥ tʃaɪnís. wʌ̃n dé, wi ǽdɪd̥ itʃ ʌ́ðɚ ʌ̃n
told her something about China and Chinese. One day, we added each other on
fésbʊk. wɛ̃n aɪ só: hɚ pɹófaɪl̥ pʰɪktʃɚ, aɪ méɪd̥ e kɑ̃mə̃nt ʌ̃n ɪt, wɪtʃ aɪ θót ðæt
Facebook. When I saw her profile picture, I made a comment on it, which I thought that
wəz e kɑ̃mplɪmə̃nt. aɪ sǽd̥ jʊ lɑ́st weɪt naʊ kʌ̃mpʰɛ́ɚĩŋ wɪð ði óld̥ pʰíktʃɚ. "ɹílɪ?"
was a compliment. I said you lost weight now comparing with the old picture. "Really?"
ʃi wəz sɚpɹáízd̥. ʃi sǽd̥ "aɪ tʰúk ðɪs pʰíktʃɚ lǽst wíkə̃nd̥". aɪ θĩŋk maɪ swǽtɚ lʊks só
she was surprised. She said "I took this picture last weekend. I think my sweater looks so
pʰʌfɪ ʌ̃n ðɪs. aɪ ɛm góĩŋ tu tʃǣndʒ tu ənʌ́ðɚ wʌ̃n." aɪ félt ɛ̃mbǽɹəsd̥ ǣnd̥ ɹɪəláízd̥ ðæt
puffy on this. I am going to change to another one." I felt embarrassed and realized that
maɪ kɑ̃mplɪmə̃nt wəz nɑt véɹɪ əpɹópɹɪeɪtlɪ. ɪf jʊ: tʰél sʌm fímeɪl ĩn tʃãĩnə, ívə̃n ɪf ʃi
my compliment was not very appropriately. If you tell some female in China, even if she
íz nɑt juɚ fɹɛ́nd̥, ʃi wɪl fíl só hǽpɪ tu híɚ ðæt. ðɪs íz haʊ wi pɹéíz ĩn aʊɚ
is not your friend, she will feel so happy to hear that. This is how we praise in our
kʌ́ltʃɚ. haʊɛ́vɚ, ɪt dʌz nɑt wɚ́k aʊt ĩn əmɛ́ɹɪkə̃n kʌ́ltʃɚ. ǽʃli dɪdnt tʰék maɪ wɚdz
culture. However, it does not work out in American culture. Ashlee didn't take my words
ɹɑ̃ŋ. ʃi ólso ʌ̃ndɚstúd haʊ dífɹə̃nt kʌ́ltʃɚz əpláí ĩn lǣŋgwɪdʒɪz. tədéɪ, ə jíɚ
wrong. She also understood how different cultures apply in languages. Today, a year
létɚ, aɪ dɔnt θĩŋk aɪ wɪl kɑ̃mplɪmə̃nt sʌ̃mwʌ̃n wɪð kɑ̃mɛnts ʌ̃n ðɛɚ wéɪt sĩns aɪ hæv
later, I don't think I will compliment someone with comments on their weight since I have

ɔlɹǽdɪ əb’sɔ́ɹbd̥ əmɛ́ɹɪkə̃n kʌ́ltʃɚ fɔɹ ə jíɚ́.   ĩn ədíʃə̃n, aɪ kæ̃nt ĩmǽdʒĩn haʊ mʌ́tʃ aɪ hæv
already absorbed American culture for a year. In addition, I can’t imagine how much I have
lɚ́nd̥ əbaʊt   tʃaɪnis kʌ́ltʃɚ ǽftɚ aɪ kéɪm hɪə.   ðə θĩŋz aɪ jʊs tu dú ĩn tʃãĩnə kə̃ntʰénz
learned about Chinese culture after I came here. The things I use to do in China contains
ə lɑ́t əv pɑɹtʰíkjʊlɚ kʌ́ltʃəɹəl ĩnfɔɹméɪʃə̃n, bʌt aɪ kʊdnt ɹɪəláíz ðoz   bɪkɔ́z ðɛɚ wɚ
a lot of particular cultural information, but I couldn’t realize those because there were
nó ənʌ́ðɚ kãĩnd əv kʌ́ltʃɚ aɪ kʊd̥ kə̃mpʰɛ́ɚ wɪð.   kʌ́ltʃɚ æs ə pʰaɹt əv læ̃ŋgwɪdʒ, ɛ́vɹɪ
no another kind of culture I could compare with. Culture as a part of language, every
læ̃ŋgwɪdʒ lɚ́nɚ   ʃʊd̥ lɚ́n   ə njú kʌ́ltʃɚ   wɛ̃n ðeɪ staɹt tu   lɚ́n ə njú læ̃ŋgʊɪdʒ,
language leaner should learn a new culture when they start to learn a new language,
wɪtʃ   íz ɔ́lso fʌ́n tu nó]
which is also fun to know.

2.0 The Acoustic Vowel Space

This portion of the portfolio focuses on eleven vowels and three diphthongs. It is based on the words from the original study by Peterson and Barney in 1952, saying the following words: <heed>, <hid>, <hayed>, <head>, <had>, <hod>, <hawed>, <hoed>, <hood>, <who’d>, <hud>, and <heard> (Koffi, 2013:153). The word <heard> is not included in the current section because it is not a phoneme but an allophone. The word <hag> is added to attest to the existence of vowel rising in my accent. I recorded myself saying each word three times in a row. The following data validates my prediction that the F1 of [æ] has been raised 35Hz from <had> to <hag>. Notice that the higher F1 of a sound, the lower the sound is actually articulated. All measurements include pitch (F0), F1, F2, and duration. These values analyzed from my data will be compared with those from General American English (GAE) in a data table form. After importing the data form into NORM, the software will create an “Individual Vowel Format Values non-normalized” graph, which shows how different my vowels are from GAE. It will demonstrate which vowel sounds in my accent may cause intelligibility problems by displaying them in a chart showing the visible distance between my English vowels and the GAE.

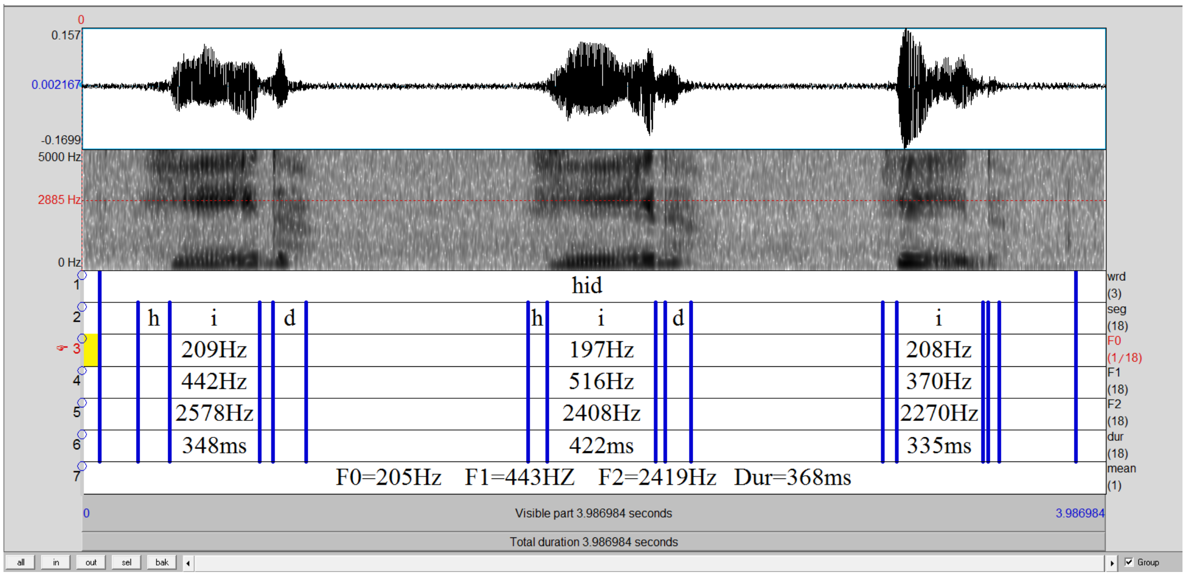

Figure 1: <heed>

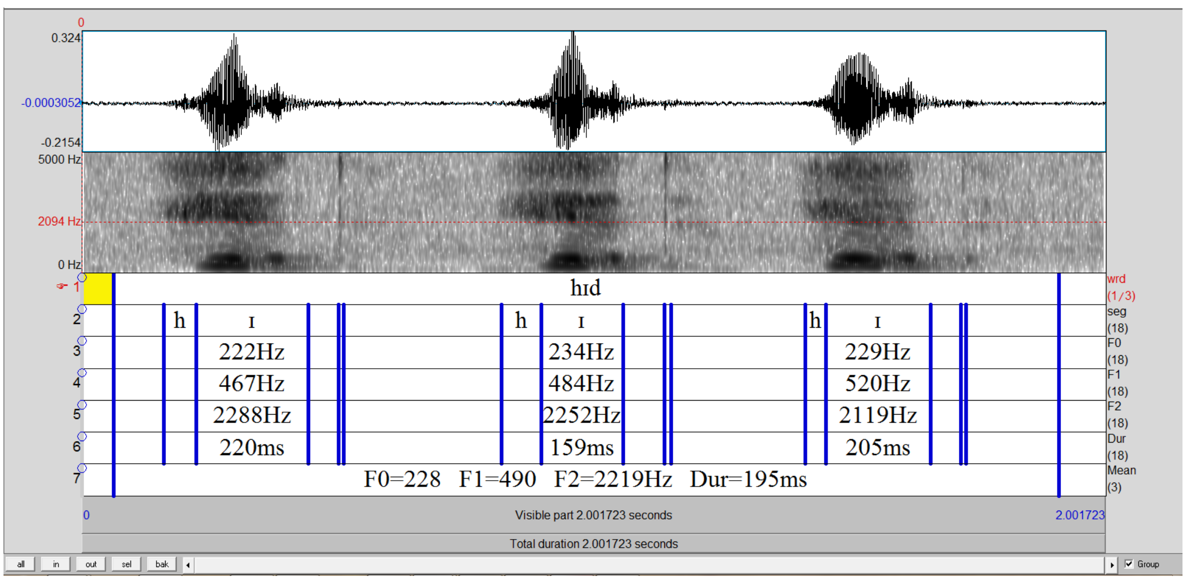

Figure 2: <hid>

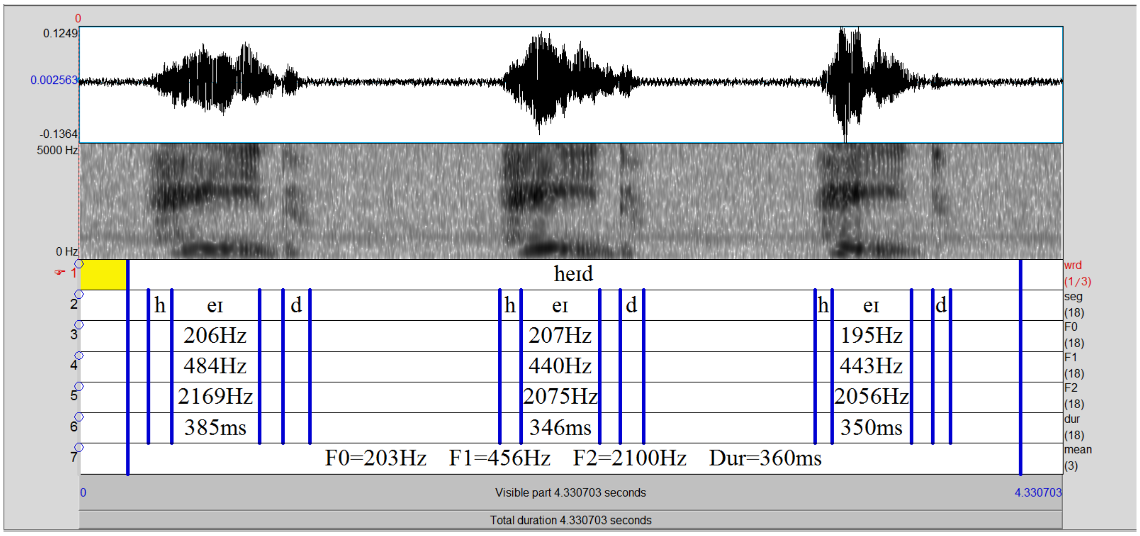

Figure 3: <hayed>

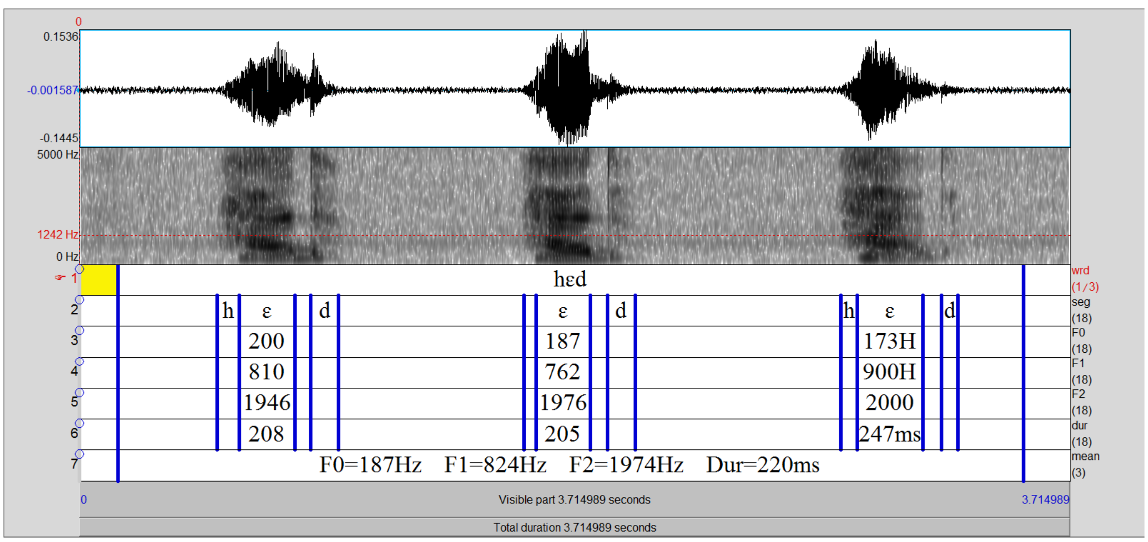

Figure 4: <head>

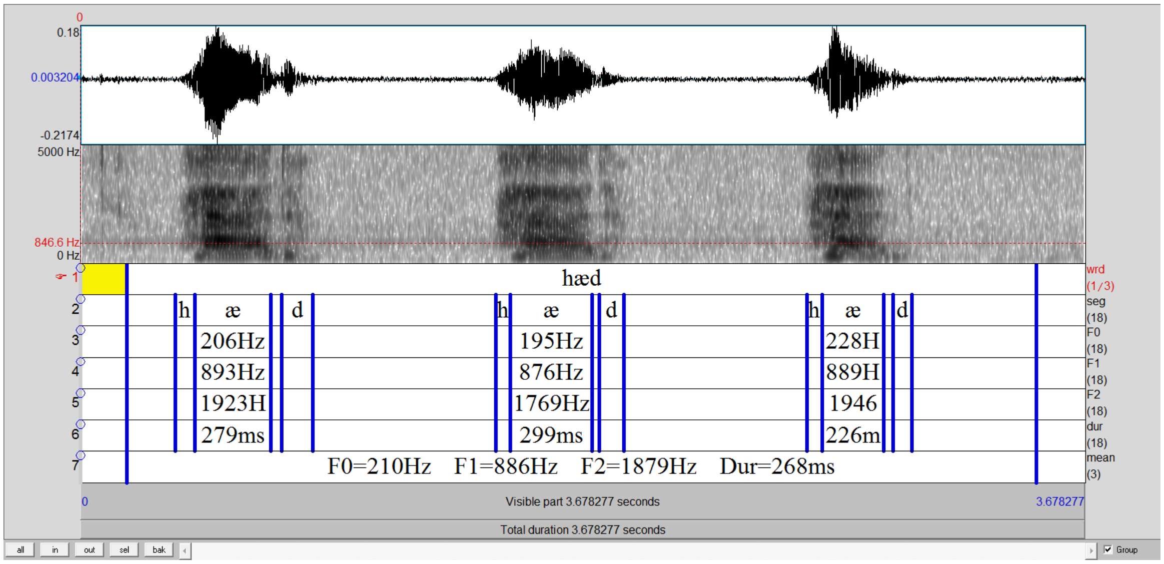

Figure 5: <had>

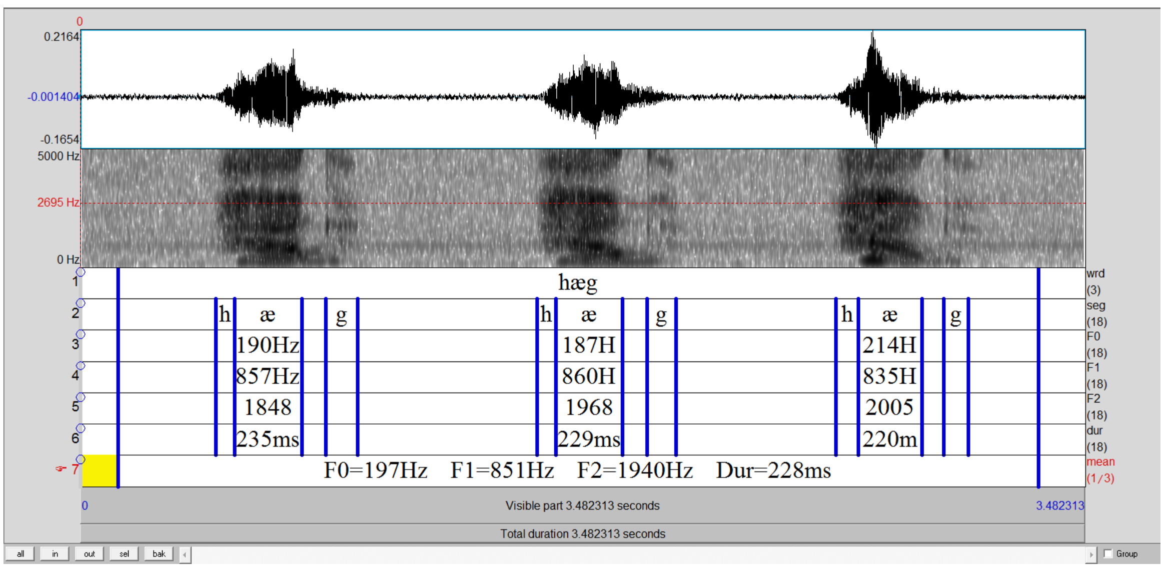

Figure 6: <hag>

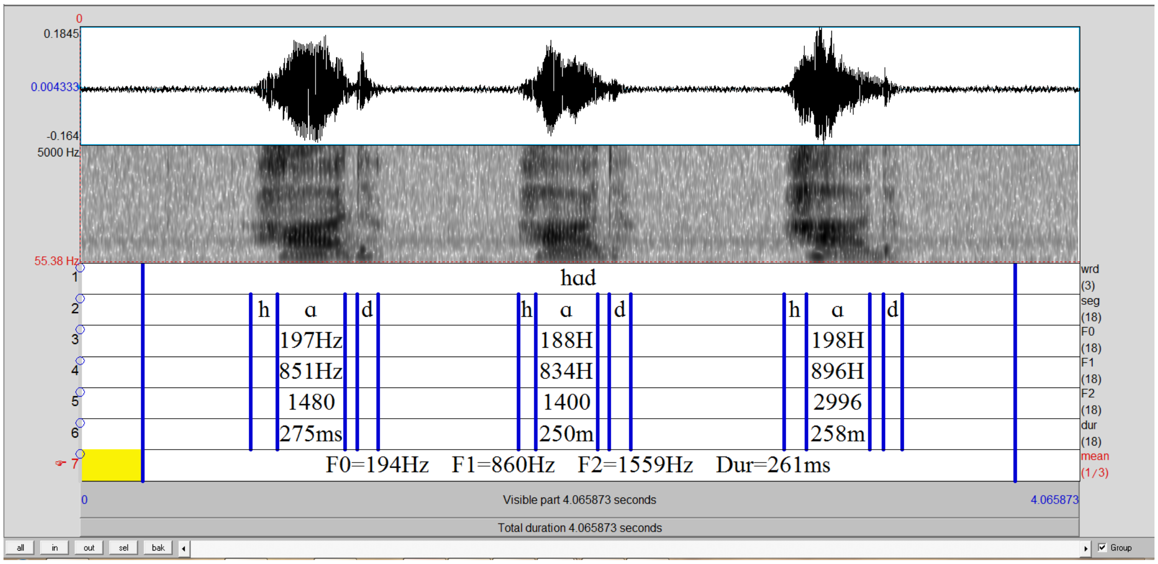

Figure 7: <hod>

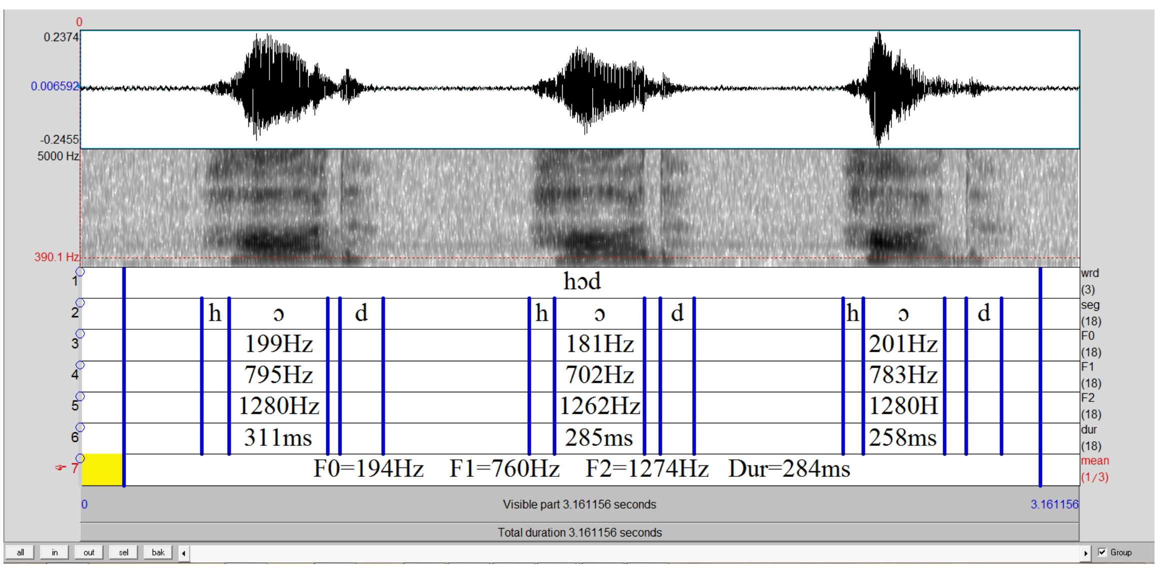

Figure 8: <hawed>

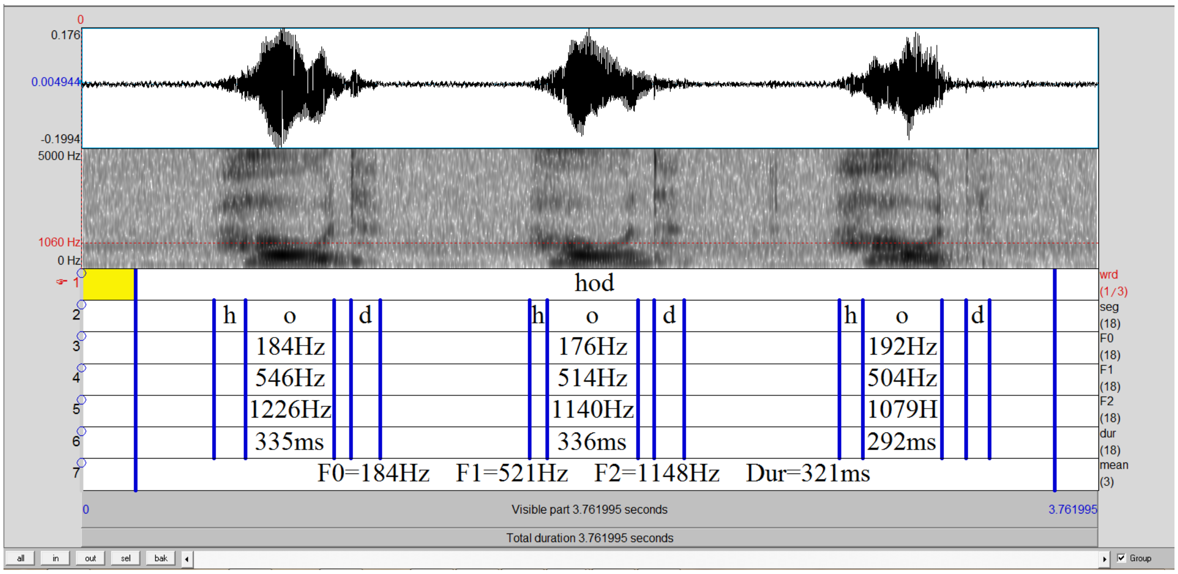

Figure 9: <hoed>

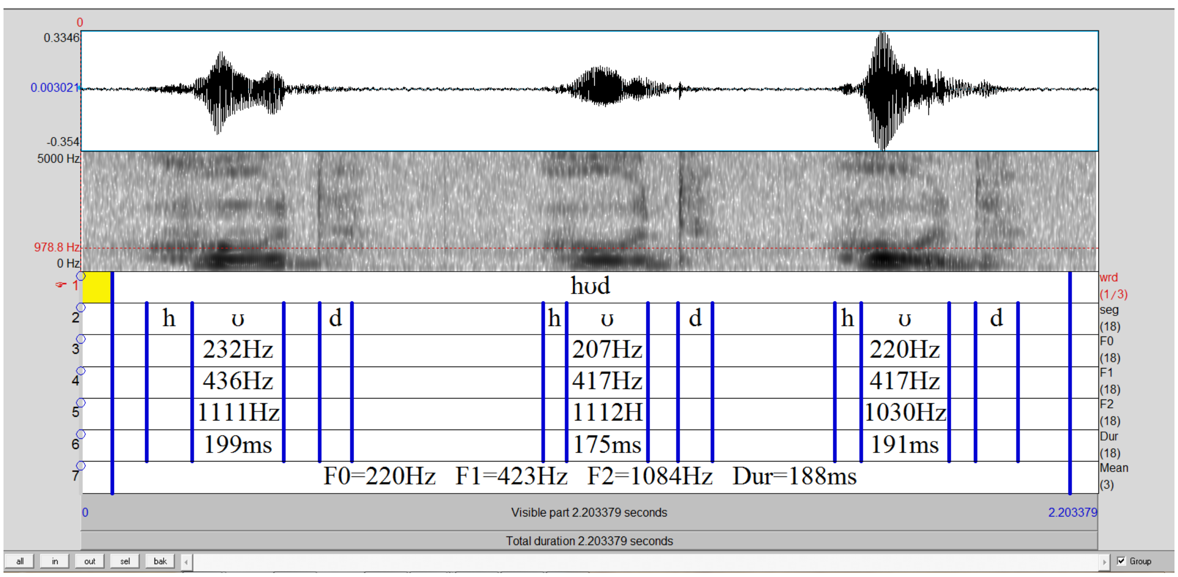

Figure 10: <hood>

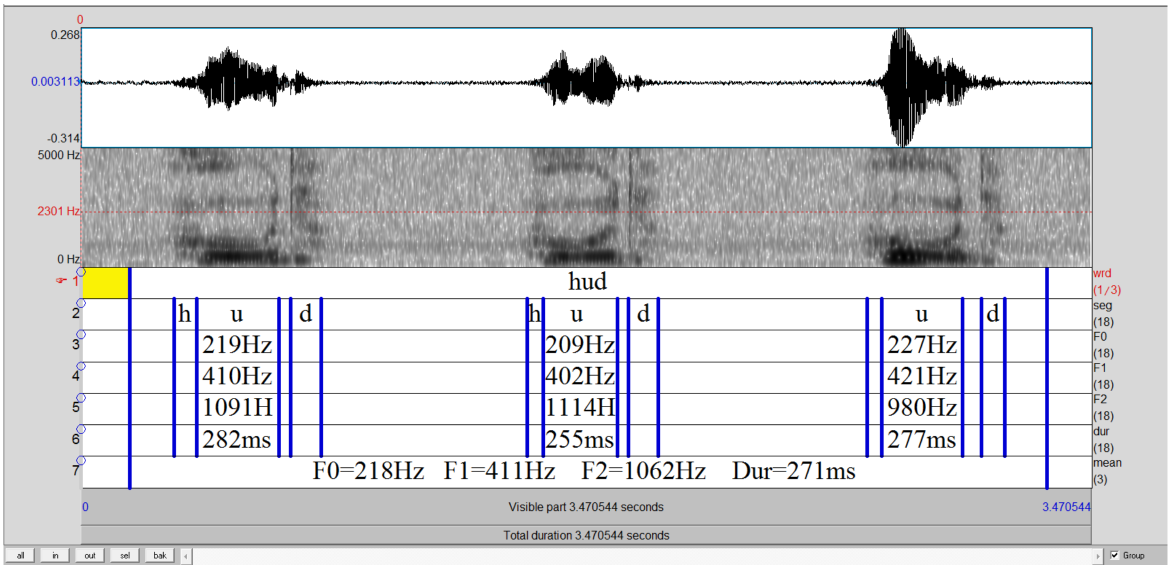

Figure 11: <who'd>

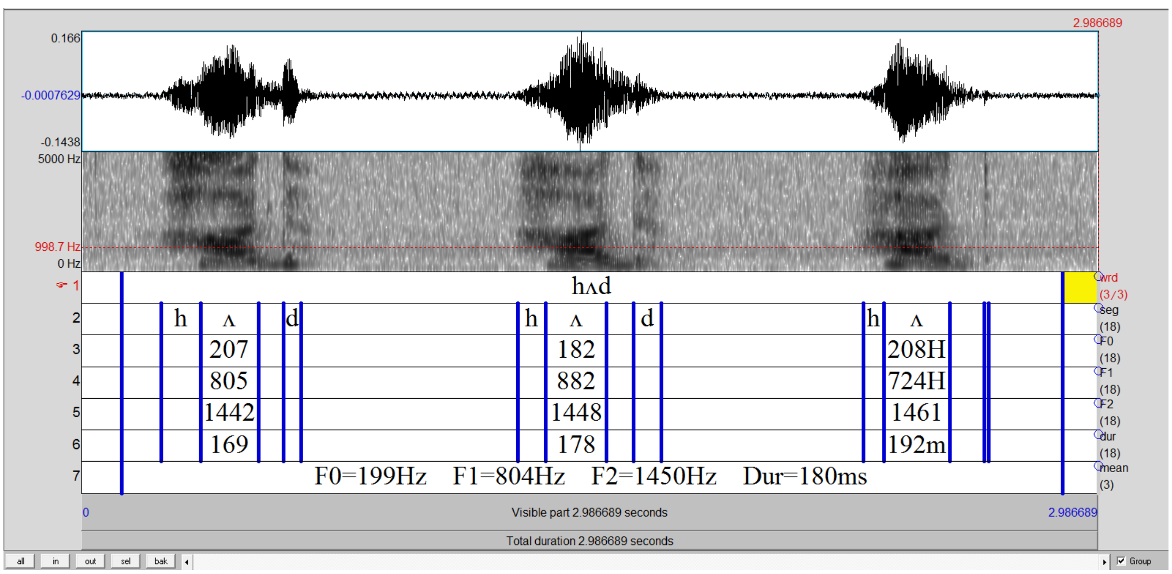

Figure 12: <hud>

Table 1 displays the mean correlates of F0 (Pitch), F1, F2, and duration for my idiolect and compares them with those of GAE:

Words		heed	hid	hayed	head	had	hag	hod	hawed	hoed	hood	who'd	hud
Vowels		[i]	[ɪ]	[e]	[ɛ]	[æ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
GAE	**F0**	**235**	**232**	**219**	**223**	**210**	**210**	**212**	**216**	**217**	**232**	**231**	**221**
Borui	F0	205	228	203	187	210	197	194	194	184	220	218	199
GAE	**F1**	**310**	**430**	**536**	**610**	**860**	**860**	**850**	**590**	**555**	**470**	**370**	**760**
Borui	F1	443	490	456	824	886	851	860	760	521	423	411	804
GAE	**F2**	**2790**	**2480**	**2530**	**2330**	**2050**	**2050**	**1220**	**920**	**1035**	**1160**	**950**	**1400**
Borui	F2	2419	2219	2100	1974	1879	1940	1559	1274	1148	1084	1062	1450
GAE	**DUR**	**306**	**237**	**320**	**254**	**332**	**332**	**323**	**353**	**326**	**249**	**303**	**226**
Borui	DUR	368	195	360	220	268	228	261	284	321	188	271	180

Table 1: Borui's Vowels vs. GAE Vowels

Vowels that I produce and those in GAE are plotted together in the same acoustic vowel space, as presented below in Figure 13:

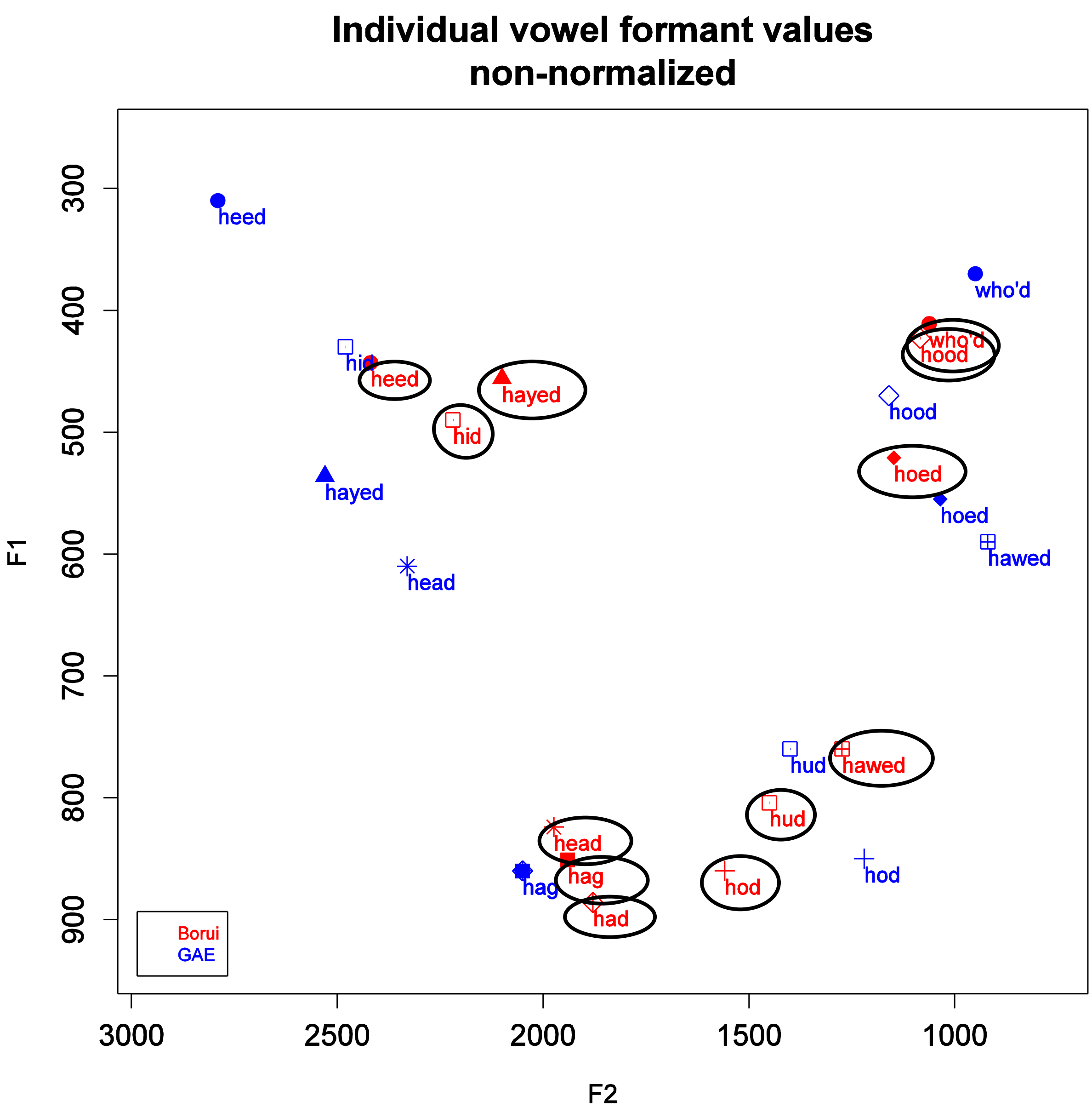

Figure 13: Acoustic Vowel Space

Vowel harmony occurs in my accent. The vowels are clustered together into several main groups on the acoustic vowel space map. The vowels [i], [ɪ], and [e] are gathered in an upper-central-and-mid area. The vowels [ɛ] and [æ] slightly overlap each other at the central-and-low area. The vowels [u] and [ʊ] mostly overlap each other at the high-and-back area on the map. We can also see from the map that my [i] and GAE's [ɪ] are very close, as well as my [ɛ] and GAE's [æ], and my [ɔ] and GAE's [ʌ].

2.1 Vowel Intelligibility Assessment

An acoustic difference of 60Hz marks the intelligibility boundary between two contiguous vowels, according to Koffi (2013:9-10). If the acoustic difference between two contiguous vowels is 60Hz or greater, the intelligibility between these two vowels is considered intact, no audible confusion will result; when the distance of two contiguous vowels ranges between 21Hz and 59Hz, people usually have slight difficulties recognizing the vowels. Intelligibility is considered absolutely absent and confusion will result if the acoustic range is 20Hz or less. F1 contains more than 80% of the acoustic energy of a vowel according to Ladefoged, as mentioned in Koffi's *Linguistic Portfolio* (2013:155). See Table 2 below for a catalog of my vowel issues through the differential F1 measurement of two adjacent vowels, taken between my idiolect and GAE:

No.	Adjacent Pairs		F1 (Hz)		Difference	Intelligibility
	Borui's	GAE	Borui's	GAE		
1	[i]	[ɪ]	443	430	13Hz	Absolute confusion
2	[ɛ]	[æ]	824	860	36Hz	Slight confusion
3	[ɔ]	[ʌ]	760	760	0Hz	Absolute confusion
4	[ʌ]	[a]	804	850	46Hz	Slight confusion
5	[u]	[ʊ]	411	470	59Hz	Slight confusion
6	[ʊ]	[u]	423	370	53Hz	Slight confusion
7	[o]	[ʊ]	521	470	51Hz	Slight confusion

Table 2: Confusion Table

On the basis of the F1 comparison, [ɔ] and [ʌ] have a 0Hz difference. This indicates that GAE listeners may confuse the words I produce that contain [ɔ] and [ʌ]. For example, they may recognize <cut> when I actually pronounce <caught>. The front vowel [i] in my accent has only a 13Hz difference with the vowel [ɪ] in GAE. This could cause GAE listeners to hear the word <ship> when I actually produce the word <sheep>. The vowel [ɛ] in my accent also may create a slight confusion with [æ] in GAE, which could be problematic for GAE listeners. They may recognize my pronunciation of the word <beg> as <bag>. The same happens in the back vowels [u] and [ʊ]. GAE listeners may hear <route> when I produce the word <rude>, and vice versa, because I pronounce these two sounds too close to each other. The back vowels [u] and [ʊ] in GAE have a 100Hz (470Hz−370Hz) difference, while those in my idiolect have only a 12Hz (423Hz−411Hz) difference.

2.2 Diphthongs

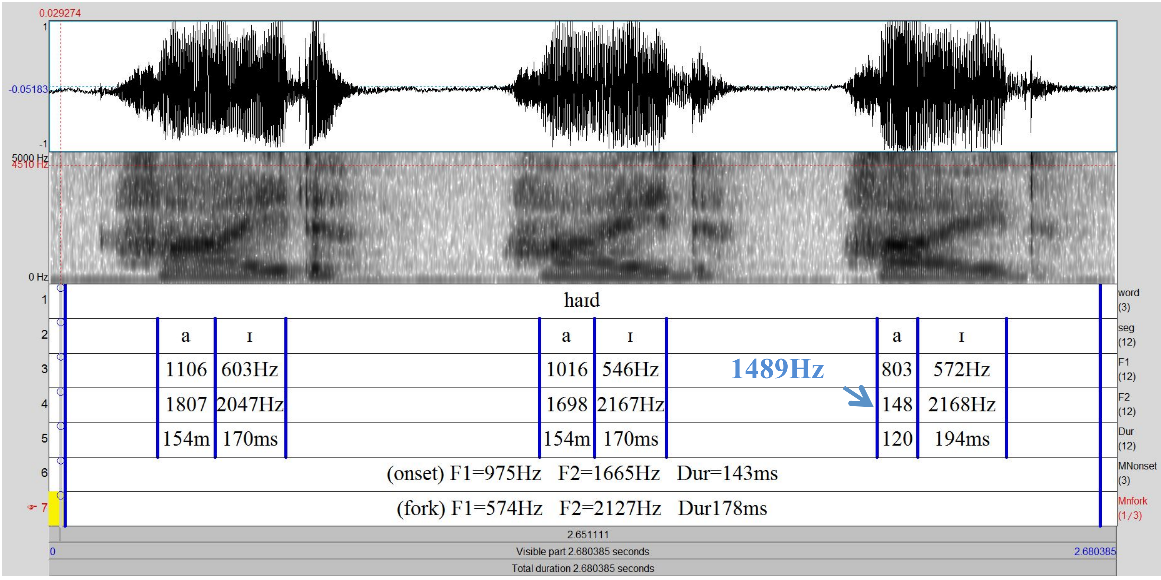

Figure 14: <hide>

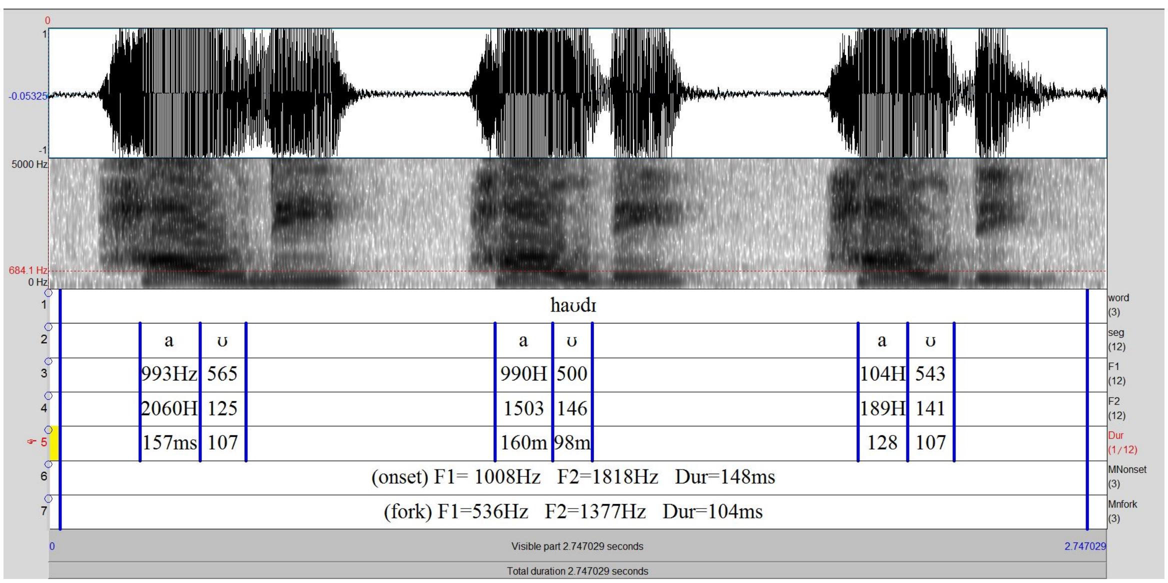

Figure 15: <howdy>

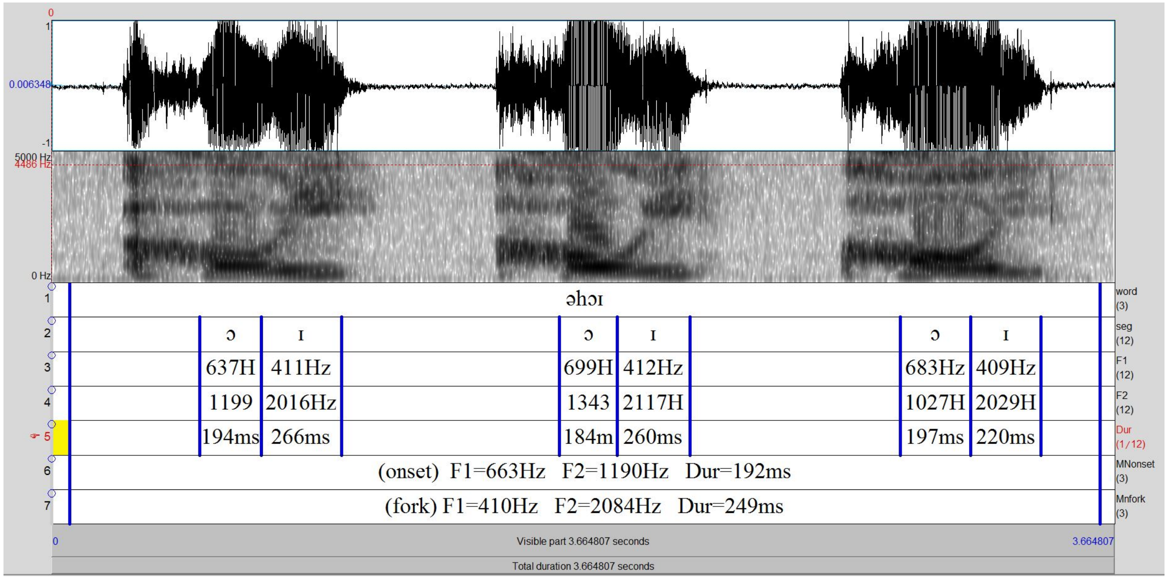

Figure 16: <ahoy>

word	hide	howdy	ahoy
F1 (onset)	975	1008	663
F2 (onset)	1665	1818	1190
F1 (Fort)	574	536	410
F2 (fort)	2127	1377	2084
Dur onset	143	148	192
Dur fork	178	104	249

Table 3: Diphthong Measurement

The data in Table 3 verifies that [aɪ], [aʊ], and [ɔɪ] exist as diphthongs in my accent, representing the clear forks shown in the spectrogram. As mentioned in the beginning of the first section, [eɪ] is also another diphthong in my idiolect. The use of this diphthong is probably related to my early English learning experience in elementary school in China, where we were taught British English. The diphthong [eɪ] is one of the diphthongs present in British English.

3.0 Consonants

Because of different manners of articulation, all English consonants show their own distinct patterns on spectrograms. In this section, consonant analysis is covered, including voiced and voiceless consonants, frication, and sonorants. A voiced consonant appears with a dark strain on the bottom of a spectrogram, while a voiceless one appears with a blank space. Stop consonants have a clear vertical line to show the burst. Fricatives appear with a strain from top to bottom, from dark to light. Affricates combine both features of stop and fricative consonants. Most of the consonants do not have the same acoustic energy as a vowel has, which means vocal pulsating doesn't show on spectrograms. However, some consonants have this characteristic and are called sonorants.

3.1 VOT

VOT stands for Voiced Onset Time, which describes the time between the sound releasing and voicing. The measurement of VOT is widely applied in stop consonant analyses. In the spectrogram, the VOT can be observed from the start of the clear breaking strain of a wave to the beginning of the following periodical wave. The value of VOT can be either positive, negative, or zero.

There are six words that I am using in this portion of the project: <pot>, <brought>, <tot>, <dot>, <cot>, and <got>.

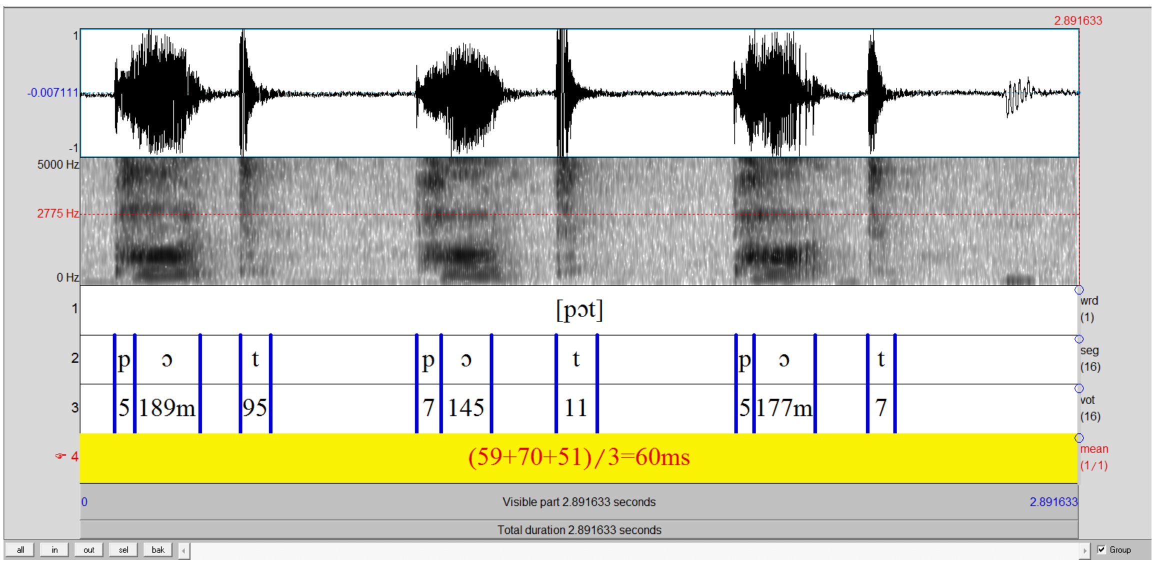

Figure 17: <pot>

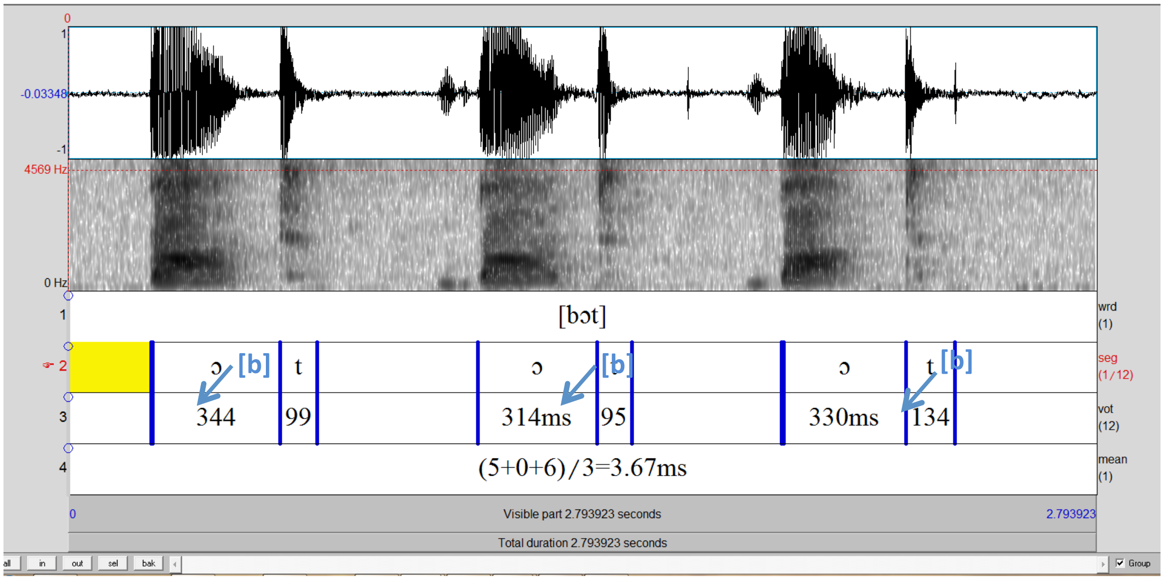

Figure 18: <bought>

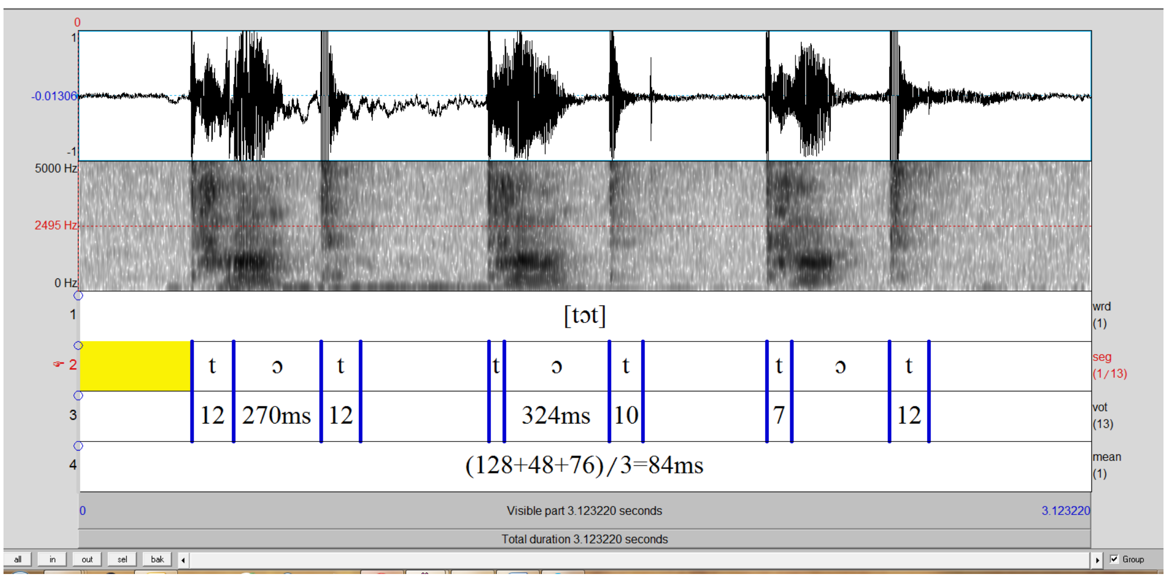

Figure 19: <tot>

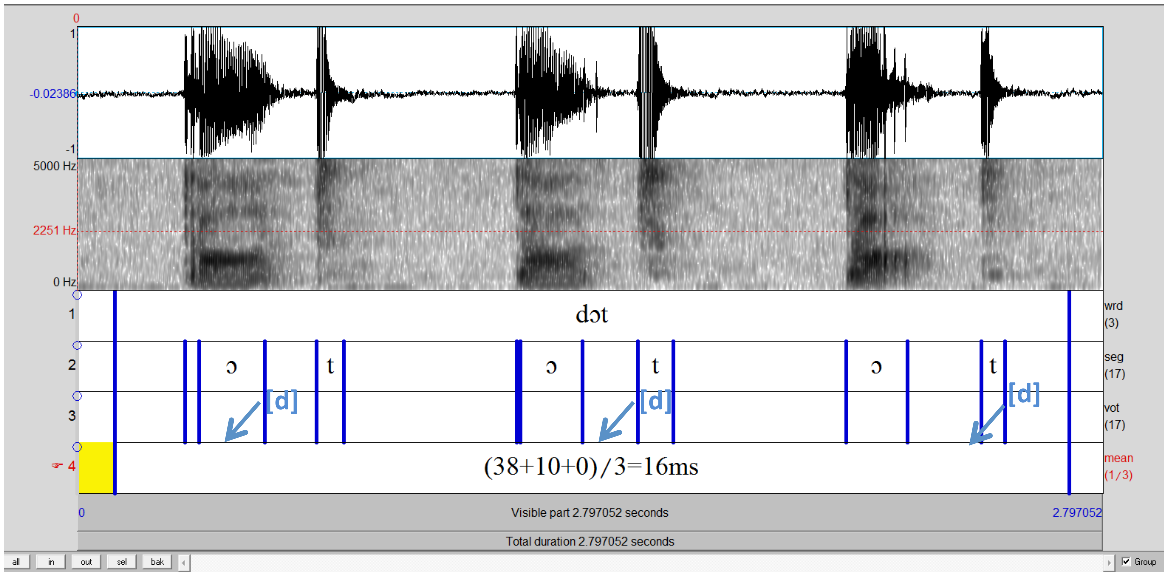

Figure 20: <dot>

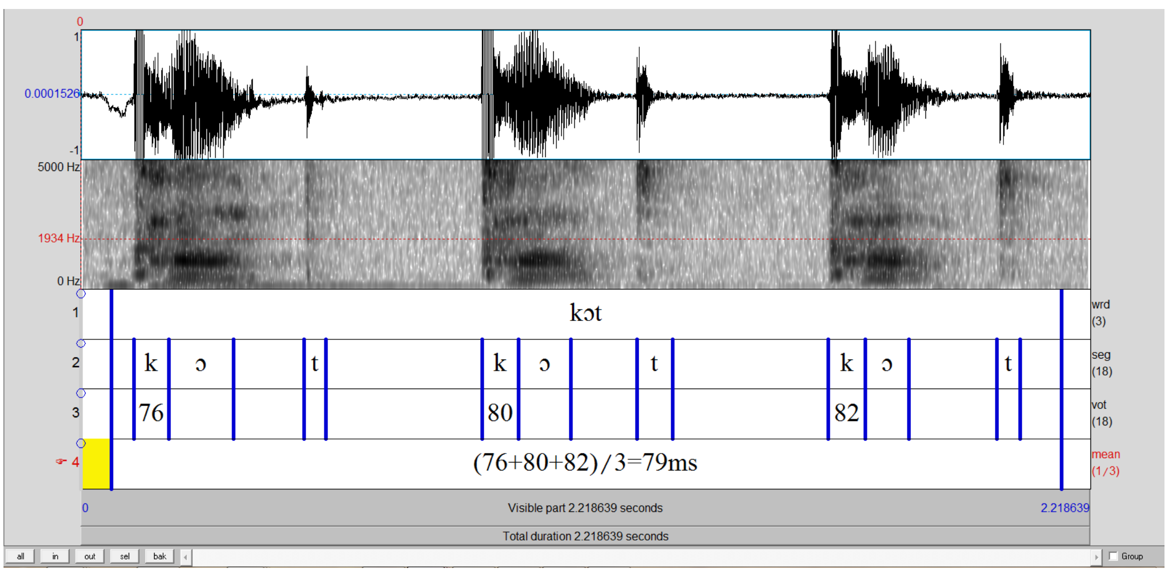

Figure 21: <cot>

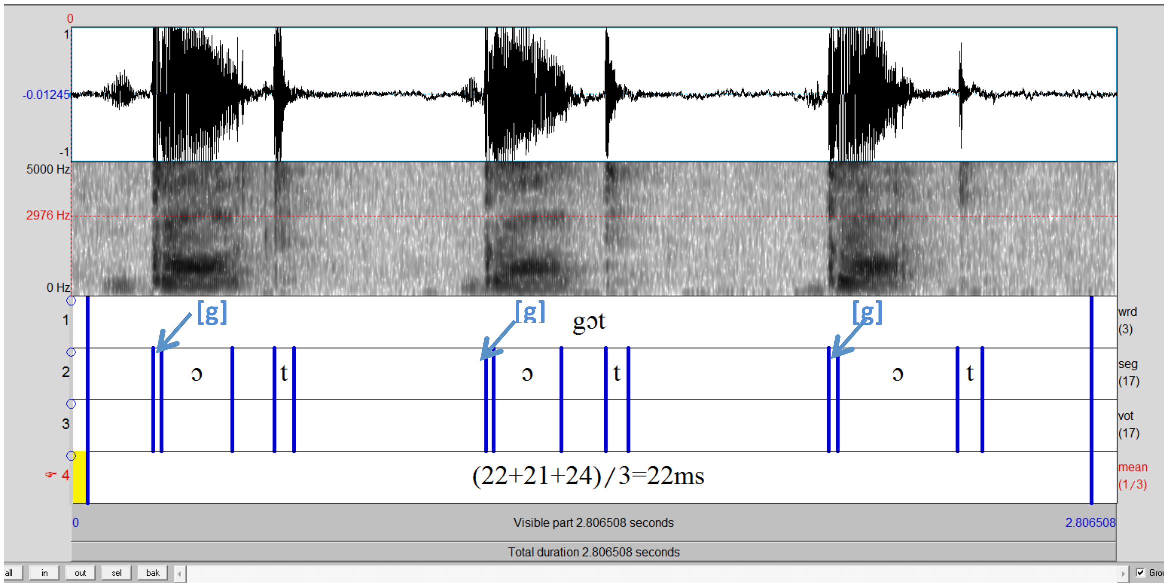

Figure 22: <got>

Word	pot	bought	tot	Dot	cot	got
Segment	/p/	/b/	/t/	/d/	/k/	/g/
Mean of VOT	60ms	3.67ms	84ms	10ms	79ms	22ms

Table 4: VOT of Stop Consonants

Table 4 shows the VOT of the six stop consonants. My voiceless stop consonants are longer than those of my voiced consonants. This follows the theory (Koffi, 2013) that the VOT of a voiced consonant is equal or less than 30ms, and that it is longer than 30ms for a voiceless consonant. In Figure 19, the duration of /t/ is 84ms, so it is aspirated and labeled as [tʰ]. The phonological rules of the phoneme /t/ are stated as follows:

/t/ → [tʰ] #__ {+ stressed, + syllable}

The other voiceless stop consonants, /p/ and /k/, are also aspirated, having a longer duration at the word-initial position.

In addition, Table 4 also shows that VOTs vary from one place of articulation to another. Usually, the closer to the front of the mouth, the shorter VOT a sound has (Koffi, 2010). For instance, the VOT of a bilabial consonant is shorter than that of a velar consonant. Duration progress is noted from shorter to longer as [pʰ] < [tʰ] < [kʰ] and [b] < [d] < [g]. In my idiolect, only the voiced consonants follow this rule, /b/ < /d/ < /g/, while the voiceless consonant [tʰ] is longer than [kʰ].

In Figure 20, the voiced consonant /d/ is not obvious in the spectrograms, which might be because my first language is Mandarin Chinese. In Mandarin, /b/, /d/, and /g/ are not phonemes (IPA for Mandarin, n.d.). In my recordings, two out of three times the voiced

/d/ is close to a non-aspirated sound [t], which exists in Mandarin. However, according to Figures 18 and 22, I have acquired the voiced sounds [b] and [g] in my second language. I suppose that is because of the two years that I have spent living in the US. In addition, /t^h/, /p^h/, and /k^h/ are also different phonemes in Mandarin, so I don't have problems with these sounds in English. In Figures 18 and 22, extra-long vibrations appear in the beginning before the burst. I assume that this is because I consciously pronounced these words and elongated them as I spoke.

3.2 Vowel Initial

In American English, glottal stops occur optionally before vowel-initial words (Garellek, 2012). This section focuses on whether a glottal stop occurs in the word-initial position of the phrase "oozing blood".

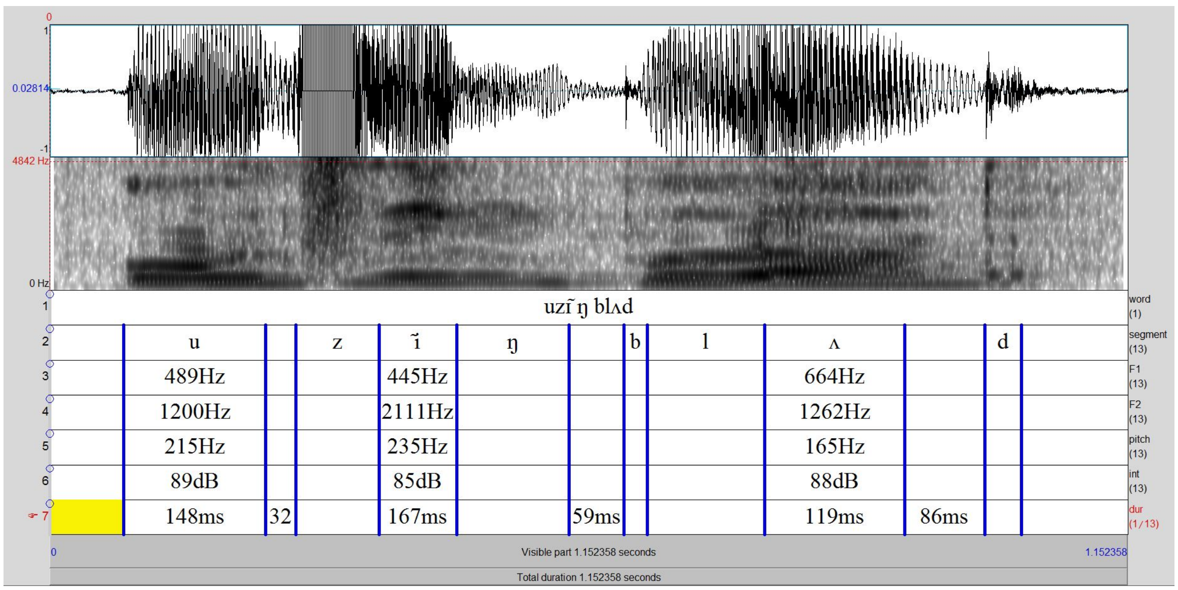

Figure 23: <oozing blood>

In Figure 23, there is not a glottal stop preceding the initial vowel [u]. However, a space between [u] and [z] shows a glottal stop caused by a vowel merging to a consonant. Another glottal stop occurs between [ʌ] and [d].

3.3 Frication

Frication can be observed as a dark stream from the top of an acoustic spectrogram. An affricate is a combination of a stop and a fricative, so it is presented as a sharp, narrow bar followed by a dark stream. In this section, affricates [tʃ], [dʒ], [tɹ], and [dɹ] in the words <church>, <judge>, <trial>, and <drive> are analyzed to see whether I have acquired the pronunciation of these affricates.

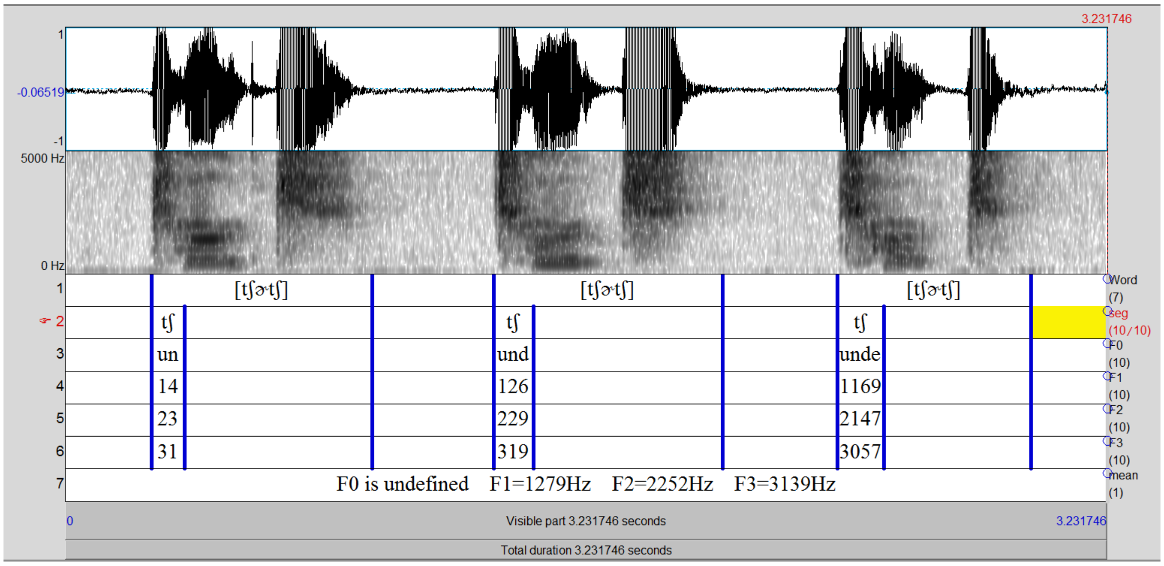

Figure 24: <church>

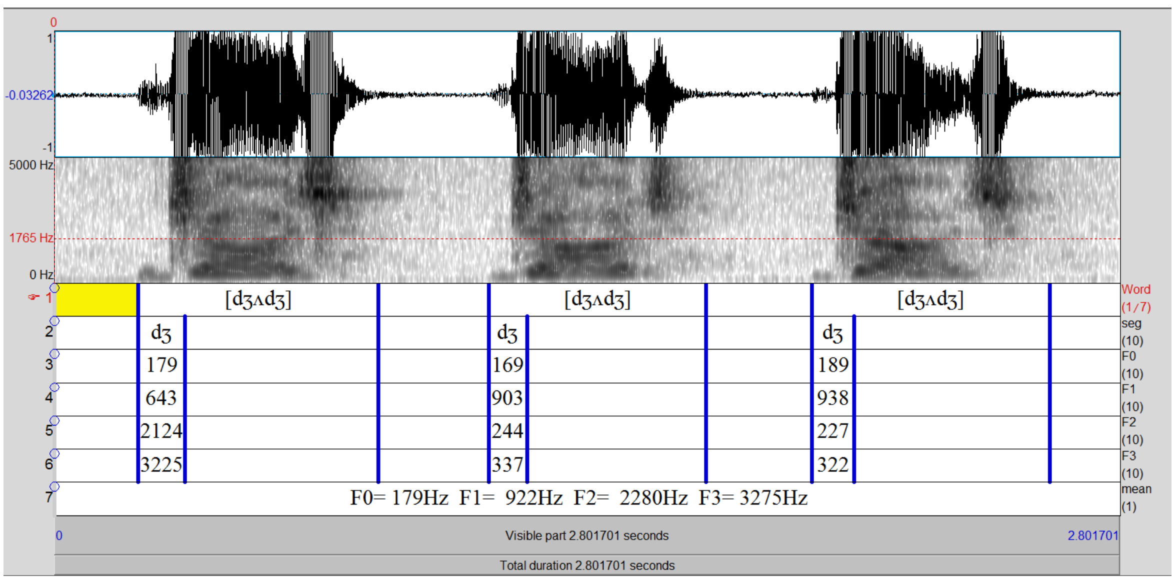

Figure 25: <judge>

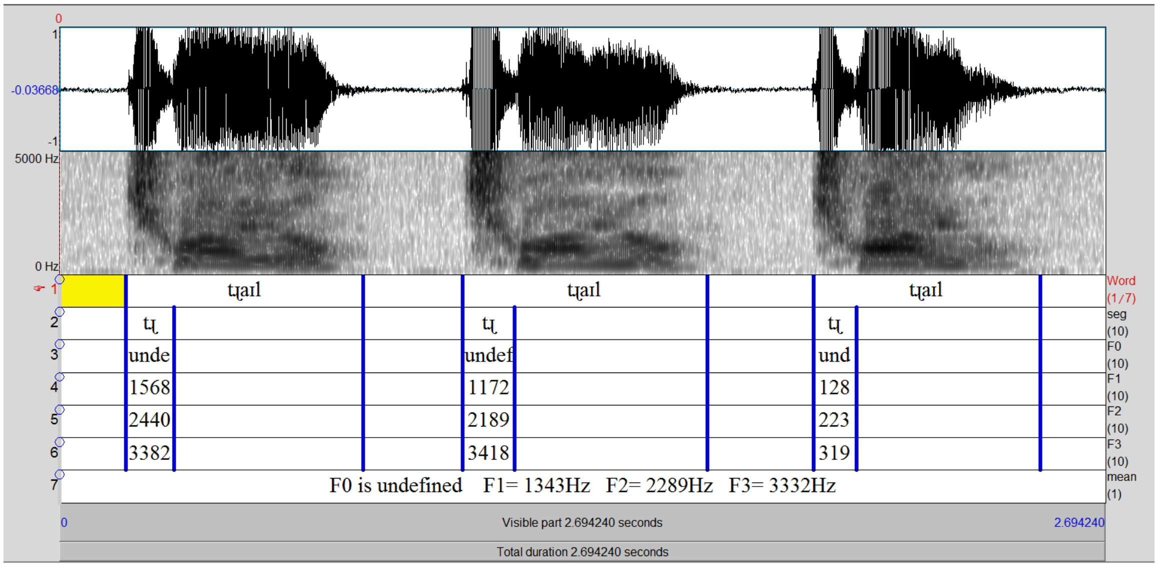

Figure 26: <trail>

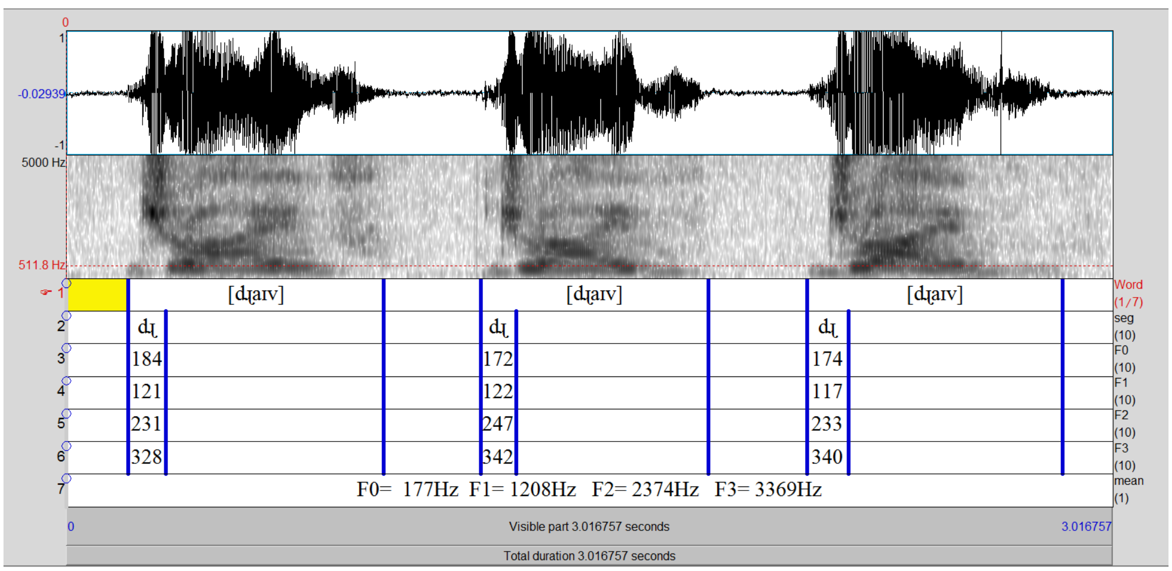

Figure 27: <drive>

The stop feature in affricates does not always exist in my idiolect, and the voiceless stop [t] occurs more often than the voiced stop [d] in word-initial position. As mentioned in the VOT section, I have not acquired the voiced stop consonants, which may lead to difficulties with pronouncing voiced affricates. Examining Figures 24 through 27, without a stop, my pronunciation of affricates makes them look more like fricatives.

3.4 Sonorants

All of the vowels and some consonants such as approximants, nasals, and flaps are considered sonorants. Approximants [l] and [r/ ɹ] and nasals [n] and [m] in the words <lull>, <rear>, <nun>, and <mom> are tested in this project. These sonorants can be seen as syllabic in English by displaying vowel-like periodical waves in spectrograms. Koffi (2013) has hypothesized that the sonorants in the coda tend to be more vowel-like than those in the onset.

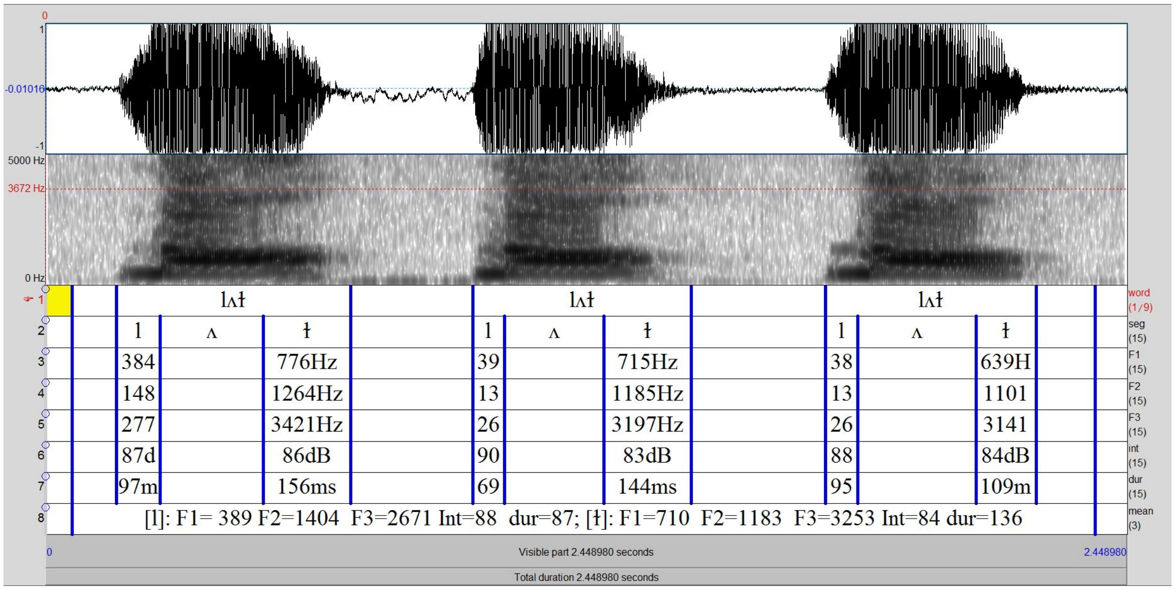

Figure 28: <lull>

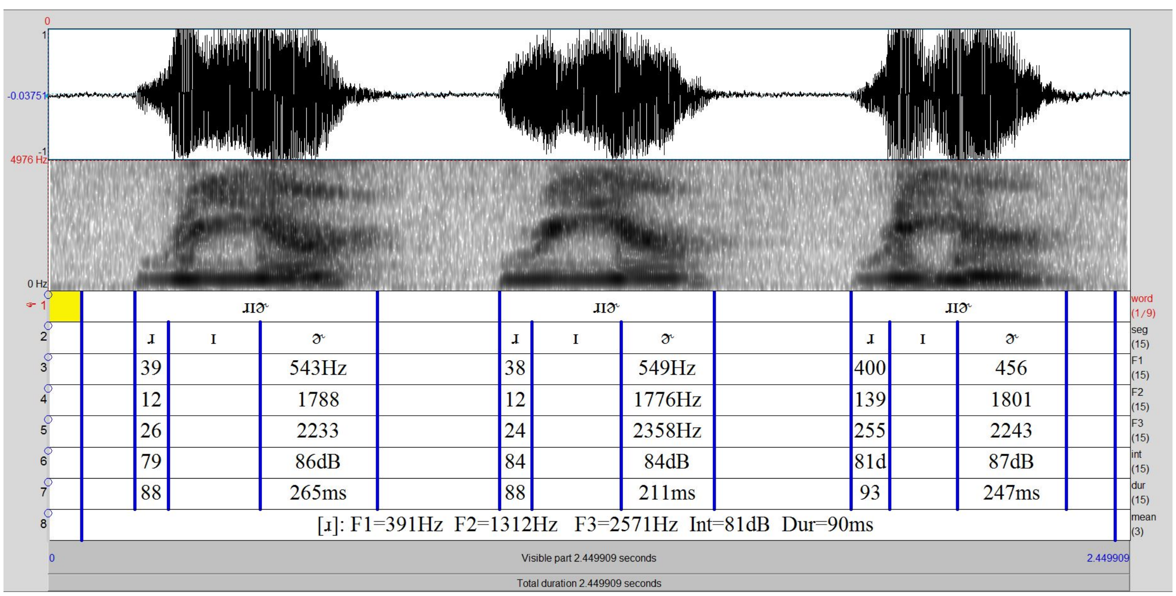

Figure 29: <rear>

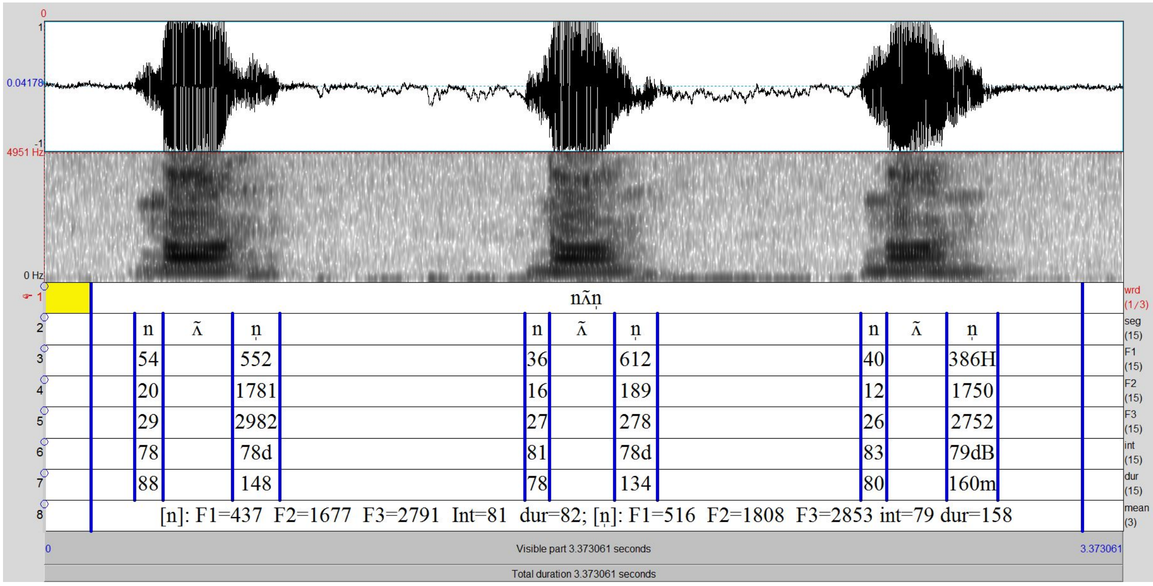

Figure 30: <nun>

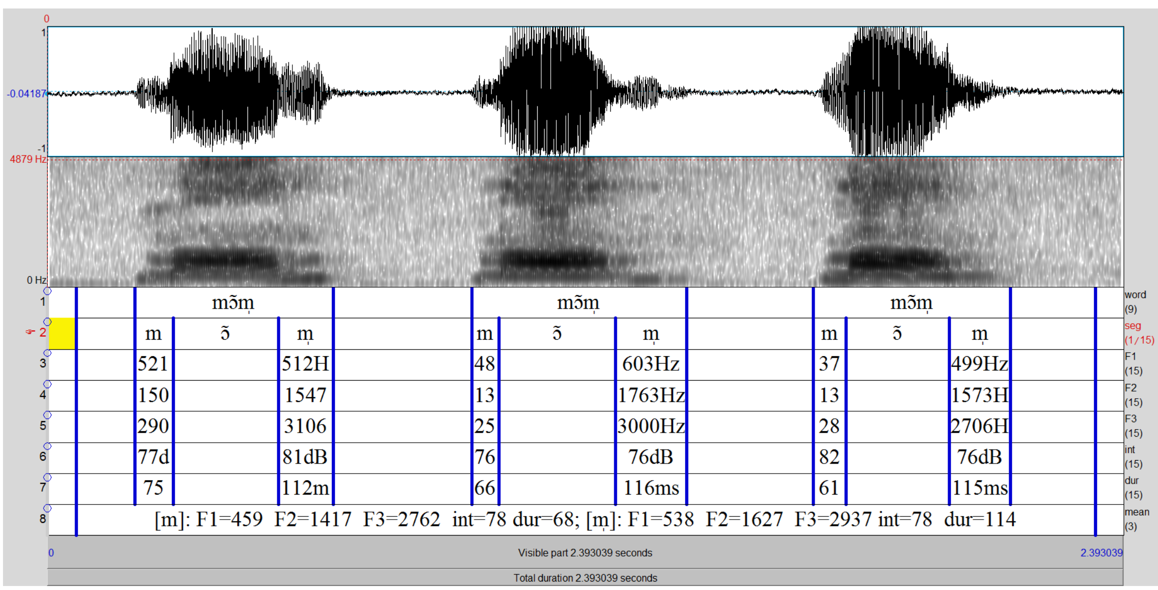

Figure 31: <mom>

Word	<lull>		<rear>		<nun>		<mom>	
Segment	Onset [l]	Coda [l]	Onset [ɹ]	Coda [ɹ]	Onset [n]	Coda [n]	Onset [m]	Coda [m]
Duration (ms)	87	136	90	241	82	158	68	114
Difference Between Onset and Coda (ms)	49		151		76		46	

Table 5: Duration Difference by Borui

All of my post-vocalic sonorants are longer. The example of the largest difference between onset and coda is 151ms for the approximant [ɹ]. As can well be seen in Figure 29, the vowel [ɚ] is actually pronounced in the coda with a higher pitch and a louder intensity.

4.0 Suprasegmentals

Besides place and manner of articulation, suprasegmental features generally include pitch, intensity, and duration. English is considered a stress-timed language, which means that at least one syllable is stressed in a word, and the duration of syllables is either longer or shorter to make sense for listeners (See https://www.youtube.com/watch?v=PrAe07KluZY for an example). Unlike English, languages such like French, Spanish, and Cantonese are syllable-timed languages in which all syllables maintain the same pitch, intensity, and duration. Mandarin Chinese is a tonal language in which pitch is used for distinguishing meaning. In this section, I attempt to find out which of these three language methods I primarily use to signal the strong syllable in my idiolect. The following spectrograms show the acoustic strategies that I use to pronounce three types of English words.

4.1 Homographs

Homographs are defined as words that have the same spelling but different pronunciations and different meanings. The different pronunciations may be caused by suprasegmental changes or/and segmental changes. Fry's study shows a hierarchy correlation among suprasegmental (Koffi, 2013:164) features that signal the stress of words, represented as Pitch > Intensity > Duration.

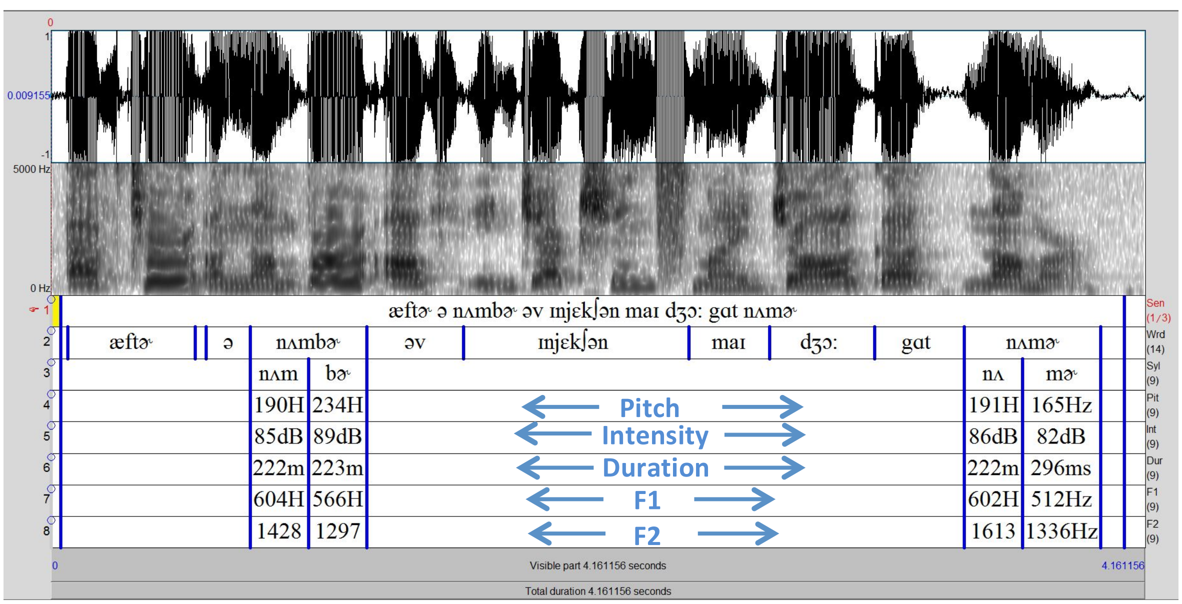

Figure 32: <after a number of injections my jaw hot number>

No.	Word	subject		Subject	
1	IPA	/nʌm bər/		/nʌmər/	
2	POS	Noun		Comparative Adjective	
3	Syllables	[nʌm]	[bər]	[nʌm]	[mər]
4	Pitch (Hz)	190	234	191	165
5	Intensity (dB)	85	89	86	82
6	Duration (ms)	222	223	222	296

Table 6: Disyllable Words

In Figure 32, the word <number> is a homographic disyllable word. The higher pitch and intensity show that the stress falls on the second syllable in the word as a noun. In the word as an adjective, the same features show that the first syllable is stressed. This result agrees with the Fry's findings.

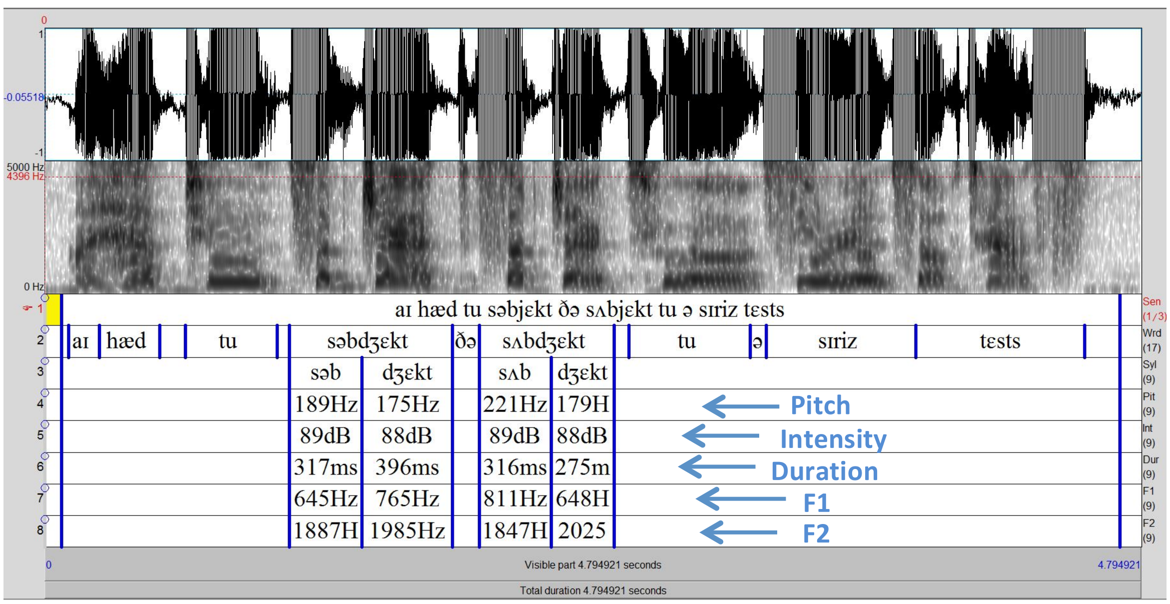

Figure 33: <I had to subject the subject to a series tests>

No.	Word	subject		subject	
1	IPA	/səbdʒɛkt/		/sʌbdʒɪkt/	
2	POS	Verb		Noun	
3	Syllables	[səb]	[dʒɛkt]	[sʌb]	[dʒɛkt]
4	Pitch (Hz)	189	175	221	179
5	Intensity (dB)	89	88	89	88
6	Duration (ms)	317	396	316	275

Table 7: <subject>

In Figure 33, duration is the key for identifying the stressed syllable. The difference of F1in the two syllables indicates that I produce the two words differently. According to Figures 32 and 33, Fry's hierarchy features do not apply to me consistently.

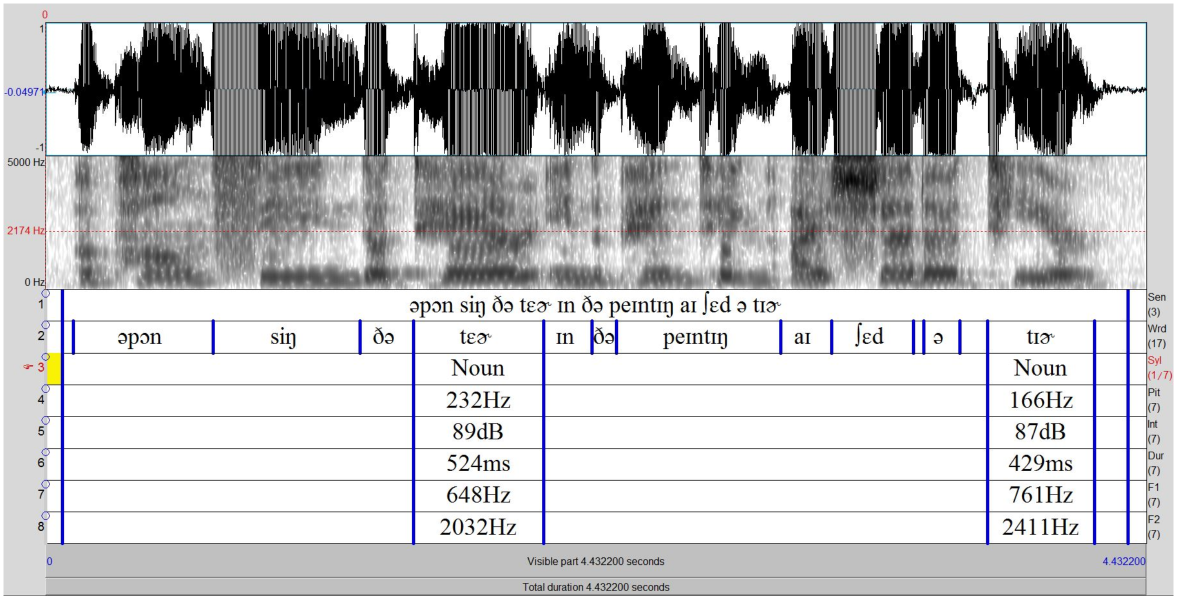

Figure 34: <upon seeing the tear in the painting I shed a tear>

The first <tear> is a synonym of the word 'rip,' and the second <tear> refers to moisture produced by ducts of the eye. There is a significant difference of 388Hz in F2 between these two words, which means that I produce these two words differently.

4.2 Stress Pattern with Three-Syllable words

In this section I pronounce the words <academe>, <universe>, <element>, <fundament>, and <activate> each three times in a row. I reference phonemic transcripts from dictionary.com, which provide examples of both primary and secondary stress for GAE. In this project, I only concentrate on the primary stress.

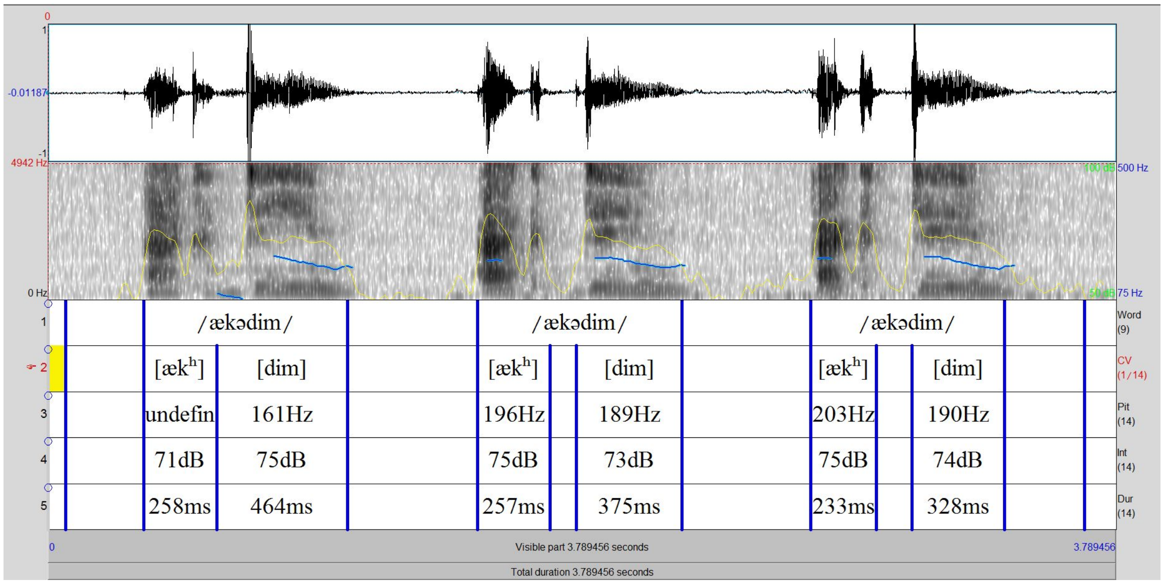

Figure 35: <academe>

No.	Word	Academe	
1	IPA	/ˌækə’dim/	
2	Syllables	[æk^h]	[dim]
3	Pitch (Hz)	199[4]	180
4	Intensity (dB)	74	74
5	Duration (ms)	249	389

Table 8: <academe>

The stressed syllable in my idiolect is the last syllable [dim], which agrees with dictionary.com. I use duration to identify the stress in this word. However, I reduce the three syllables to two, which does not match with the IPA transcript /ækə’dimiə/. The schwa does not exist in my idiolect; instead I aspirate the stop consonant [k] to [k^h].

[4] Since the pitch is undefined the first time, the mean is calculated from the latter two.

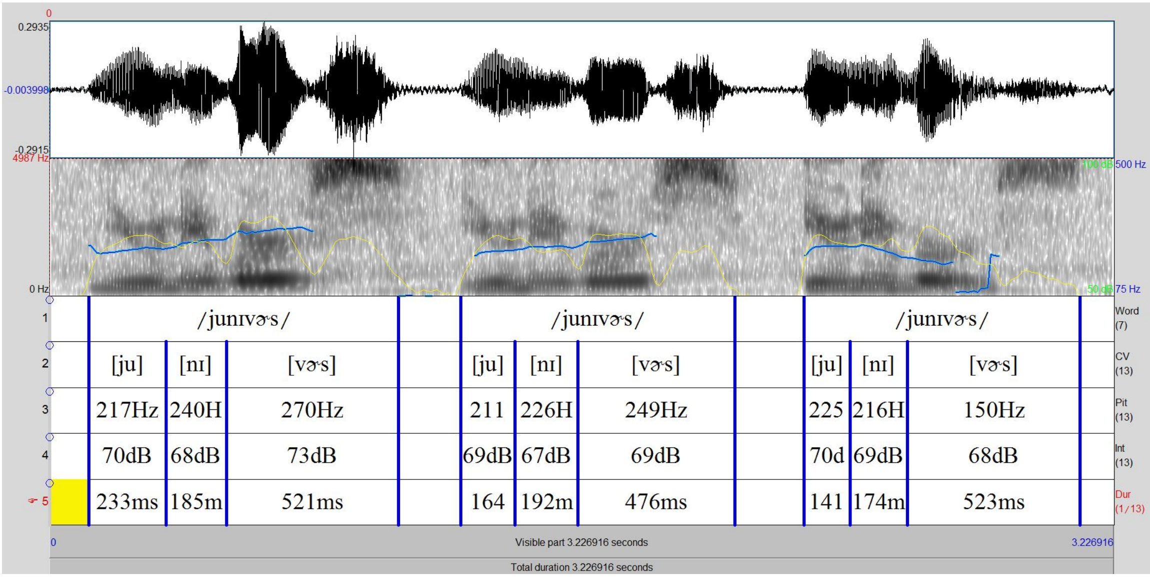

Figure 36: <universe>

No.	Word	universe		
1	IPA	/ˈyunəˌvɜrs/		
2	Syllables	[ju]	[nɪ]	[vɚs]
3	Pitch (Hz)	218	227	223
4	Intensity (dB)	70	68	71
5	Duration (ms)	179	184	506

Table 9: <universe>

According to dictionary.com, the stress falls on the first syllable [yu], but in my idiolect the last syllable [vɚs] is stressed with greater intensity and a longer duration. Also, I pronounce the front high vowel [ɪ] instead of a schwa in the unstressed syllable [nɪ].

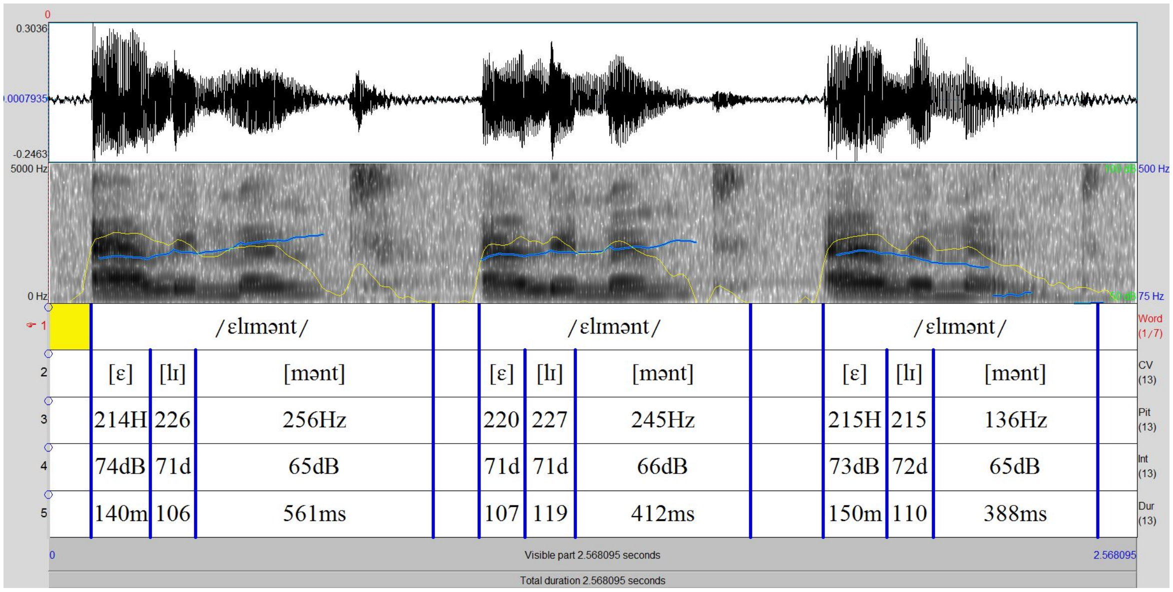

Figure 37: <element>

No.	Word	element		
1	IPA	/ˈɛl ə mənt/		
2	Syllables	[ɛ]	[lɪ]	[mənt]
3	Pitch (Hz)	216	223	212
4	Intensity (dB)	73	71	65
5	Duration (ms)	132	112	266

Table 10: <element>

My strong syllable is [ɛ] which matches the phonemic transcript. I again used intensity and duration but not pitch to signal the stressed syllable. Although the duration of the strong syllable is shorter than the last one [mənt], 132ms is long enough for a nucleus to be a syllable.

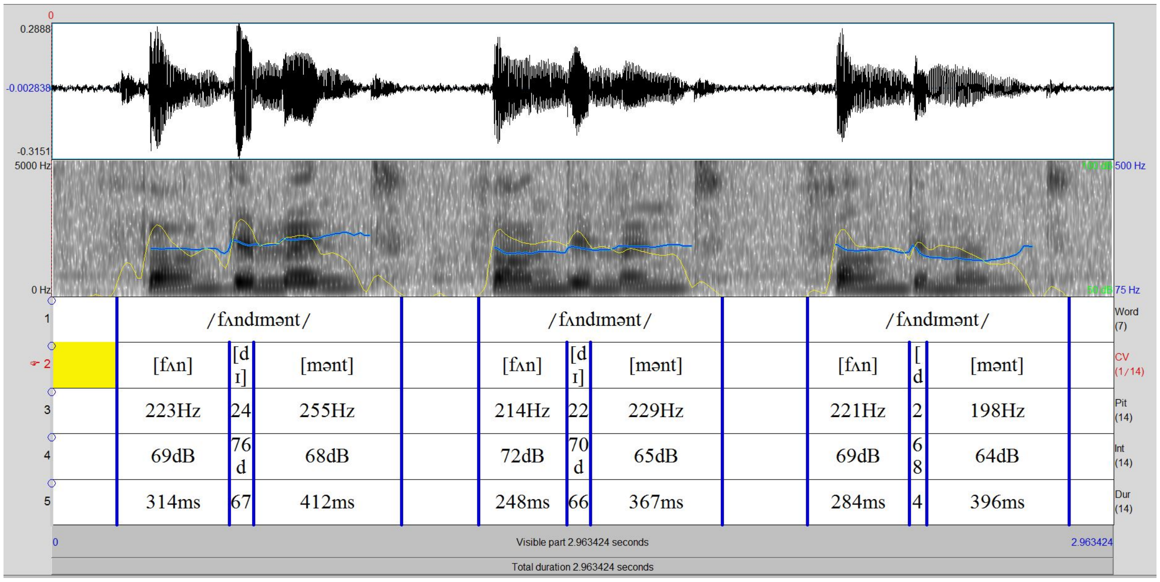

Figure 38: <fundament>

No.	Word	fundament		
1	IPA	/ˈfʌn də mənt/		
2	Syllables	[fʌn]	[dɪ]	[mənt]
3	Pitch (Hz)	219	228	227
4	Intensity (dB)	70	71	65
5	Duration (ms)	282	60	391

Table 11: <fundament>

The strong syllable is [fʌn] from dictionary.com, while the one in my idiolect is either [fʌn] or [dɪ]. The acoustic correlate I use is not very predictable in this word. In Figure 38, the syllable [fʌn] maintains the highest intensity among syllables two out of three times. However, according to Table 11, both the mean pitch and the intensity of the strong syllable [dɪ] are slightly higher than those of the other syllables.

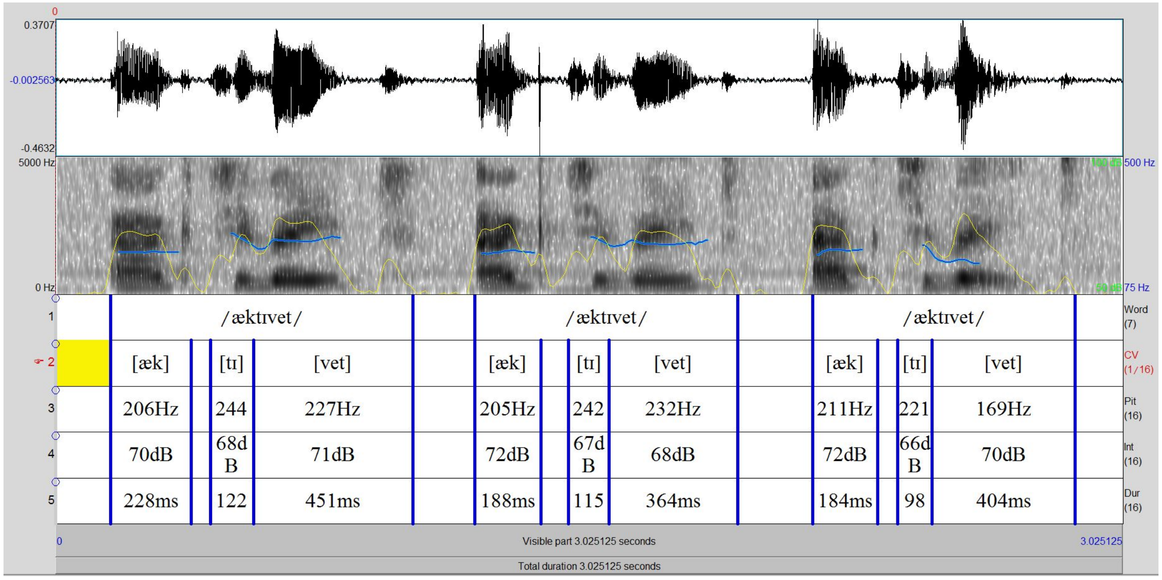

Figure 39: <activate>

No.	Word	Activate		
1	IPA	/ˈæk tə ˌveɪt/		
2	Syllables	[æk]	[tɪ]	[vet]
3	Pitch (Hz)	208	236	200
4	Intensity (dB)	71	67	70
5	Duration (ms)	200	112	405

Table 12: <activate>

My strong syllable is [æk] which matches with dictionary.com. Again, the acoustic correlates of intensity and duration assist to support this result.

Although the testing words are limited, the evidence above indicates that I tend to use intensity primarily and duration secondarily to produce three-syllable words. Therefore, Fry's theory of acoustic correlate hierarchy does not apply in my idiolect of English. Besides, I substitute a schwa for the high front lax vowel [ɪ] in some unstressed syllables.

4.3 Polysyllabic Stress

The words that I use in this project are based on the words used in the previous section. All of the words used in this section are words from the previous section with added derivational suffixes such as <-ic>, <-ation>, <-ity>, etc. Kreidler's study has found some patterns of changing stress through suffixation (Koffi, 2013:171). Suffixes like <-al> are marked as "neutral" which means that the stress syllable of a word will not change to another syllable after adding a neutral suffix. Other suffixes move stress, referred to as "stress-shifting" suffixes, such as <-ic> and <-ity>. Another type of suffix directly attracts the stress to themselves, such as <-ation>. In this section, I attempt to investigate the acoustic correlates that I use to produce the stress on multi-syllable words.

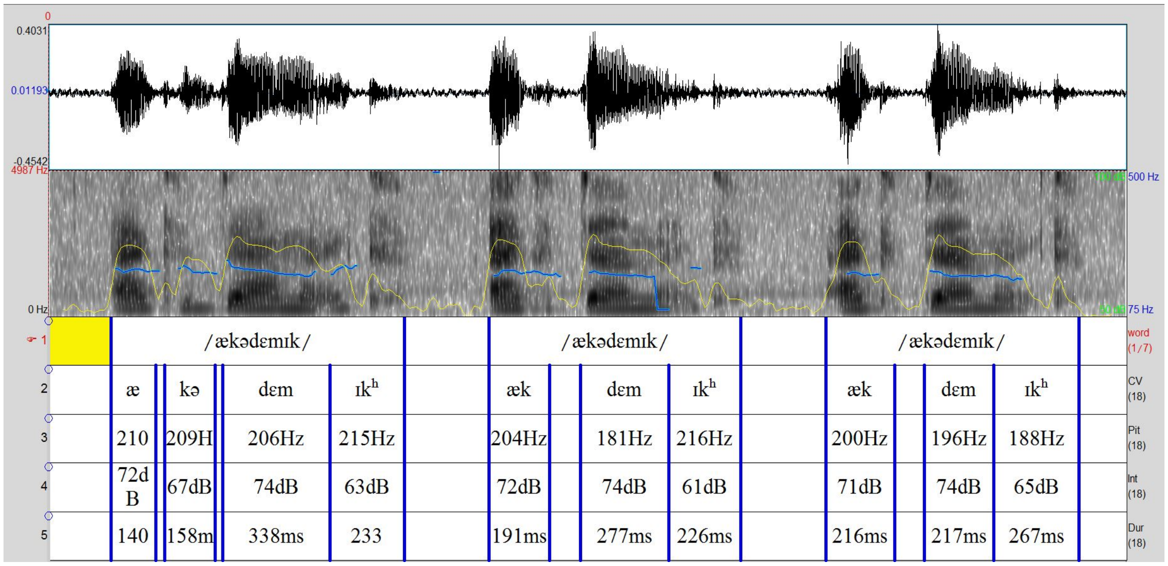

Figure 40: <academic>

No.	Word	academic		
1	IPA	/ˌækəˈdɛmɪk/		
2	Syllables	[æk]	[dɛm]	[ɪkʰ]
3	Pitch (Hz)	202	188	202
4	Intensity (dB)	72	74	63
5	Duration (ms)	204	247	246

Table 13: <academic>[5]

The strong syllable is [dɛm] in my idiolect, which matches with dictionary.com. Although the other two syllables have higher pitch, [dɛm] has the loudest intensity. Compared with the word <academe> in Figure 35, the deviational suffix <-ic> does not shift stress in my pronunciation. Besides, I again reduce the syllable [kə] and only produce a three-syllable word.

[5] The mean values are calculated from the latter two trials.

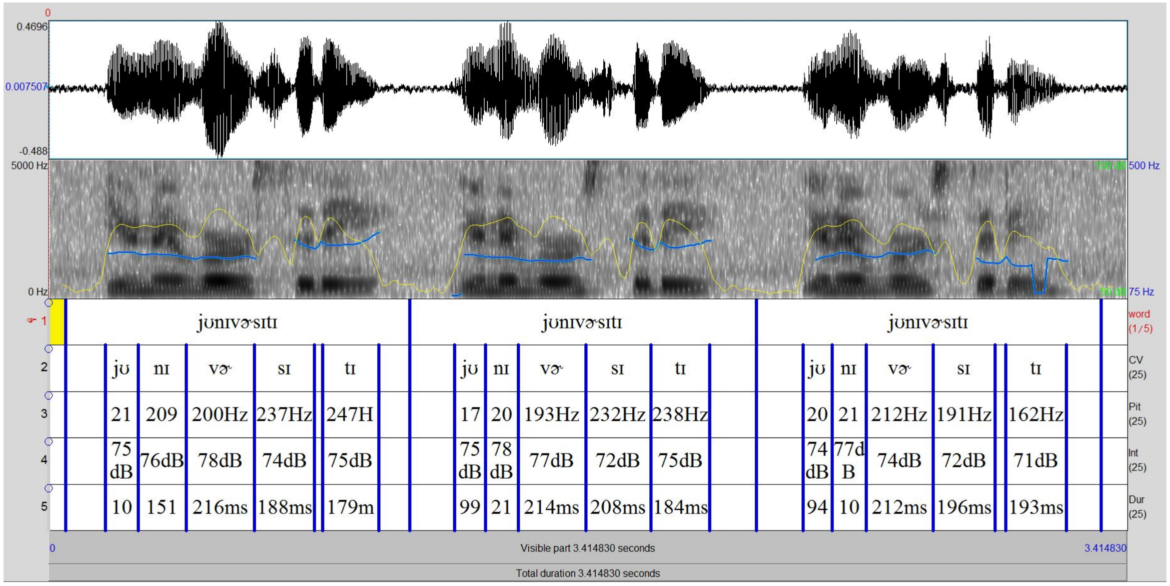

Figure 41: <university>

No.	Word	University				
1	IPA	/ˌyu nəˈvɜr sɪ ti/				
2	Syllables	[jʊ]	[nɪ]	[vɚ]	[sɪ]	[tɪ]
3	Pitch (Hz)	198	207	202	220	216
4	Intensity (dB)	75	76	76	73	73
5	Duration (ms)	99	120	214	197	185

Table 14: <university>

In this example I am not very confident in identifying the strong syllable in my idiolect because the highest pitch falls on [sɪ], the greatest intensity falls on the syllables [nɪ] and [vɚ], and the longest duration falls on [vɚ]. Based on my previous conclusion that intensity is the primary strategy that I use and duration is the secondary one, [vɚ] can be the strong syllable in this case. This matches the transcript from dictionary.com. Comparing this example with Figure 36, the derivational suffix <-ity> does not make the stress move onto another syllable.

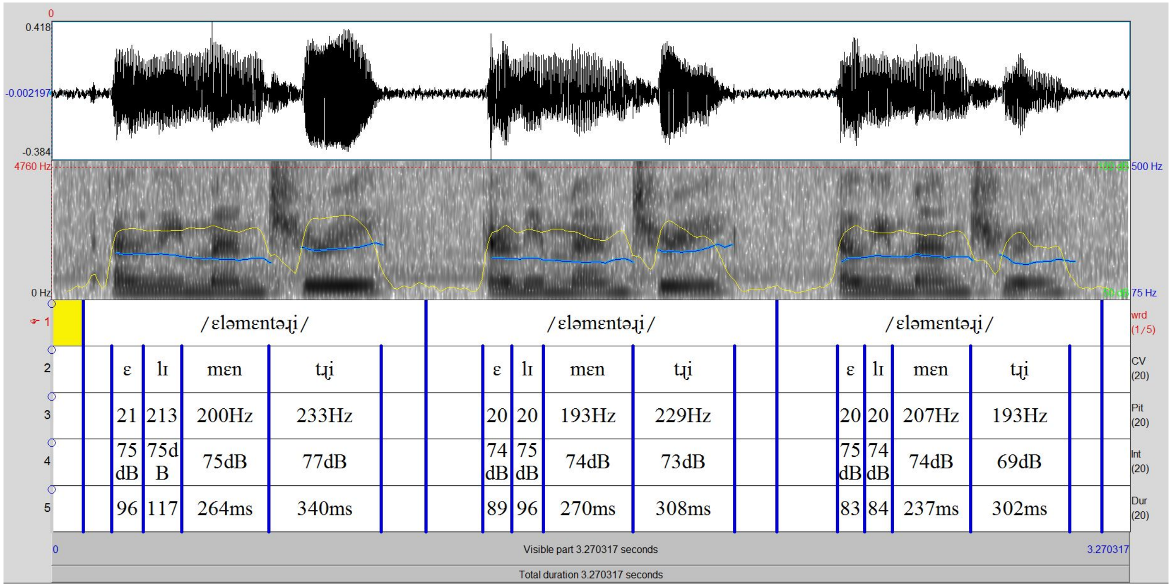

Figure 42: < elementary>

No.	Word	elementary			
1	IPA	/ˌɛlə'mɛn tə ri/			
2	Syllables	[ɛ]	[lɪ]	[mɛn]	[tɹi]
3	Pitch (Hz)	206	207	200	218
4	Intensity (dB)	75	75	74	73
5	Duration (ms)	89	99	257	317

Table 15: < elementary>

This is another example where it is difficult to predict the strong syllable in my idiolect because the highest pitch falls on [tɹi], the greatest intensity falls on the syllables [ɛ] and [lɪ], and the longest duration falls on [mɛn]. Therefore, the stress is unpredictable in this case, and it is difficult to predict whether the derivational suffix <-ary> caused a stress movement. In addition, I substitute the syllables [tə] and [ri] with one syllable [tɹi] in my idiolect.

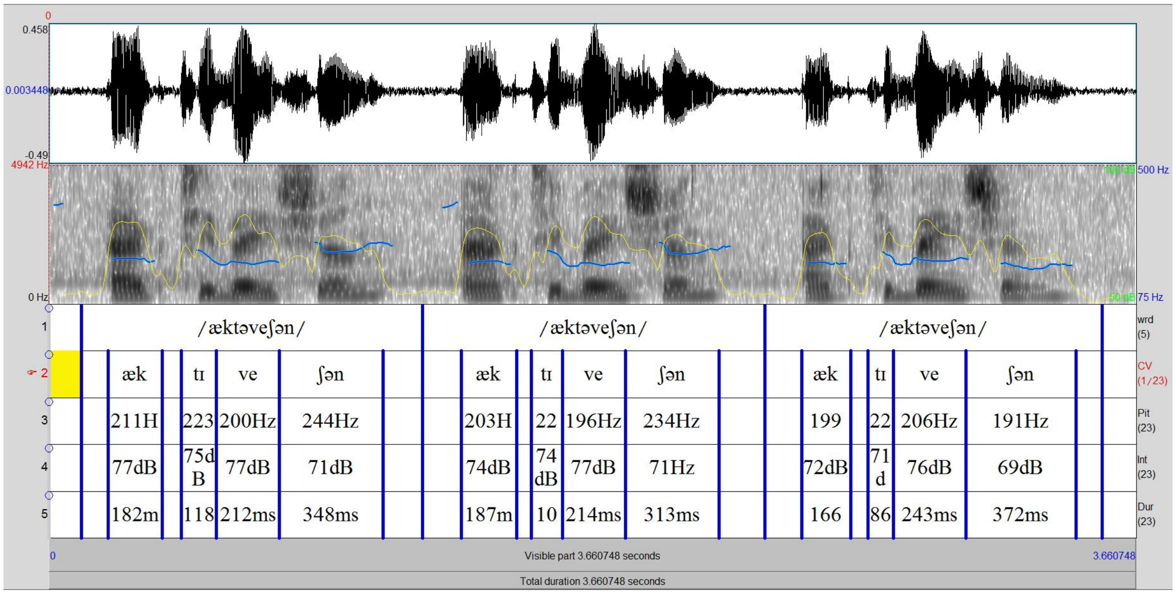

Figure 43: < activation >

No.	Word	activation			
1	IPA	/æktə’veʃən/			
2	Syllables	[æk]	[tɪ]	[ve]	[ʃən]
3	Pitch (Hz)	204	225	201	223
4	Intensity (dB)	74	73	77	70
5	Duration (ms)	178	103	223	344

Table 16: <activation>

The stress syllable in this example is [ve] in my idiolect, which matches with dictionary.com. It maintains the greatest intensity, though the pitch is not the highest one. Again, from my personal observation so far, pitch does not seem to be the strategy that I use to produce words in English.

In addition, a schwa is in the unstressed syllable [tə] from dictionary.com, but I replace it with the front high lax vowel [tɪ] in my idiolect. Moreover, comparing this example with Figure 39, the derivational suffix <-ation> successfully attracts the stress onto itself.

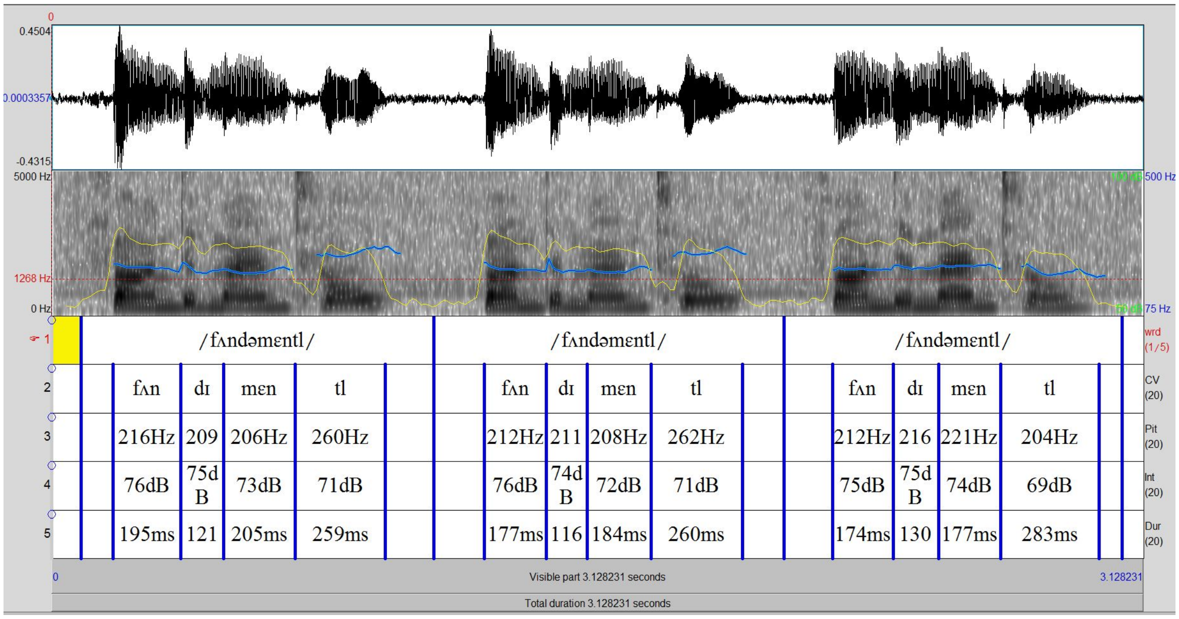

Figure 44: <fundamental>

No.	Word	fundamental			
1	IPA	/ˌfʌn dəˈmɛn tl/			
2	Syllables	[fʌn]	[dɪ]	[mɛn]	[tl̩]
3	Pitch (Hz)	213	212	212	242
4	Intensity (dB)	76	75	73	70
5	Duration (ms)	182	122	189	267

Table 17: <fundamental>

According to dictionary.com, [mɛn] is the stressed syllable in the word <fundamental>, and [fʌn] is the stressed syllable in the word <fundament>. This goes against the finding from the Kreidler study that the neutral suffix <-al> does not change the stress pattern. In my idiolect [fʌn] is the strong syllable in both of the two words, which means that Kreidler's finding about <-al> does apply in my idiolect.

Through the projects I have done in this section, I think intensity is the key to signaling primary stress in my idiolect. The schwa never occurs in any words that I produce. Instead I often replace the schwa with the high front vowel [ɪ], which occurs in unstressed syllables in British English. Difficulties of signaling stress in some cases indicate that my acoustic strategy is somewhat inconsistent.

5.0 Showing My Pronunciation

This exercise is for self-observing how phonological rules apply in my particular pronunciation, step-by-step from mental phonemic representation to the real surface phonetic structure. I first pronounce the word <tempest>, then analyze each segment on Praat and compare the values with the phonological rules in our work book. I follow the derivational model shown below (Koffi, 2013):

Underlying phonemic representation	/t ɛ m p ɛ s t/
Stress rule	ɛ́
Schwa rule	ə
Aspiration rule	tʰ
Nasalization rule	ɛ̃
Unreleasing	t̚
Surface phonetic realization	[tʰ ɛ̃ m p ə s t̚]

Figure 45: Lexical Phonology of <tempest>

The nasalization rule cannot be verified in this project due to limitations of the software tool. The existence of the other phonological rules can be proven through the spectrogram below.

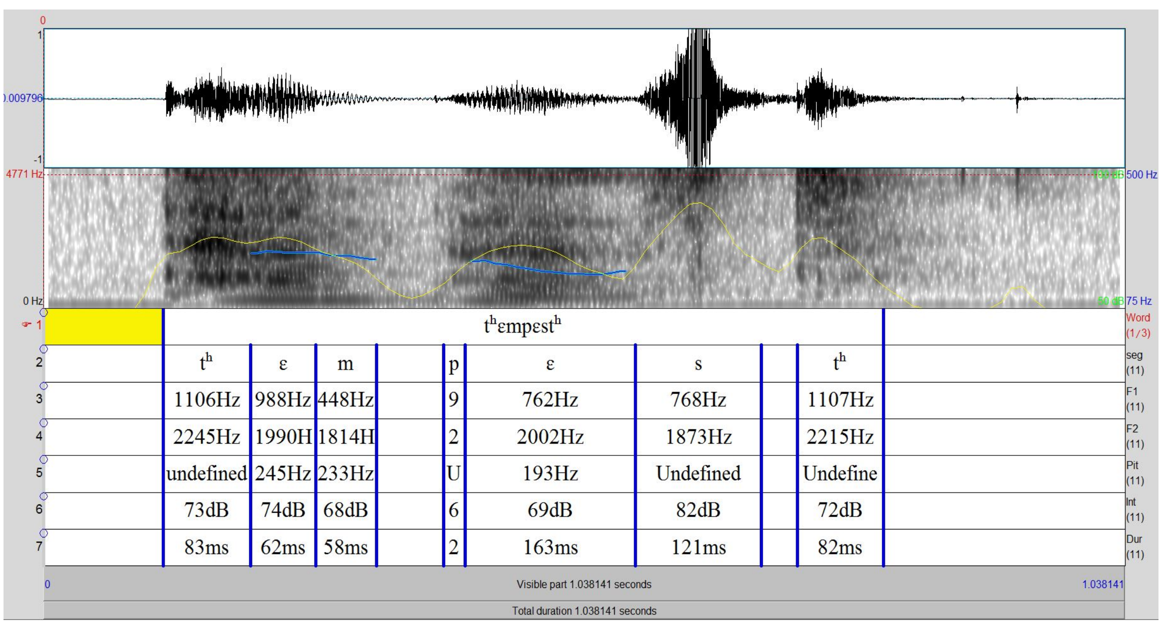

Figure 46: <tempest>

According to the measurements and spectrograms, the vowels that I produce in the two syllables of <tempest> are both [ɛ]. My accented [ɛ] is close to the /æ/ in GAE. These two sounds overlap each other in my idiolect, as shown in Figure 13. Therefore, the schwa rule does not apply in my case. In addition, unlike in the releasing rule introduced in Figure 45, the coda, the voiceless stop [t], is aspirated in my idiolect. Moreover, the voiceless stop [p] does not show a clear burst because I do not press my lips tightly to produce words in my mother tongue. The voiceless fricative [s] has the greatest intensity (82dB), which shows that I pronounce this sound very strongly. Since the first vowel [ɛ] has the second greatest intensity, I consider that the stress falls on the first syllable.

6.0 Summary

I am a 24-year old female from the Northeast region of China. My first language is Mandarin Chinese. I started learning British English when I was 10 years old and continued until senior high school when American English started being taught. I came to the US two years ago and have been practicing my American English with native speakers since I arrived.

I have discovered some interesting patterns in my idiolect throughout this portfolio. Before taking this class I just knew that I did not speak English like a native speaker; this project taught me how to investigate the way in which I speak differently by examining the details of my speech. The nonexistence of voiced stop consonants in Mandarin has caused some difficulties in my acquisition of those consonant sounds in English. Vowel harmony and vowel centralization occur in my idiolect too. Because of the difference in the aspect between a tonal language and a stress-timed language, my acoustic strategies for signaling primary stress in English are not very clear and seldom consistent. In addition, the British English that I was taught in childhood has affected my current English pronunciation by mixing American and British accents. The existence of the diphthong [eɪ], and the vowel [ɪ] replacing schwa are examples of this in my current English pronunciation.

ABOUT THE AUTHOR

Borui Zhang is a current student in the Teaching English as a Second Language graduate program at St. Cloud State University, in St. Cloud, MN. After graduation she will be pursuing further study in computational linguistics. She is interested in the pronunciation of language, voice analysis, and computational phonology. She can be reached via email at: zhbo1101@stcloudstate.edu or breezhang8160@gmail.com.

Recommendation: This work is a compilation of the various phonetic lab assignments in English 475/573 (Laboratory Phonology) course offered in Spring 2013. It was chosen for publication for its thoroughness in depicting Ms. Zhang's idiolect of English. It was recommended for publication by Professor Ettien Koffi, Ph.D., Linguistics Department, St. Cloud State University, St. Cloud, MN. Email: enkoffi@stcloudstate.edu

References

Garellek, M. (2012). Glottal stops before word-initial vowels in American English: distribution and acoustic characteristics. UCLA Working Papers in Phonetics, 110, 1-23.

Koffi, Ettien N. and Ruanglertslip, Ekkarat 2013. An Acoustic Phonetic Portfolio of A Thai-Accented English Idiolect. Linguistics Portfolios (2), 144-178.

Koffi, Ettien N. 2013. *Phonology1:[fə'netɪks and fə'nalədʒi]*. Unpublished Course Manuscript.

IPA for Mandarin. (n.d.). In Wikipedia. Retrieved October 6, 2013, from http://en.wikipedia.org/wiki/Help:IPA_for_Mandarin