TRABAJO FIN DE GRADO

GRADO EN ESTUDIOS INGLESES: LENGUA, LITERATURA Y CULTURA

The production of four diphthongs in RP and Australian English: a comparative study

MARC BRACHS RIBE

MBRACHS1@ALUMNO.UNED.ES

TUTOR ACADÉMICO: EVA ESTEBAS VILAPLANA

LÍNEA DE TFG: Línea de Fonética y fonología inglesas

FACULTAD DE FILOLOGÍA

CURSO ACADÉMICO: 2021-22- Convocatoria: Junio

Abstract

The pronunciation of English diphthongs varies across dialects. In Received Pronunciation, closing diphthongs are /eɪ/, /aɪ/, /əʊ/ and /aʊ/ whereas in Australian English rising diphthongs include /æɪ/, /ɑe/, /əʉ/ and /æɔ/, which are similar to Cockney and broad London. These accents from the south-east of England were the most predominant among the first Australian settlers, whose influence can be seen in the current production of diphthongs in Australian English.

A small corpus of 5 RP and 8 Australian English speakers has been created in this dissertation by asking them to utter 8 carrier sentences which included specific words containing the required diphthongs. The recordings have been segmented and the words with the diphthongs have been analysed using the program Praat.

The words analysed in this study exemplify how these diphthongs are produced differently in Received Pronunciation and in Australian English.

Keywords: Received Pronunciation, Australian English, diphthong, formant, spectrogram.

Index

1. Introduction

English is widely spoken all over the world and it is the first language for most of the countries that were under the British Empire, including the United States, South Africa or Australia among others. Since language is in constant evolution, English evolved differently in every place where the English settled.

Received Pronunciation, which is the English variety originally based in London and the South-east of England is to be compared with Australian English in this study. Although both varieties are geographically apart, they share a great deal of similarities, the most significant is being both accents non-rhotic. Nevertheless, the focus in this dissertation has been put on the vowel phonology, which is different in both accents. More specifically, four different diphthongs of RP have been compared with the corresponding diphthongs in Australian English.

As stated in Baugh & Cable (2013), the English entered late in the race for colonization compared with other European countries. It was when the Spanish and Portuguese started pouring wealth into their coffers that the ambition of the English grew. The English settlements of Jamestown and Plymouth were the starting of the process of colonization of North America where they controlled most of the territory after Wolfe's Victory (1759) over the French commanded by Montcalm.

In 1600, England established settlements in India, in Madras, Bombay and Calcutta and the East India Company was founded. Again, the French were the main rivals in the region, but a series of triumphs gave control of the country to the English and India became part of the British Empire.

Although the American Revolution deprived England of one of its most promising colonies at the second half of the eighteenth century, English remained the most spoken language in the region.

Also, in the second half of the eighteenth century, the Royal Society persuaded the king to sponsor an expedition to the Pacific to observe some astronomical patterns. After the observations had been completed, they explored the neighbourhood, sailed around New Zealand and finally reached Australia.

In 1770, Captain James cook sailed along the Australian east coast which he named New South Wales and was claimed for Great Britain. Later on, in 1788 a new penal colony was established at Sydney Cove, Port Jackson. Convicts were sent to Australia in order to work as labourers or as servants to the settlers.

It is probable that the first speakers of Australian English were the children born in the new colony, since adults have more well established social identity and are less likely to modify their speech. The most predominant accent was the south-eastern English and the innate tendency of humans to accommodate linguistically during interactions forged the new dialect probably the first 40 years after the settlement (Cox & Fletcher, 2012).

With an area of 7,617,930 square kilometers, Australia is a quite large country. Yet, Australian English is remarkably homogeneous. There are various theories explaining this homogeneity. One of them is that the same combination of input dialects was present in each major centre, thus, the same dialect developed independently in different places (Bernard 1981, Trudgill, 2004). Another theory is that Australian English was firstly spoken in Sydney and spread to other places through the mobility of the population mostly by sea (Bernard, 1969).

Nonetheless, existed accent variability in Australia, but rather being geographical, was more social and stylistic. As Mitchell & Delbridge (1965) described, this new dialect of English that had developed by the 1830s may have diverged between the 1850s and 1880s into a continuum where three main varieties could be identified, i.e. Cultivated, General and Broad. Cultivated shared many features with RP and was considered the most prestigious form. General Australian English, which shared some features with Cockney and popular London speech and finally the Broad Australian, which was mostly spoken in rural areas and was quite similar to General, but lengthened the first elements of the diphthongs.

In the late nineteenth century, RP was considered a symbol of prestige, the correct English pronunciation and was spoken by the social elite. Australian English fell short to these standards. Thus, Australians built up a social solidarity against their Britain-based officials and administrators, whether they were convicts or free migrants and the number of Cultivated accent speakers lowered as the young population preferred to avoid it (Wells b, 1982). In the second half of the twentieth century, this tendency continued and Cultivated Australian was soon to be supplanted by General Australian English, as the new internal model (Cox & Palethorpe, 2007).

Australian society has changed significantly in the last 50 years. Cox & Fletcher (2012) state that parallel developments between accent and sociopolitical events, such the rise of republican sentiment, increased immigration from non-English speaking countries and a new era embracing multiculturalism have affected the fabric of Australian society. Thus, the traditional division between Cultivated, General and Broad is no longer valid today.

Nowadays, the label “Australian English” should be considered a superordinate, which embraces three major dialect types, namely, Standard (Mainstream) Australian English, used by the majority and the Australian participants of this dissertation; Australian Aboriginal English, used by Indigenous Australians; and Ethnocultural Australian English, which are forms used to express non-mainstream or ethnic identity.

Currently, the relationships between England and Australia remain tight thanks to a shared heritage and common values. Their English varieties also remain similar and nearly undistinguishable for the non-native ear. This dissertation has the purpose of bringing to light the phonemic contrasts of four different diphthongs produced in both accents through a series of spectrogram analysis.

1.2 Received Pronunciation

At the beginning of the 20th century the term “Received Pronunciation” or “RP” was coined by Daniel Jones as representing standard spoken British English. By then, the BBC started broadcasting and its head, John Reith, who was concerned with prestige, set up a committee with phoneticians seeking a common denominator of educated speech (Cruttenden, 2014).

Although only 3 per cent of the population in Britain use an RP accent regularly, the BBC played a huge role spreading it and it is the most employed accent when it comes to teaching English as a foreign language.

Received Pronunciation is currently what anyone who lives in the United Kingdom hears daily from radio and television announcers, newsreaders and from many other public figures. Apart from that, Many English people is also regularly exposed to RP in face-to-face interactions. However, for just a small minority is their own speech (Wells a, 1982).

Although the origin of RP was in the south-east of England, it is now a regionless accent and it is likely to be encountered throughout the country. Received Pronunciation is more of a social accent and it is associated with the upper-middle and upper classes (Trudgill & Hannah, 2017).

Like all accents, RP is subject to change, Cruttenden (2014) and Wells a (1982) distinguish between “conservative” or “advanced RP”. Whereas conservative pronunciations are more typical of older speakers, advanced pronunciations are more typical of younger speakers.

Egalitarian ideals have led the BBC to give more voice to non-RP accents than they did forty years ago and younger generations might perceive Received Pronunciation as “posh” and “snobbish”. Thus, one of the features of advanced RP is the increase of T-glottaling, which is common in lower-class accents (Trudgill & Hannah, 2017).

1.3. Australian English

Throughout this dissertation the abbreviation of Australian English will be used when referring to Standard (Mainstream) Australian English which is the variety used in education, government, broadcasting, trade and the courts. It is the dominant dialect and used by the majority who are born in Australia or who immigrate at an early stage.

English language was established in Australia around the beginning of the nineteenth century and many different accents from the British Isles were present at the moment. Nevertheless, the south-eastern English dialects were the most abundant. Thus, Australian English share some similarities with the current popular London speech and the Cockney accent. Some of these characteristics are the Diphthong Shift and H Dropping (Wells b, 1982).

Australia was colonized considerably later than North America, so it is no surprising that Australian English is much more similar to the accents of present-day England than are those in North America (Wells b, 1982).

However, enough time has passed for Australia to develop its own accent, which is distinctive to any accent spoken in England. Some of these differences can be lexical, as the following examples taken from Trudgill & Hannah (2017).

Australian English	**English English**
- to barrack for	- to support
- trady	- skilled workman
- stoush	- fight
- goodday	- hello
- unit	- flat, apartment
- to retrench	- to fire, sack
- offsider	- partner, workmate

Regarding phonology, there is two main differences: Australian English avoids unstressed /ɪ/ in favour of word-internal /ə/ as in the words *horses*, *wanted* or *naked* (Wells b, 1982). The other main difference is that Australian English has /i:/ instead of /ɪ/ in *very*, *many*, etc. (Trudgill & Hannah, 2017).

When it comes to grammar there are just a few differences between RP and Australian English, one of them is the use of *have* when expressing possession. In Received Pronunciation is more used *I have a new car*, and in Australian English *I've got a new car* (Trudgill & Hannah, 2017).

The focus of this study has been put on in the phonetic differences that have RP and Australian English in their diphthongs. Some of the diphthongs are wider in Australian English than in RP. As in the closing diphthongs, where the difference between the open first element and closed second element is greater in Australian English than in RP. The diphthong is also slower in Australian English, with a longer first element and a higher probability of becoming monophtongized (Trudgill & Hannah, 2017).

1.4 Comparison between the two vocalic systems

There are considerable differences between the diphthongs of RP and Australian English as can be seen in the following figures.

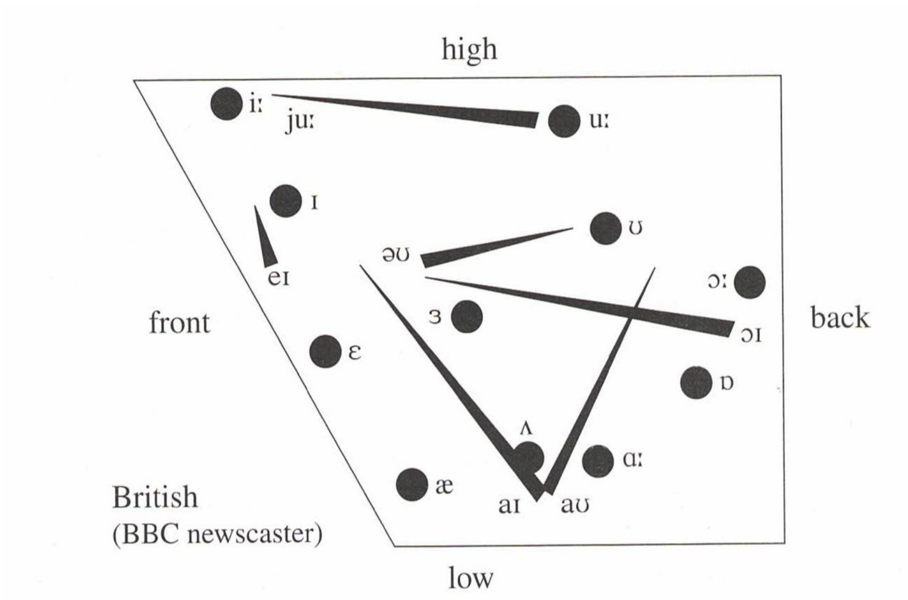

Figure 1: Diphthongs of RP (Ladefoged & Johnson, 2018)

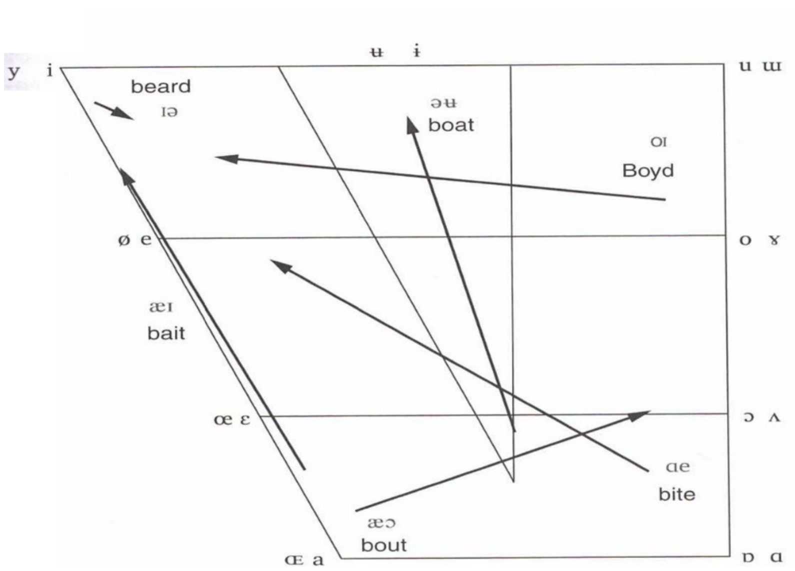

Figure 2: Diphthongs of Australian English (Cox & Fletcher 2012)

In this study, the following diphthongs are to be compared.

RP	Australian English
/eɪ/	/æɪ/
/aɪ/	/ɑe/
/əʊ/	/əʉ/
/aʊ/	/æɔ/

1.5. Objectives

The objectives of this study are as it follows:

- Record a small sample of Received Pronunciation and Australian English speakers.
- Exemplify how diphthongs are produced in Received Pronunciation and Australian English and analyse them.
- Determine if the phonemes produced by the participants correspond with their accents' patterns.

2. Methodology

For this study, a small corpus of recordings was gathered with the help of friends. RP1, RP2, RP3 and RP4 were recorded by some author's friends who work as English teachers in the province of *Lleida* and the speakers themselves are their work colleagues. Although they are all native speakers from England, they currently live in the province of *Lleida*. On the other hand, RP5's recording was sent straight to the author: He is a personal friend of him and has been living in the province of *Girona* for the last twenty years.

As for the AE speakers, most of the data was gathered by another friend, namely, speaker AE1 who was able to acquire recordings from his personal friends. These participants are AE2, AE3, AE4, AE5, AE7 and AE8. These informants were native Australian English speakers from different regions, where they are currently living. The exception is participant AE4, who was born in Greece, Europe, and brought to Australia at a very early stage. The participant AE6 was also born in Australia but he is currently living in the province of *Barcelona.* He was recorded by an author's friend.

All the participants were informed that their recordings would be used for a phonetic study. However, they were not informed of the main purpose of the study just in case they modified their pronunciation unconsciously. Their verbal consent can be found in their full recordings in point 6.3 of the annex.

2.1. Participants

5 RP and 8 Australian English speakers were recorded for this study. RP participants included 2 females and 3 males with ages ranging from 41 to 60. They were born in different places in England and 4 out 5 work or have worked as English teachers. English teachers have been recorded in order to obtain good examples of the RP accent.

A detailed list of the participants and further information can be found down below.

Participant RP1: Female, 60 years old, born in Stoke-on-trent, Staffordshire. She has a BA in English Literature (equivalent to *grado universitario*) and a PGCE in Education (equivalent to *posgrado*). She worked as an English teacher in Spain. She is currently retired.

Participant RP2: Female, 59 years old, born in Hillingdon, Greater London. She studied drama and German at the university. She is currently working as an English teacher in Spain.

Participant RP3: Male, 41 years old, born in Shoreham-by-sea, West Sussex. He moved to Spain when he was a teenager, where he finished *Bachillerato* and currently lives working as an English teacher.

Participant RP4: Male, 48 years old, born in Peterborough. He has a BA in Modern Languages and works as an English teacher in Spain.

Participant RP5: Male, 65 years old, born in Bideford, Devon. He completed the O level (equivalent to Spanish *ESO*) and works in the building site.

The 8 Australian English speakers included 3 females and 5 males with ages ranging from 21 to 68. They were all born in different places in Australia except for the participant AE4.

Participant AE1: Male, 45 years old, born in Bateman's Bay, New South Wales. He completed year 10 (equivalent to Spanish *ESO*). He takes random jobs and travels as much as he can.

Participant AE2: Male, 32 years old, born in Sydney, New South Wales. He completed year 10. He works as a bore driller in the mining industry.

Participant AE3: Female, 65 years old, born in Perth, Western Australia. She has a BA in Education. She has recently retired from being a teacher.

Participant AE4: Male, 68 years old, born in Greece, Europe. he has a High School Certificate (equivalent to Spanish *Bachillerato*). He is currently retired.

Participant AE5: Male, 44 years old, born in Melbourne, Victoria. He has a High School Certificate. He works as a drone pilot for the Australian Department of Home Affairs.

Participant AE6: Male, 25 years old, born in Tairee, New South Wales. He has a BA in Business.

Participant AE7: Female, 43 years old, born in Sydney, New South Wales. She has a High School Certificate. She is currently unemployed.

Participant AE8: Female, 21 years old, born in Perth, Western Australia. She has a High School Certificate. She is currently studying a degree.

2.2. Data

The participants were asked to utter 8 words containing a different diphthong. Words were included in a carrier sentence in order to avoid differences regarding patterns of intonation. The chosen sentence was "the word….is difficult to spell". The diphthongs to be analysed are the ones that appear in the following words.

	RP	**Australian English**
- Mate	/eɪ/	/æɪ/
- Face	/eɪ/	/æɪ/
- Light	/aɪ/	/ɑe/
- Hide	/aɪ/	/ɑe/
- Cold	/əʊ/	/əʉ/
- Boat	/əʊ/	/əʉ/
- Mouth	/aʊ/	/æɔ/
- Howl	/aʊ/	/æɔ/

2.3. Data recording and analysis

The required sentences were sent to the participants through either Facebook (Mark Zuckerberg et al., 2004) or Whatsapp (Acton & Koum, 2009). The sentences were read and recorded by the participants and sent back to the author of this dissertation as a voice message. Facebook Messenger voice messages were sent back as Mp4 files and the Whatsapp voice messages as Opus files. Afterwards, the files were converted into Wav format on the website of Convertio (www.convertio.co), which is the format that Praat (Boersma & Weenink, 2005) can read. Praat is a computer program employed for analysing, synthesizing and manipulating speech and other sounds. For this study, a spectrogram was created for each word, where only the diphthong is to be analysed, and the differences in their formants will be seen.

Formants are the resonances of the vocal tract. The first formant, which is the lowest, is called F_1 and deals with mouth opening. Vowels that require small mouth opening have low-frequency first formants. On the other hand, open vowels have higher frequency first formants. The second formant, labelled as F_2, deals with the changes in the size of the oral cavity. Tongue backing or lip rounding is going to lower the frequency of this formant because the constriction takes place in areas of high velocity. However, any tongue fronting or jaw movement that narrows the region of the oral cavity where the pressure is high, will raise the frequency of the formant (Borden & Harris, 1980).

In English, the quality of a vowel can change within a single syllable. Such vowel-glides are known as diphthongs (Ashby & Maidment, 2005). Vowels can have many formants. However, in this study the focus has been put on the first and second formant, along with their shifts, since are key to distinguishing the characteristics of each diphthong.

The graph below illustrates the RP diphthong frequencies measured in Hertz. The term Received Pronunciation has been substituted in the graph by General British (GB) since in Cruttenden, A. (2014) this term is considered outdated.

Table 5 Formant frequencies for GB diphthongs (in citation form).

	First component				Second component			
	F1		*F2*		*F1*		*F2*	
Diphthongs	*Male*	*Female*	*Male*	*Female*	*Male*	*Female*	*Male*	*Female*
/eɪ/	587	581	1,945	2,241	413	416	2,130	2,204
/aɪ/	734	822	1,117	1,275	439	359	2,058	2,591
/ɔɪ/	477	428	824	879	443	334	1,924	2,520
/əʊ/	537	545	1,266	1,573	379	380	1,024	1,267
/aʊ/	780	901	1,368	1,538	372	403	1,074	1,088
/ɪə/	382	399	2,096	2,514	578	417	1,643	1,846
/ʊə/	426	420	1,028	1,157	587	485	1,250	1,258

Figure 3: RP diphthong frequencies (Cruttenden, 2014)

The graph below illustrates the Australian English vowel and diphthong frequencies measured in Hertz and produced by males.

Males

Vowel	F1 (Hz)		F2 (Hz)		F1 (Hz)		F2 (Hz)		Duration (msec)
	Target 1				Target 2				
	mean	sd	mean	sd	mean	sd	mean	sd	
/iː/	304	32	2349	122					223
/ɪ/	360	28	2195	89					114
/e/	560	45	1956	108					135
/eː/	512	53	2007	110	499	40	1787	110	261
/æ/	762	57	1633	109					167
/ɐ/	711	42	1252	71					118
/ɐː/	728	44	1204	79					265
/ɔ/	592	47	944	57					134
/oː/	419	39	704	68					246
/ʊ/	383	46	901	90					124
/ʉː/	332	31	1656	107					229
/ɜː/	503	46	1468	61					250
/ɑe/	710	64	1088	106	498	44	1864	152	272
/æɪ/	683	71	1751	142	367	28	2226	100	246
/oɪ/	453	42	809	98	379	34	2068	117	237
/æɔ/	738	66	1682	139	589	82	1038	129	275
/əʉ/	597	56	1260	79	356	27	1664	93	246
/ɪə/	343	28	2214	110	410	40	1869	119	239

Figure 4: Australian English vowel and diphthong frequencies (Cox & Fletcher 2012).

The graph below illustrates the Australian English vowel and diphthong frequencies measured in Hertz and produced by females.

Females

Vowel	F1 [Hz]		F2 [Hz]		F1 [Hz]		F2 [Hz]		Duratio [msec]
	Target 1				Target 2				
	mean	sd	mean	sd	mean	sd	mean	sd	mean
/iː/	376	51	2900	120					244
/ɪ/	421	33	2675	127					122
/e/	657	94	2274	134					143
/eː/	645	97	2357	109	625	104	2130	88	273
/æ/	1005	88	1811	95					177
/ɐ/	924	89	1476	123					139
/ɐː/	937	91	1353	119					277
/ɔ/	724	90	1122	105					143
/oː/	462	51	837	68					259
/ʊ/	441	28	1012	125					141
/ʉː/	402	41	2011	156					236
/ɜː/	620	72	1785	87					264
/ɑe/	973	107	1375	105	563	95	2309	150	286
/æɪ/	809	99	2156	167	414	65	2716	143	259
/oɪ/	519	69	963	134	424	54	2537	155	257
/æɔ/	987	107	1872	98	670	125	1197	178	288
/əʉ/	712	95	1507	153	426	71	1983	198	259
/ɪə/	405	50	2712	119	463	70	2282	137	258

Figure 5: Australian English vowel and diphthong frequencies (Cox & Fletcher 2012).

3. Results

With the use of Praat, a spectrogram of each word containing the required diphthong of all participants has been created an analysed. Consequently, the formant frequencies of the glides have been compared between the speakers of Received Pronunciation and Australian English and discussed in the following subsections.

3.1 The diphthongs in the words “mate” and “face”

“Mate” is produced as /meɪt/ in RP and as /mæɪt/ in Australian English. “Face” is produced as /feɪs/ in RP and as /fæɪs/ in Australian English. The vowel glide is quite similar in both diphthongs. The first formant falls as the opening of the mouth reduces and the tongue rises. Whereas the second formant rises as the tongue moves to the front of the mouth and the lips spread.

Mate

The glide in /meɪt/ starts from slightly below the close-mid front position as the phoneme /e/ is uttered and moves to the position of /ɪ/, where a slight closing movement of the lower jaw takes place along with a spreading of the lips.

For /e/, the average starting point of the F_1 frequency for the male RP participants is located around 592 Hz and around 670 Hz for the female participants. This increase in frequency in the female participants is due to a smaller size of the vocal tract.

As the mouth opening reduces the F_1 falls to an average of 370 Hz for the male speakers and 361 Hz for the female participants.

The second formant begins with an average frequency of 1833 Hz for the male participants and increases up to 2017 Hz for the female speakers.

There is a rise of the F_2 as the tongue moves upwards and frontwards and the lips spread to an average of around 2244 Hz for the male speakers and 2839 Hz for the female participants. The movements of the formants can be seen in figure 6.

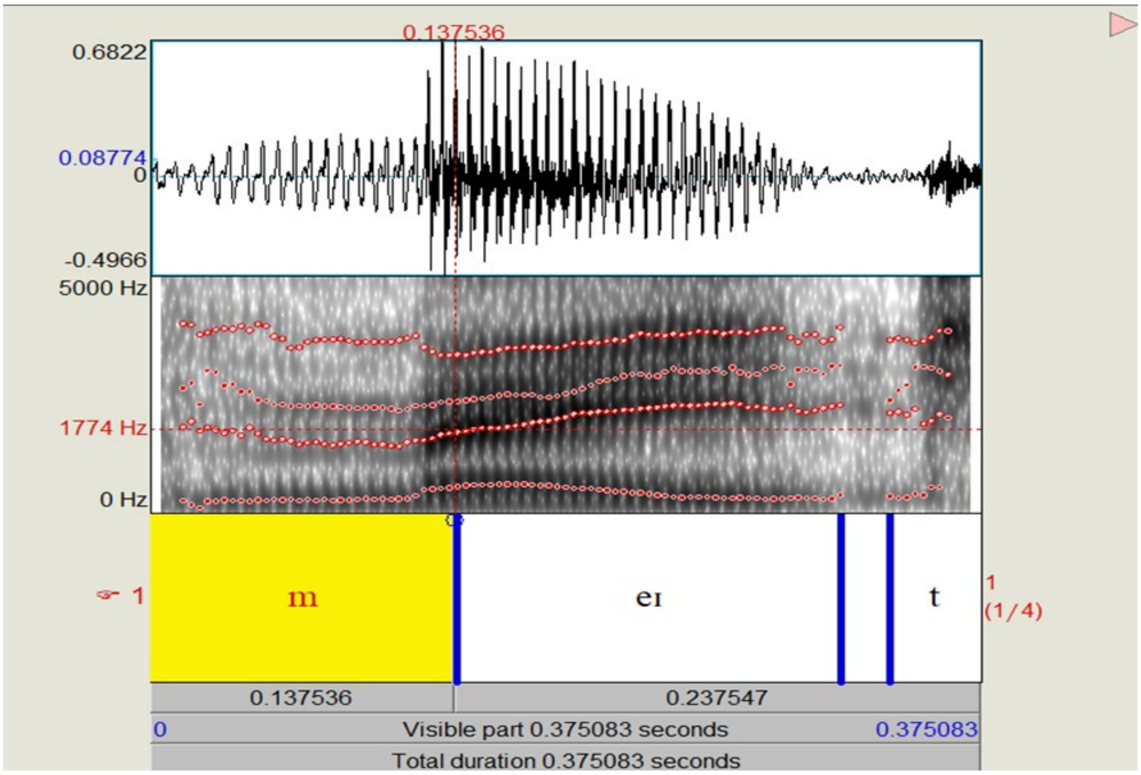

Figure 6: Waveform and spectrographic representation for the word "mate" produced by speaker RP5.

On the other hand, the glide in /mæɪt/ is more extensive. The starting point of the first vowel of the diphthong is between an open-mid and open position and slighter centred than /e/. The combination of a narrower pharyngeal cavity and a greater opening of the mouth results in a higher F_1 starting point frequency at an average of 724 Hz for the male participants and 814 Hz for the female speakers. As the mouth closes, the F_1 falls to an average of 376 Hz for the male speakers and around 352 Hz for the female speakers. The end point of the first formant is similar to the RP speakers since both diphthongs have /ɪ/ in the final position.

Due to a bigger opening of the mouth and a lowered and slighter centred position of the tongue when uttering /æ/, the frequency of the second formant is lower than when uttering /e/. The average frequency for the male speakers is located around 1474 Hz and around 1900 Hz for the female participants.

As the tongue rises and moves frontwards and the lips change from a neutral to a spread position, the second formant rises to an average frequency of around 2131 Hz for the male speakers and 2732 Hz for the female speakers. Again, this is a similar end point to the RP speakers.

The more extensive falls and rises of the first and second formants of the diphthong in the word “mate” when uttered by the Australian English speakers analysed in this dissertation can be seen in figure 7.

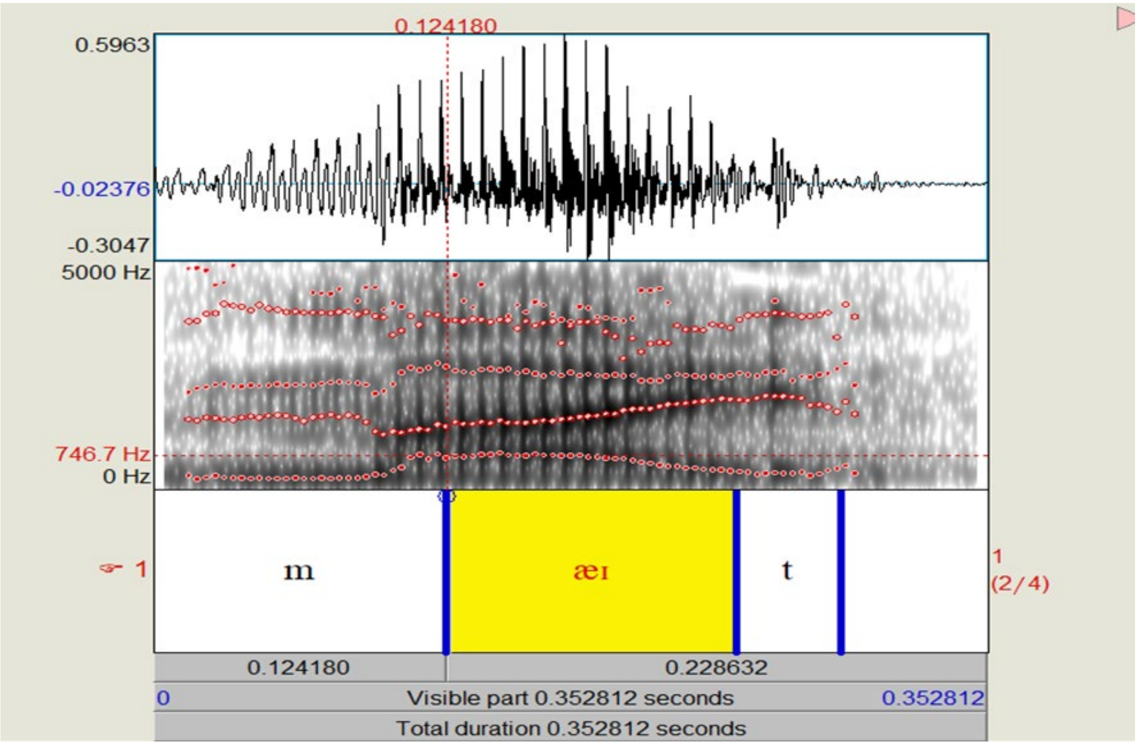

Figure 7: Waveform and spectrographic representation for the word “mate” produced by speaker AE5.

Face

The glide in /feɪs/ starts from slightly below the close-mid front position. However, due to /f/, it is somewhat higher than in /meɪt/. Then it moves up to the position of /ɪ/, with the closing of the lower jaw and the spreading of the lips.

For /e/, the average starting point of the F_1 frequency for the RP male speakers is around 515 Hz and 631 Hz for the female participants. The frequencies are similar to /meɪt/.

F_1 falls to an average of 361 Hz for the male participants and to 348 Hz for the female speakers as the mouth opening reduces.

The average starting point of F_2 is around 1705 Hz for the male speakers and 1812 Hz for the female participants. This is slightly lower than in /meɪt/, since when /f/ is produced, the tongue is in a lower position than in /m/.

The F_2 frequency increases up to 2150 Hz for the male speakers and to 2287 Hz for the female participants as the tongue rises and moves frontwards to the position of /ɪ/. This is somewhat lower than in /meɪt/, since the tongue does not move as frontwards and upwards as to reach the /t/ position. The movements of the formants can be seen in Figure 8.

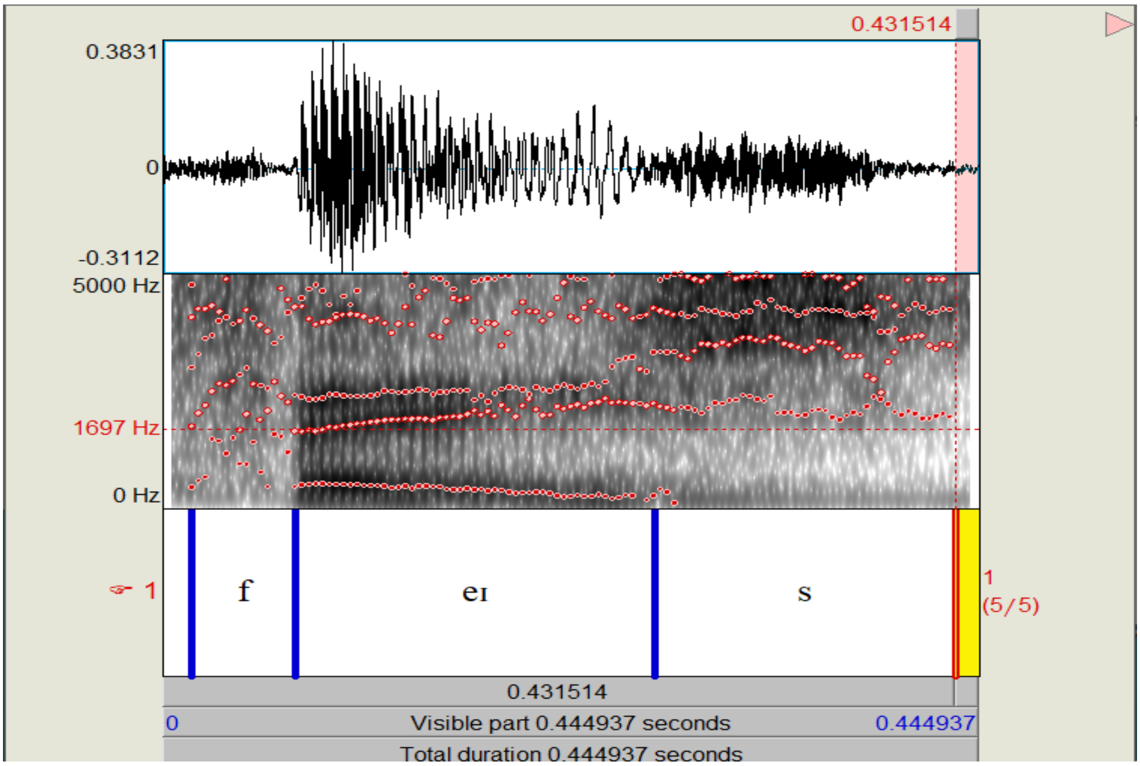

Figure 8: Waveform and spectrographic representation for the word "face" produced by speaker RP3.

On the other hand, the glide in /fæɪs/ is larger. The starting point of the first part of the diphthong is between an open-mid and open position and slighter centred than /e/. When Australian English speakers produce /æɪ/, the starting point of the F_1 frequency is around 633 Hz for the male participants and 643 Hz for the female participants. The frequency is higher than in the RP's /eɪ/ since the opening of the mouth is greater.

As the opening of the mouth reduces the formant of F_1 falls to an average of 319 Hz for the male participants and to 378 Hz for the female participants.

For /æ/, the average starting point of the F_2 frequency is around 1368 Hz for the male participants and 1798 Hz for the female speakers. This is lower than /e/ since the tongue is in a lower and slighter centred position.

As the tongue rises towards the front of the mouth and the lips spread the F_2 frequency increases up to an average of 2056 Hz for the male speakers and to 2295 Hz for the female participants.

The endpoints of both formants are similar in RP and in Australian English, since both diphthongs end with /ɪ/. The movements of the first and second formant for the production of /fæɪs/ can be seen in figure 9.

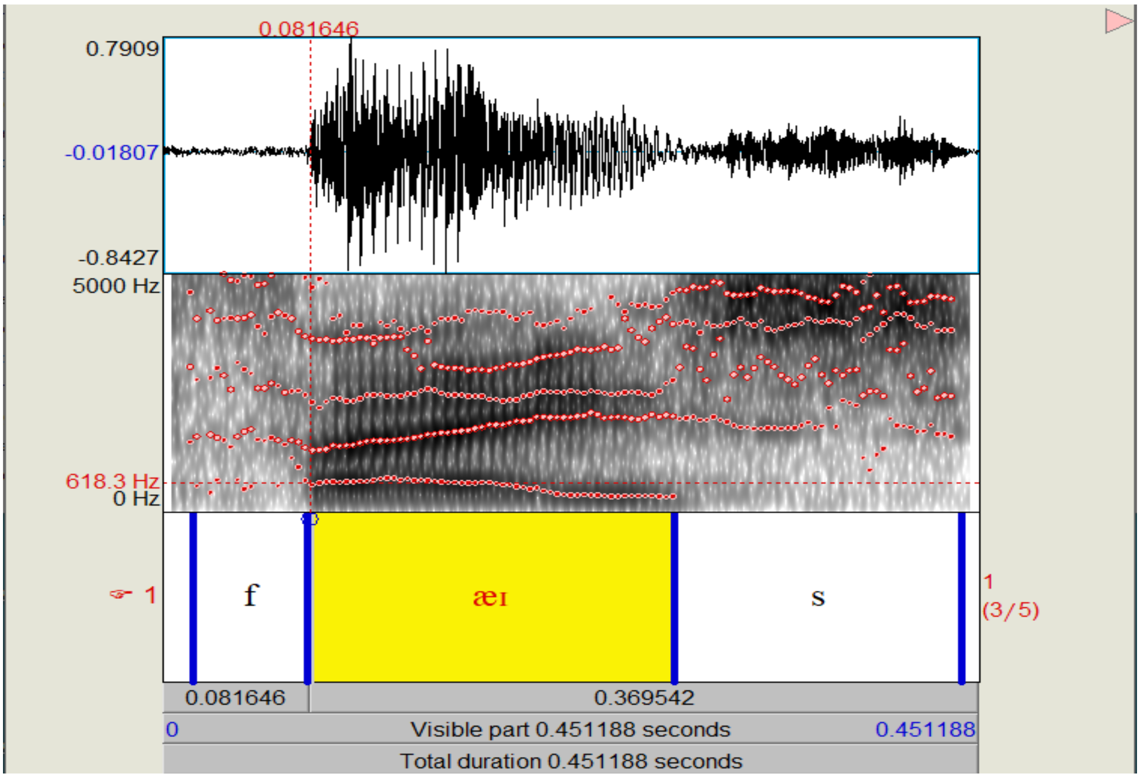

Figure 9: Waveform and spectrographic representation for the word "face" produced by speaker AE6.

3.2. The diphthongs in the words “light” and “hide”

“Light” is produced as /laɪt/ in RP and as /lɑet/ in Australian English. “Hide” is produced as /haɪd/ in RP and as /hɑed/ in Australian English. The glide in these two diphthongs is similar. The F_1 frequency falls as the mouth closes and the F_2 frequency rises as the tongue moves higher and frontwards.

Light

The glide in /laɪt/ begins at a point slightly behind the front open position as /a/ is produced and moves to the closer position of /ɪ/, where a substantial closing of the lower jaw takes place along with the spreading of the lips.

For /a/, the average starting point of the first formant for the RP male speakers is around 659 Hz and for the female speakers at 861 Hz. The frequency is higher in the female speakers due to a smaller size of the vocal tract.

Along with the closing of the mouth, F_1 falls to an average of 448 Hz for the male participants and 367 Hz for the female participants.

The F_2 frequency begins at an average point of 1367 Hz for the male speakers and 1330 Hz for the female participants.

As the tongue moves frontwards and the lips spread, the F_2 formant increases up to 2248 Hz for the male speakers and 2735 Hz for the female speakers. The movement of the formants in /laɪt/ can be seen in figure 10.

As to /lɑet/, the glide in the diphthong is less extensive. Although the starting point is still open, the tongue is nearly at the back position when /ɑ/ is produced. Either in /a/ or /ɑ/ the pharyngeal cavity is equally narrowed and the tongue lowered, which results in a similar F1 frequency. The starting average point for the male speakers is located around 618 Hz and 847 Hz for the female speakers. As the mouth closes the F_1 frequency falls to 541 Hz for the male participants and to 530 Hz for the female speakers. This involves a higher frequency than in /ɪ/, since the mouth does not close that much.

The average starting point of the second formant for the male participants is located around 1027 Hz. This is lower than /a/, since the tongue is nearly at the back position.

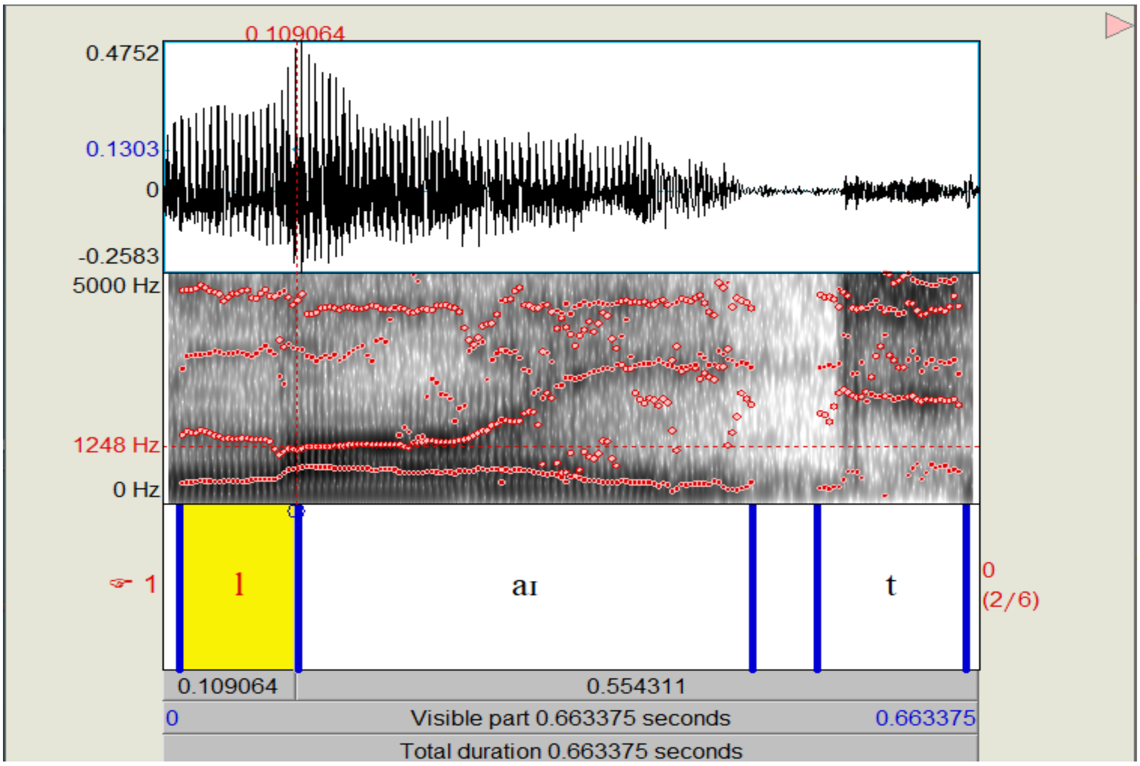

Figure 10: Waveform and spectrographic representation for the word "light" produced by speaker RP1.

In this case the female speakers examined in this dissertation have an average F_2 starting point of 1495 Hz, even higher than the RP female speakers. This means that the tongue is not as backed as when "light" is produced by the male Australian English speakers.

As the tongue moves upwards and frontwards, the second formant increases up to 1739 Hz for the male participants and 2308 Hz for the female participants. This is lower than the RP speakers, since the second vowel of the diphthong is /e/ and the tongue is slightly lower and the lips less spread than in /ɪ/. The lower increase of the F_2 formant can be seen in figure 11.

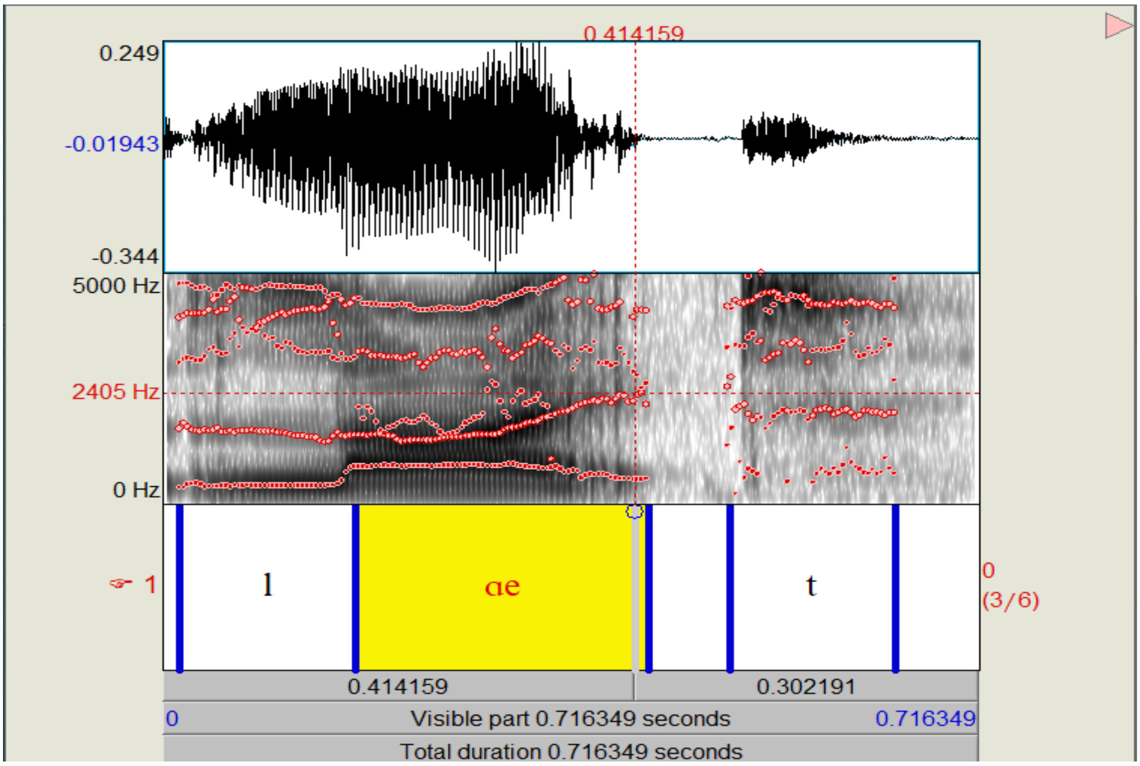

Figure 11: Waveform and spectrographic representation for the word "light" produced by speaker AE7.

Hide

The average starting point of the first formant for the male RP speakers when producing /haɪd/ is around 631 Hz and 834 Hz for the female participants. This is nearly the same as in /laɪt/. As the mouth opening reduces, the F_1 frequency drops to an average point of 448 Hz for the male speakers and 407 Hz for the female speakers. The fall is similar as in /laɪt/, since /t/ and /d/ are both alveolar and produced in the same place.

The average starting point of the second formant is located around 1183 Hz for the male speakers and 1315 Hz for the female speakers. This is quite similar to /laɪt/ for the female speakers. However, the frequency is somewhat lower in the male speakers and this is due to the fact that /h/ is produced with the tongue at the back position, whereas /l/ is placed in the front position. This affects the production of /a/.

As the tongue moves frontwards and upwards to the position of /ɪ/, the lips spread and the F2 frequency rises to an average of 2083 Hz for the male speakers and 2653 Hz for the female participant. Again, this is similar to /laɪt/, since /t/ and /d/ are both alveolar and the tongue is equally fronted. The trajectories of the formants in /haɪd/ can be seen in figure 12.

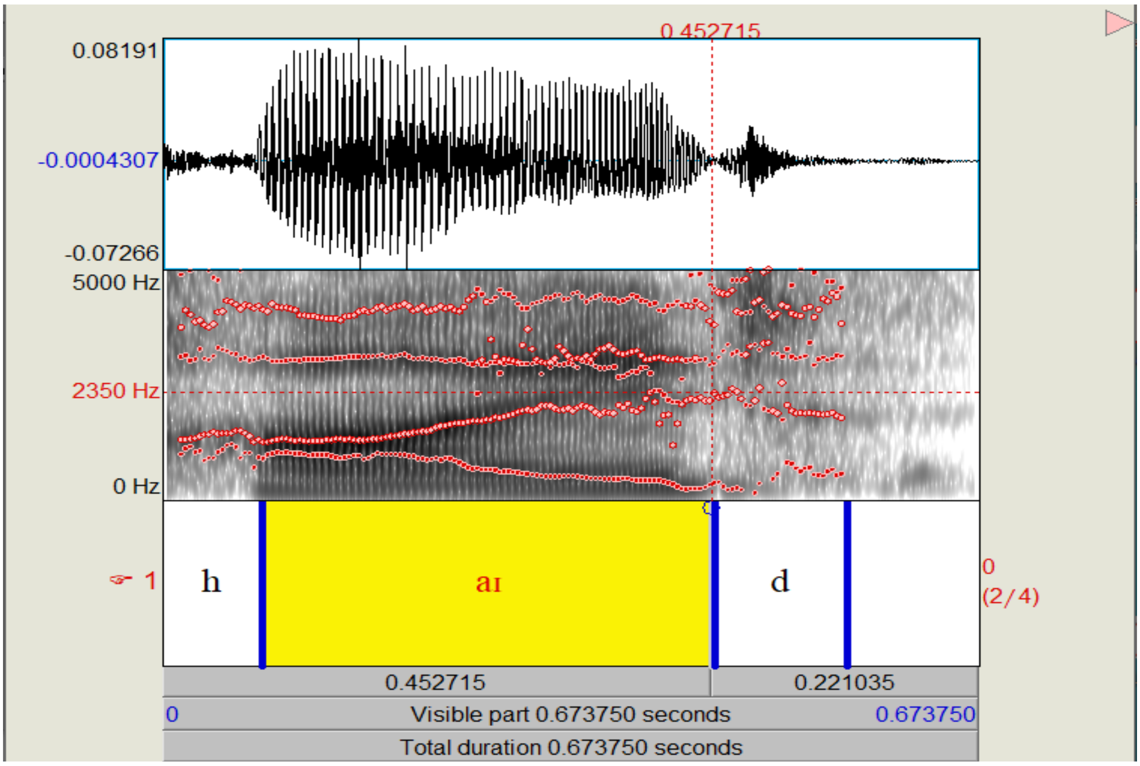

Figure 12: Waveform and spectrographic representation for the word “hide” produced by speaker RP2.

The average starting point of the F1 frequency for the male Australian English speakers when producing /hɑed/ is located around 663 Hz and 889 Hz for the female participants. This is similar to /haɪd/, since the opening of the mouth is alike.

The first formant falls as the mouth closes to an average of 530 Hz for the male speakers and 549 Hz for the female speakers. The F1 frequency does not go as low as in /aɪ/, since when /e/ is produced the mouth is not that closed as in /ɪ/.

The second formant starts at an average of 993 Hz for the male speakers and 1430 Hz for the female speakers. As in /lɑet/, the male Australian English speakers produce /hɑed/ with the tongue at the back position and this results in a lower frequency than in /haɪd/ when produced by the RP speakers. However, the female Australian English speakers have a higher frequency starting point. As in /lɑet/, the female Australian English speakers examined in this dissertation do not place the tongue as back as the male Australian English speakers and this results in a higher frequency, even higher than the female RP speakers.

As the tongue moves frontwards and upwards to the position of /e/, the lips spread and the second formant increases up to 1787 Hz for the male speakers and 2202 Hz for the female speakers. This is not as high as in /haɪd/, since the second vowel of the diphthong is /e/ and the tongue position is not that high. The movements of the formants in /hɑed/ can be seen in figure 13.

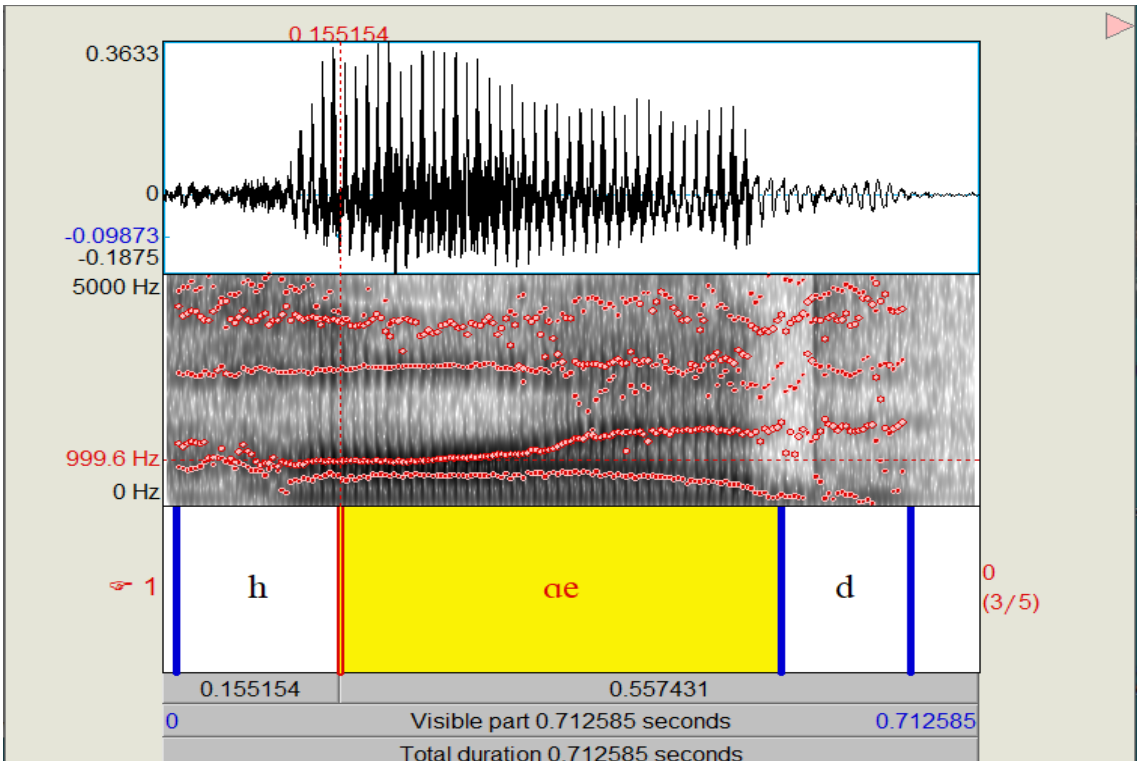

Figure 13: Waveform and spectrographic representation for the word “hide” produced by speaker AE1.

3.3. The diphthongs in the words “cold” and “boat”

“Cold” is produced as /kəʊld/ in RP and as /kəʉld/ in Australian English. “Boat” is produced as /bəʊt/ in RP and as /bəʉt/ in Australian English. In the case of “cold”, the diphthong is particularly similar in both accents. The first formant falls as the mouth closes and the second formant also falls as the tongue moves backwards both in RP and in Australian English. This diphthong is somewhat retracted in both accents due to the influence of dark [ɫ], which is the following allophone.

On the other hand, the glide differs significantly when “boat” is produced. The first formant falls in both accents as the mouth closes. However, the second formant remains in the same frequency in the case of RP and increases in the case of Australian English as the tongue moves upwards and frontwards to the position of /ʉ/.

Cold

The glide in /kəʊld/ begins at a central position, between close-mid and open-mid, and moves in the direction of /ʊ/ with a slight closing of the lower jaw. The lips move from a neutral position for the first element to a slightly round position on the second element.

For /ə/, the average starting point of the F_1 for the male RP speakers is located around 585 Hz and for the female participants around 627 Hz. As the mouth closes to the position of /ʊ/, the first formant falls to 356 Hz for the male speakers and to 379 Hz for the female participants.

The F_2 frequency starts at an average point of 1109 Hz for the male speakers and 1040 Hz for the female participants. As the tongue moves backwards, the second formant falls to 907 Hz for the male speakers and to 847 Hz for the female speakers. Speakers RP2 and RP3 merge /ʊ/ and dark [ɫ]. This feature can be seen in figure 14.

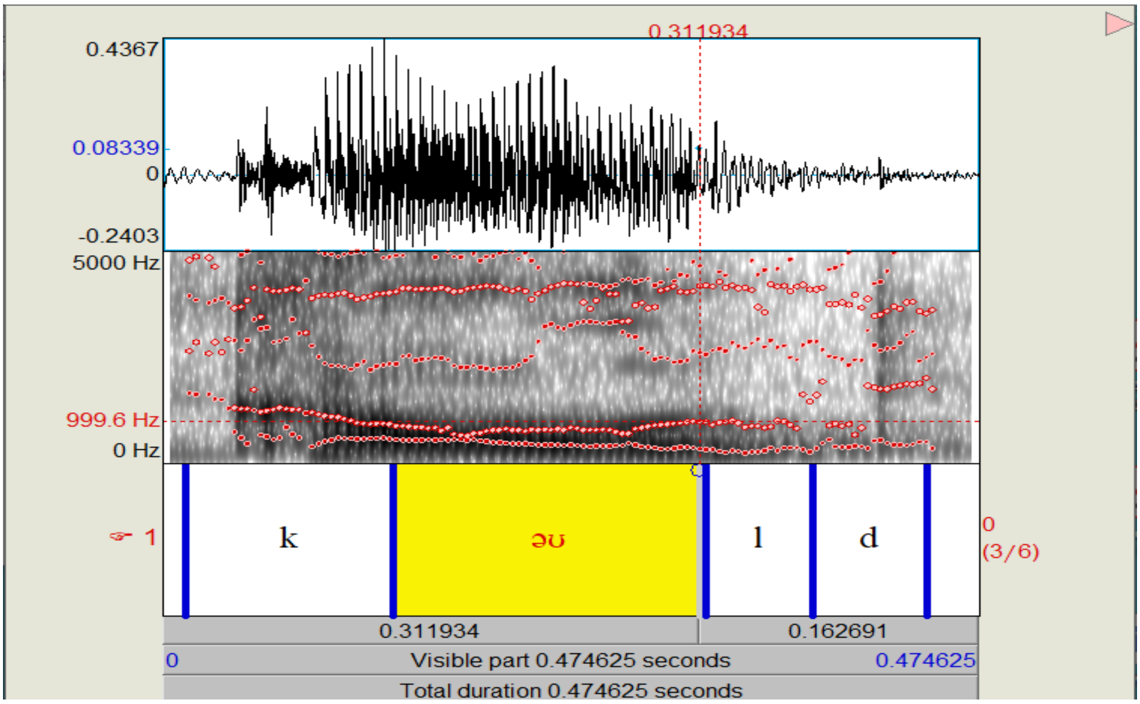

Figure 14: Waveform and spectrographic representation for the word "cold" produced by speaker RP3.

The Australian English speakers and the RP speakers examined in this dissertation utter the word "cold" in a similar way. However, in order to maintain community patterns, a broad transcription has been used and has been transcribed as /kəʉld/ when produced by the Australian English speakers.

The average starting point of the first formant for the Australian male speakers when /əʉ/ is produced is located around 613 Hz and 723 Hz for the female speakers. Slightly higher than in /kəʊld/, this means that the opening of the mouth is somewhat greater in the first element of the diphthong.

As the mouth closes, the F_1 frequency falls to 413 Hz for the male speakers and to 457 Hz for the female participants.

The F_2 frequency, starts at an average point of 1054 Hz for the male speakers and 1109 Hz for the female speakers. as the tongue moves backwards, the second formant falls to an average point of 930 Hz for the male speakers and 980 Hz for the female participants. Both the RP speakers and the Australian English speakers analysed in this study have an undistinguishable second formant when uttering "cold". As the RP speakers, some of the Australian English speakers merge the last element of the diphthong with the allophone dark [ɫ]. This is the case of the participants AE1, AE4, AE6 and AE7. An example of this feature can be seen in figure 15.

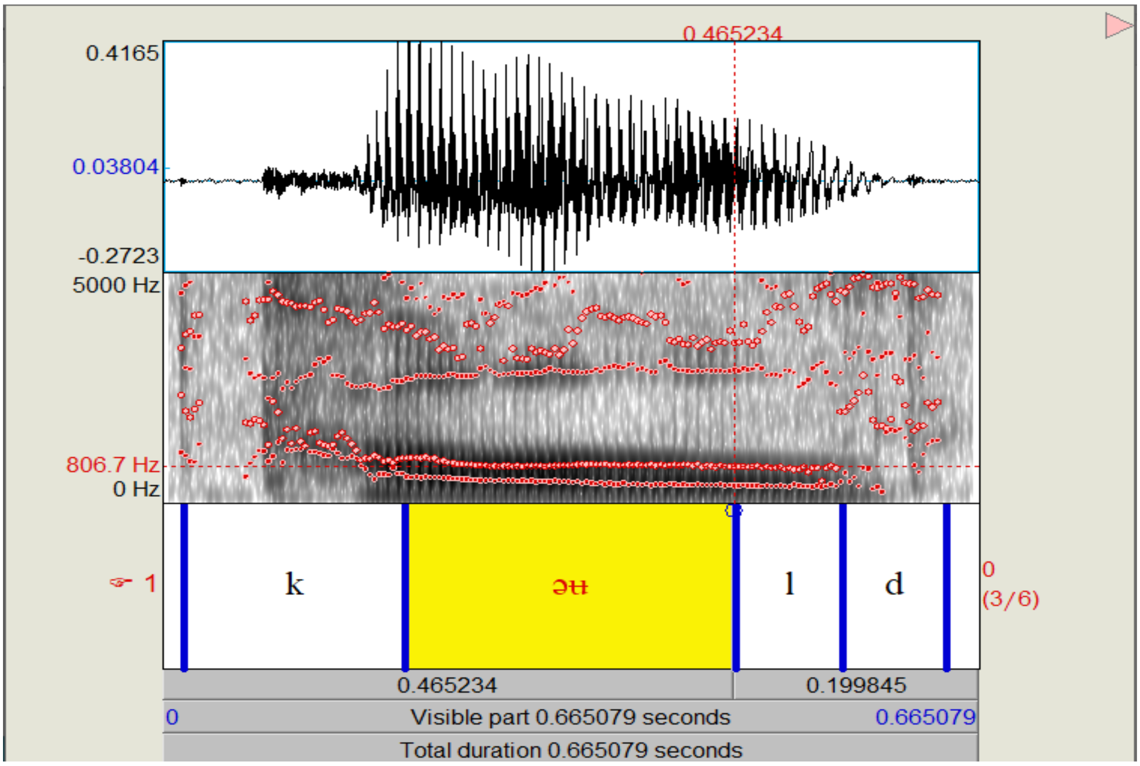

Figure 15: Waveform and spectrographic representation for the word "cold" produced by speaker AE1.

Boat

The glide in /bəʊt/ starts at a central position, between close-mid and open-mid, and moves in the direction of /ʊ/ with a slight closing of the lower jaw and a rounding of the lips.

The average starting point of the F_1 frequency for the male RP speakers is located at 540 Hz and for the female speakers at 641 Hz. As the mouth closes, the first formant falls to an average point of 365 Hz for the male speakers and 323 HZ for the female participants. Nearly the same as in /kəʊld/.

The average starting point of the second formant for the RP male speakers when /bəʊt/ is produced is around 1394 Hz and for the female speakers around 1564

Hz. These values are higher than those in /kəʊld/. This means that the phoneme /k/ has an influence over /ə/. Since the phoneme /k/ is velar, the back of the tongue rises towards the velum and it clearly affects the following phoneme with a lower frequency.

The trajectory of the second formant when /bəʊt/ is produced also differs from /kəʊld/. In /bəʊt/, without the influence of the dark [ɫ], the second formant remains in the same frequency. The average ending point for the male participants is around 1431 Hz and for the female participants around 1509 Hz. This means that the tongue barely moves during the glide. The trajectories of the first and second formants in the production of /bəʊt/ can be seen in figure 16.

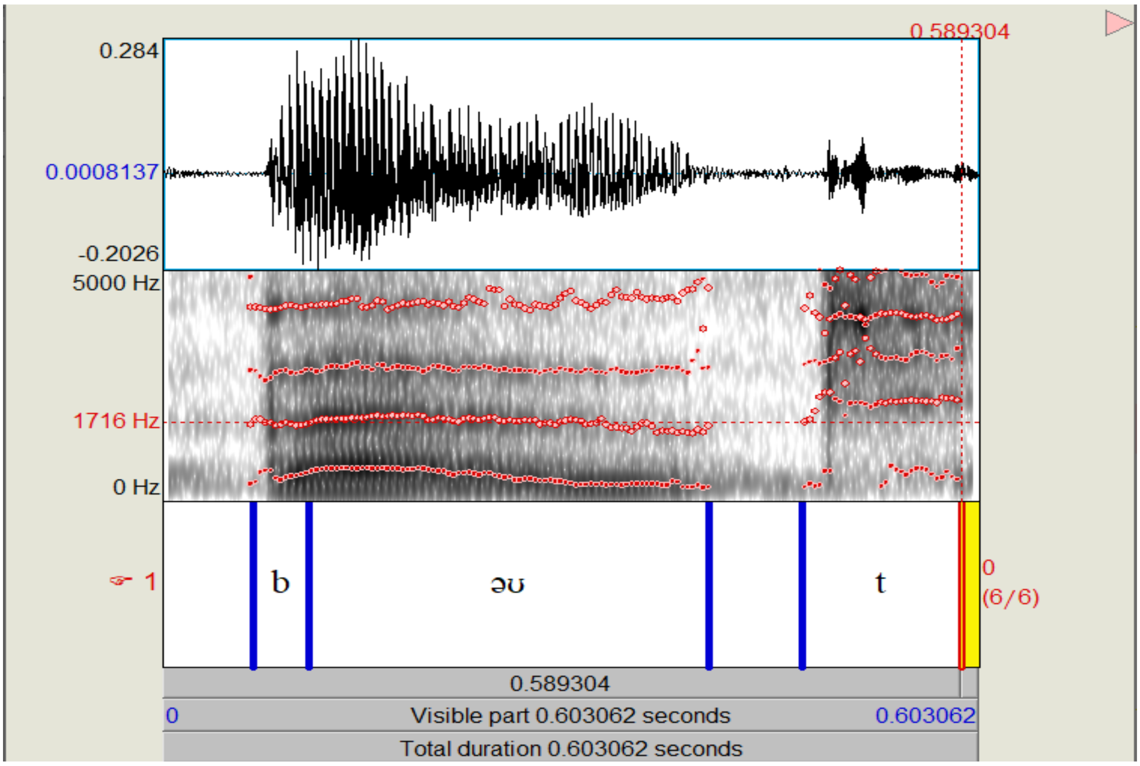

Figure 16: Waveform and spectrographic representation for the word "boat" produced by speaker RP1.

The glide in /bəʉt/ begins at a central position, between close-mid and open-mid, and moves in the direction of /ʉ/ with a slight closing of the lower jaw, a rounding of the lips and a movement of the tongue to a more fronted position than in /ʊ/.

The average starting point of the first formant for the male Australian English speakers is located around 579 Hz and for the female speakers around 661Hz.

As the opening of the mouth reduces, the F_1 frequency falls to an average point of 358 Hz for the male speakers and 413 Hz for the female speakers. These values are nearly the same as those in /bəʊt/.

On the other hand, the second formant of the glide in /bəʉt/ differs significantly from the second formant of the glide in /bəʊt/. In spite of having a similar starting point, the trajectory of the former reaches a higher frequency due to the fronting of the tongue.

The average starting point of the F_2 frequency is around 1143 Hz for the male Australian English participants and 1591 Hz for the female speakers. Then, it increases up to 1680 Hz for the male speakers and 1940 Hz for the female participants. The movements of the formants in the production of /bəʉt/ can be seen in figure 17.

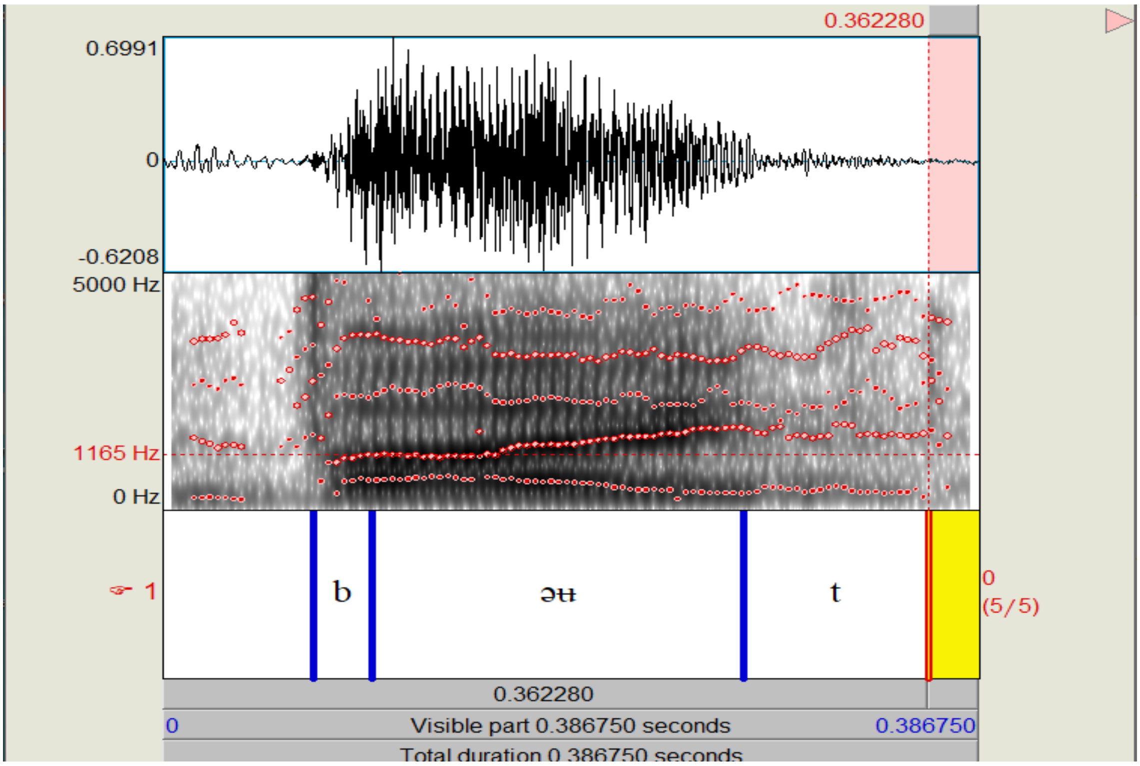

Figure 17: Waveform and spectrographic representation for the word “boat” produced by speaker AE6.

3.4. The diphthongs in the words "mouth" and "howl"

"Mouth" is produced as /maʊθ/ in RP and as /mæɔθ/ in Australian English. "Howl" is produced as /haʊl/ in RP and as /hæɔl/ in Australian English. The trajectories of the formants in these two diphthongs are similar. Both frequencies fall, except for /mæɔθ/. In this case, the first formant rises.

Mouth

The glide in /maʊθ/ begins at a central open position and moves in the direction of /ʊ/ with a closing movement of the lower jaw and a rounding of the lips.

For /a/, the average starting point of the F_1 frequency for the RP male speakers is around 678 Hz and 820 Hz for the female RP participants. This increase in frequency in the female speakers is due to a smaller size of the vocal tract.

As the mouth closes to the position of /ʊ/, the first formant falls to an average point of 531 Hz for the male speakers and 806 Hz for the female participants.

The second formant begins with an average frequency of 1569 Hz for the male speakers and 1495 Hz for the female participants. As the tongue rises and moves backwards and the lips are rounded, the F_2 frequency falls to an average point of 1367 Hz for the male participants and 1054Hz for the female speakers. The trajectories of the formants can be seen in figure 18.

On the other hand, the glide in /mæɔθ/ when produced by the Australian English speakers examined in this dissertation, differs significantly from the one in /maʊθ/. The former, starts at a point between the open mid and mid position, quite close to the mid position of /ɛ/ and moves in the direction of /ɔ/, which is a back open mid position.

For /æ/, the average starting point of the first formant is located at 574 Hz for the male speakers and 724 Hz for the female participants. These values are lower than those in /a/, since the mouth is not that open.

As the mouth opens to the position of /ɔ/, the first formant rises to an average point of 674 Hz for the male speakers and 870 Hz for the female participants. These values are higher than in /ʊ/, since the mouth is not that close.

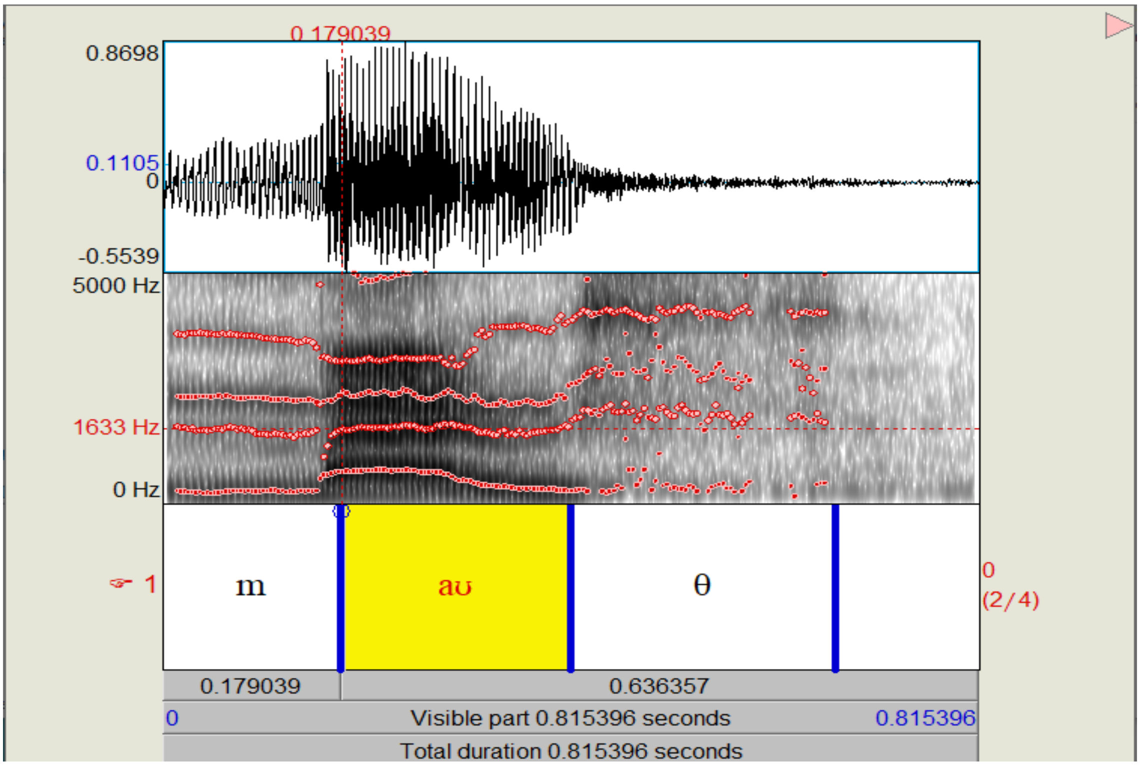

Figure 18: Waveform and spectrographic representation for the word "mouth" produced by speaker RP5.

The frequency of the second formant in the diphthong /æɔ/ has an average starting point of 1804 Hz for the male participants and 1991 Hz for the female participants. These values are higher than in /a/ since the tongue is more raised and fronted.

As the tongue moves backwards, the F_2 frequency falls to an average point of 1208 Hz for the male participants and 1468 Hz for the female participants. The decrease of the second formant is more extensive than in /aʊ/, since the tongue makes a longer movement from the front to the back of the mouth. The trajectories of the formants can be seen in figure 19.

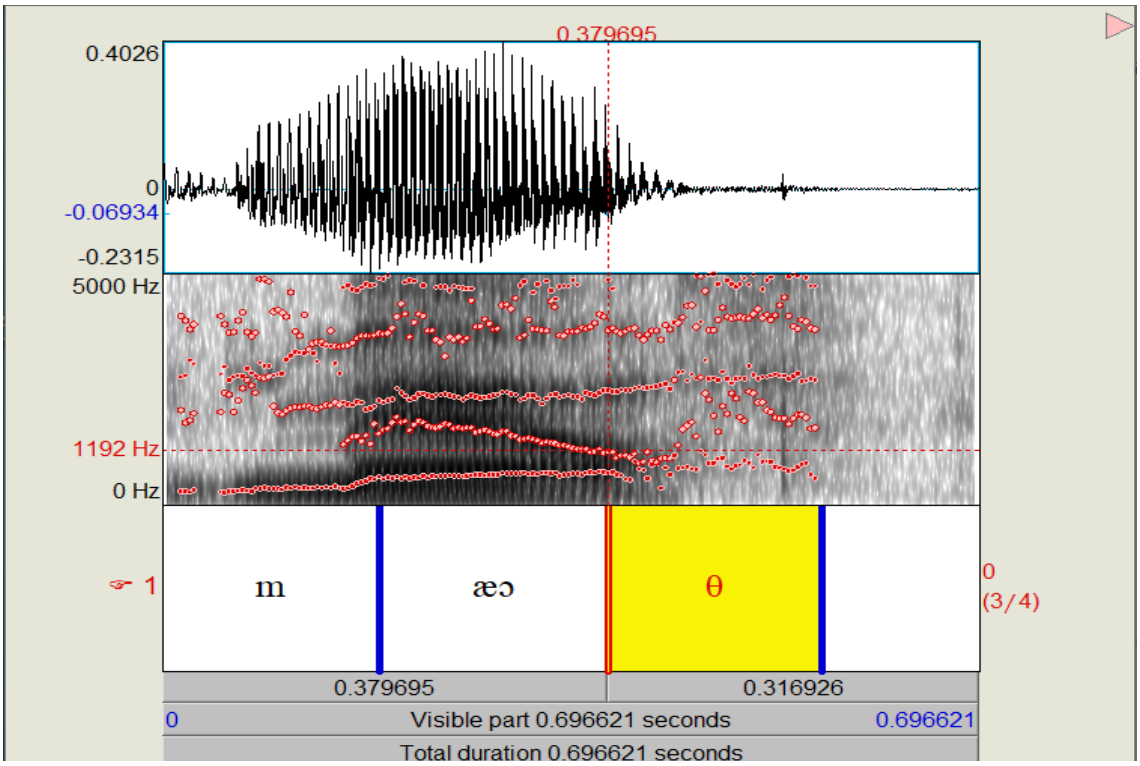

Figure 19: Waveform and spectrographic representation for the word "mouth" produced by speaker AE4.

Howl

The glide in /haʊl/ starts at a point between the back and front open positions and moves towards /ʊ/ with a closing movement of the lower jaw and a rounding of the lips.

The average starting point of the F_1 frequency for the male RP speakers is located at 733 Hz and 999 Hz for the female speakers. As the mouth closes, the first formant falls to an average point of 595 Hz for the male speakers and 517 Hz for the female participants.

The average starting point of the second formant for the male speakers is around 1403 Hz and for the female speakers 1509 Hz. As the tongue moves backwards to the position of /ʊ/, the second formant falls to an average point of 999 Hz for the male speakers and 944 Hz for the female speakers. These end point values in the F_2 frequency of the second element of the diphthong are lower than those in /maʊθ/. This is due to the influence of the following allophone, which is dark [ɫ]. The movements of the formants can be seen in figure 20.

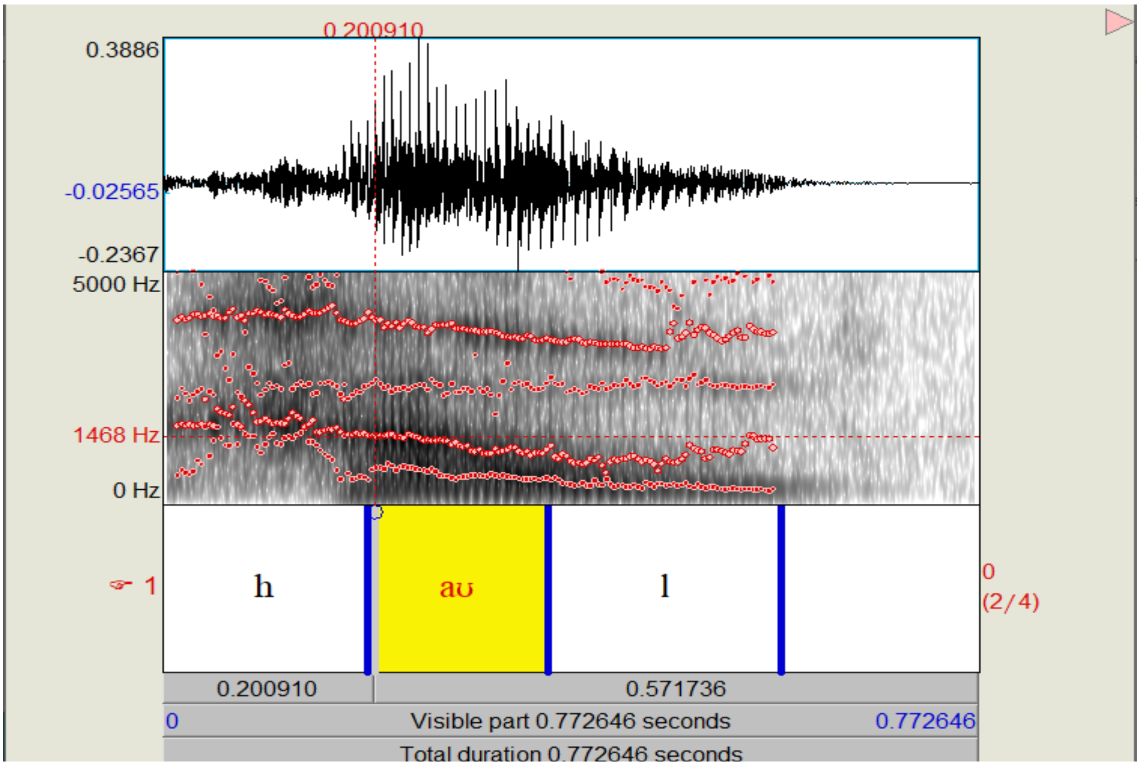

Figure 20: Waveform and spectrographic representation for the word "howl" produced by speaker RP4.

The glide in /hæɔl/ begins at a front near open position and moves towards /ɔ/, a back open mid position. In this diphthong there is a movement of the tongue from the front to the back of the mouth along with a rounding of the lips.

For /æ/, the average starting point of the F_1 frequency for the male Australian English participants is located around 657 Hz and for the female speakers around 972 Hz. These values are lower than those in /haʊl/, since the mouth is not that open.

As the glide moves to the position of /ɔ/, the mouth closes slightly. Thus, the first formant also falls slightly to an average point of 607 Hz for the male participants and 837 Hz for the female speakers. The frequencies are higher than in /haʊl/, since the mouth does not close that much.

The average starting point of the second formant for the male speakers is located at 1782 Hz and for the female speakers at 1890Hz. These values are higher than those in /haʊl/, since the tongue is more fronted.

As the tongue moves backwards, the F_2 frequency falls to an average point of 1010 Hz for the male participants and 1348 Hz for the female speakers. These values are similar to those in /haʊl/ for the male speakers. However, they are significantly higher for the female speakers. This means that when /ɔ /is produced by the female English Australian speakers examined in this dissertation the tongue is not as backed as when /ʊ/ is produced by the female RP participants. The movements of the formants can be seen in figure 21.

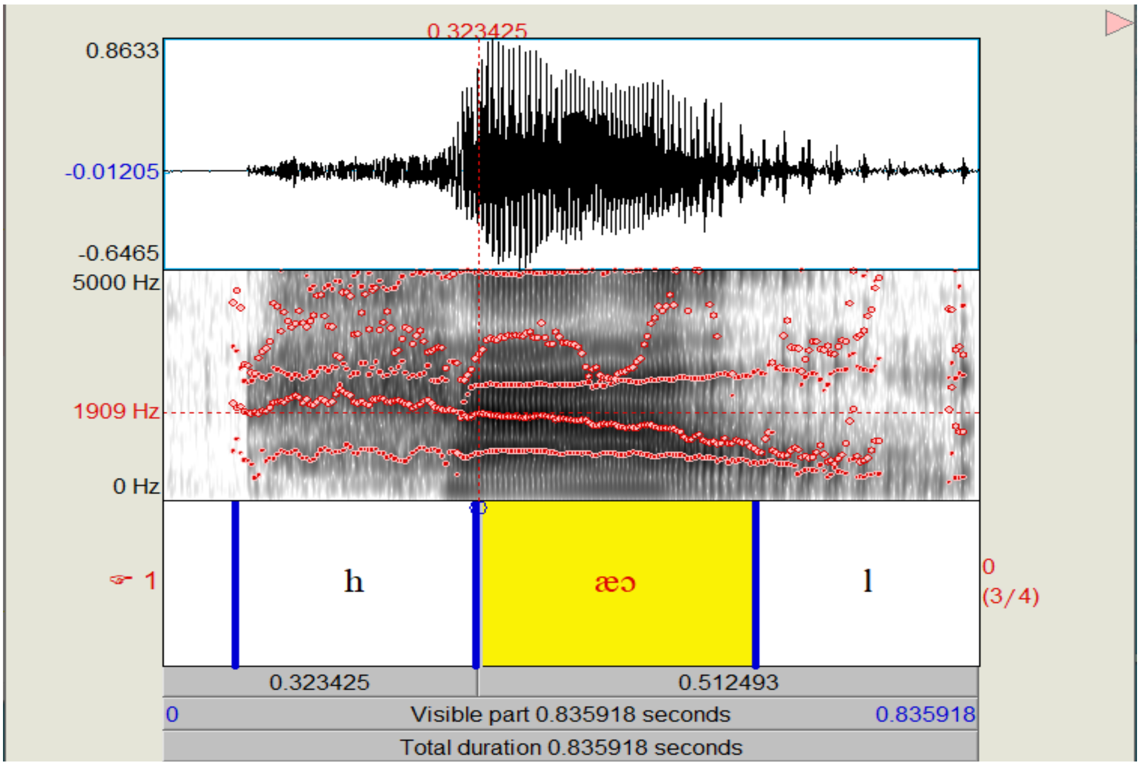

Figure 21: Waveform and spectrographic representation for the word "howl" produced by speaker SAE3.

4. Conclusions

In this dissertation, 5 Received Pronunciation speakers from different regions of England and 8 Australian English speakers from different regions of Australia have produced 8 words with a closing diphthong embedded in a carrier sentence. The productions were examined using the computer program Praat. Only four different diphthongs were finally analysed due to the limitation on the extension of this study.

The diphthongs /eɪ/, /aɪ/, /əʊ/ and /aʊ/, produced by the Received pronunciation participants, were compared with the corresponding /æɪ/, /ɑe/, /əʉ/ and /æɔ/ produced by the Australian English participants examined in this study. For the comparison, a spectrogram of each of these specific words was created and the formant frequencies of the diphthongs were analysed.

The purpose of this dissertation was to highlight the differences in the pronunciation of these diphthongs between the two accents. Although only 13 participants were analysed and some of the values do not match perfectly with those in the data section grids, the differences in the average frequencies illustrate the differences in pronunciation between the two accents when the diphthongs are produced. The fact that the frequencies do not match perfectly with the frequencies in the figures 3, 4 and 5 is associated to connected speech and the influence of the previous and following phonemes. There is only one exception, which is the diphthong in the word “cold”. The values in the frequencies show that the participants of both accents examined in this study pronounce the word “cold” in a very similar way.

It is also worth pointing out that even though Australia is a significantly large country and the participants are from distant places and different backgrounds, their English is quite homogenous and shares similarities with Cockney and contemporary broad London, as stated in J.C.Wells b (1982). The south-eastern English accent was the most predominant with the first settlers and, as described in Cruttenden (2014), these similarities with Australian English can be seen in the diphthongs produced by the Australian English speakers examined in this study.

5. References

- Acton, B. & Koum, J. 2009. Whatsapp. Available at https://www.whatsapp.com
- Ashby, M. & Maidment, J. 2005. *Introducing Phonetic Science*. Cambridge: Cambridge University Press.
- Baugh, A & Cable, T. 2013. *A History of the English Language*. London: Routlegde
- Bernard, J. 1969. On the uniformity of spoken Australian English, *Orbis* Vol. 18, pp. 62-73.
- Bernard, J. 1981. Australian pronunciation. In Delbridge, A (ed.), *The Macquarie Dictionary*, pp.18-27, North Ryde
- Boersma, P. & Weenink, D. 2005. Praat: doing phonetics by computer [Computer program]. Available at http://www.praat.org/
- Borden, G., Lawrence, R. & Harris, K. 1980. *Speech Science Primer*. Baltimore: Lippincott Williams & Wilkins. p. 87.
- Convertio [File Converter], available at https://convertio.co/es/
- Cox, F. & Fletcher, J. 2012. Introduction to Australian English speech production. In Cambridge University Press. *Australian English Pronunciation and Transcription*, pp. 13-20, Cambridge.
- Cox, F. & Palenthorpe, S. 2007. Illustration of the Ipa: Australian English, *Journal of the International Phonetic Association.* Vol. 37, pp. 341-50. DOI: https://doi.org/10.1017/S0025100307003192
- Cruttenden, A. 2014. *Gimson's Pronunciation of English.* Abingdon: Routlegdge
- Mark Zuckerberg et al. 2004. Facebook. Available at https://www.facebook.com
- Mitchell, A.G. & Delbridge, A. 1965. *The speech of Australian Adolescents*. Sydney: Magnus & Robertson.
- Trudgill, P. & Hannah, J. 2017. English, Australasian, South African and Welsh English. In Routledge. *International English*, pp. 14-31, Abingdon
- Trudgill, P. 2004. *New-Dialect Formation: The Inevitability of Colonial Englishes*. New York: Oxford University Press.

- Wells, J. C. 1982a. *Accents of English 2*. Cambridge: Cambridge University Press.
- Wells, J. C. 1982b. The Southern Hemisphere. In Cambridge University Press. *Accents of English 3*, pp. 592-604, Cambridge.

6. Annexes

6.1. Diphthong frequencies

Mate	F1 (Hz)	F2 (Hz)	F1 (Hz)	F2 (Hz)
	Target 1		Target 2	
RP1	669	2107	361	2929
RP2	670	1928	361	2749
RP3	541	1902	412	2390
RP4	618	1825	387	2082
RP5	618	1774	310	2261
AE1	709	1675	274	2205
AE2	700	1457	450	1824
AE3	772	1870	335	2440
AE4	772	1282	387	2339
AE5	746	1466	438	2030
AE6	695	1491	335	2261
AE7	849	1851	361	2878
AE8	823	1979	361	2800

Face	F1 (Hz)	F2 (Hz)	F1 (Hz)	F2 (Hz)
	Target 1		Target 2	
RP1	644	1928	335	2441
RP2	618	1697	361	2133
RP3	515	1697	310	2261
RP4	515	1645	412	1979
RP5	515	1774	361	2210
AE1	695	1440	258	2236
AE2	669	1440	310	1953
AE3	541	1722	361	2261
AE4	618	1260	335	2133
AE5	567	1414	361	1928
AE6	618	1286	335	2030
AE7	644	1744	361	2441
AE8	746	1928	412	2184

Light	F1 (Hz)	F2 (Hz)	F1 (Hz)	F2 (Hz)
	Target 1		Target 2	
RP1	779	1248	421	3066
RP2	944	1413	310	2405
RP3	641	1385	476	2074
RP4	613	1413	448	2405
RP5	724	1303	421	2267
AE1	669	1055	476	1633
AE2	613	999	558	1633
AE3				
AE4	613	972	586	1881
AE5	613	1082	558	1774
AE6	586	1027	531	1744
AE7	834	1523	558	2405
AE8	861	1468	503	2212

Hide	F1 (Hz)	F2 (Hz)	F1 (Hz)	F2 (Hz)
	Target 1		Target 2	
RP1	724	1303	421	3011
RP2	944	1330	393	2295
RP3	558	1248	503	1909
RP4	641	1055	476	2019
RP5	696	1248	365	2322
AE1	696	999	503	1661
AE2	586	861	586	1661
AE3	972	1303	558	2239
AE4	669	972	503	2047
AE5	669	1110	476	1826
AE6	696	1027	586	1744
AE7	724	1575	586	2212
AE8	972	1413	503	2157

Cold	F1 (Hz)	F2 (Hz)	F1 (Hz)	F2 (Hz)
	Target 1		Target 2	
RP1	613	944	365	751
RP2	641	1137	393	944
RP3	558	889	338	999
RP4	558	1110	365	889
RP5	641	1330	365	834
AE1	586	1027	393	806
AE2				
AE3	696	1027	393	751
AE4	613	999	393	944
AE5	641	1110	393	1027
AE6	613	1082	476	944
AE7	696	999	476	916
AE8	779	1303	503	1275

Boat	F1 (Hz)	F2 (Hz)	F1 (Hz)	F2 (Hz)
	Target 1		Target 2	
RP1	641	1716	282	1551
RP2	641	1413	365	1468
RP3	503	1275	393	1248
RP4	531	1496	421	1468
RP5	586	1413	282	1578
AE1	592	1183	335	1697
AE2	576	1114	320	1676
AE3	704	1574	422	1625
AE4	550	1114	397	1676
AE5	576	1139	371	1651
AE6	602	1165	371	1702
AE7	678	1625	422	1932
AE8	602	1574	397	2265

Mouth	F1 (Hz)	F2 (Hz)	F1 (Hz)	F2 (Hz)
	Target 1		Target 2	
RP1	669	1358	531	1165
RP2	972	1633	1082	1358
RP3	669	1523	669	1192
RP4	641	1551	586	1303
RP5	724	1633	338	1468
AE1	641	1909	669	1192
AE2	558	1771	696	1248
AE3	669	1881	999	1440
AE4	613	1744	724	1192
AE5	503	1826	724	1137
AE6	558	1771	558	1275
AE7	779	2019	861	1496
AE8	724	2074	751	1468

Howl	F1 (Hz)	F2 (Hz)	F1 (Hz)	F2 (Hz)
	Target 1		Target 2	
RP1	861	1358	448	861
RP2	1137	1661	586	1027
RP3	669	1275	669	1027
RP4	779	1468	586	1137
RP5	751	1468	531	834
AE1	641	1771	669	1192
AE2	696	1826	613	972
AE3	1055	1909	843	1220
AE4	669	1854	558	861
AE5	613	1826	586	1027
AE6	669	1633	613	999
AE7	751	1936	889	1468
AE8	1110	1826	779	1358

6.2. Other spectra

Mate

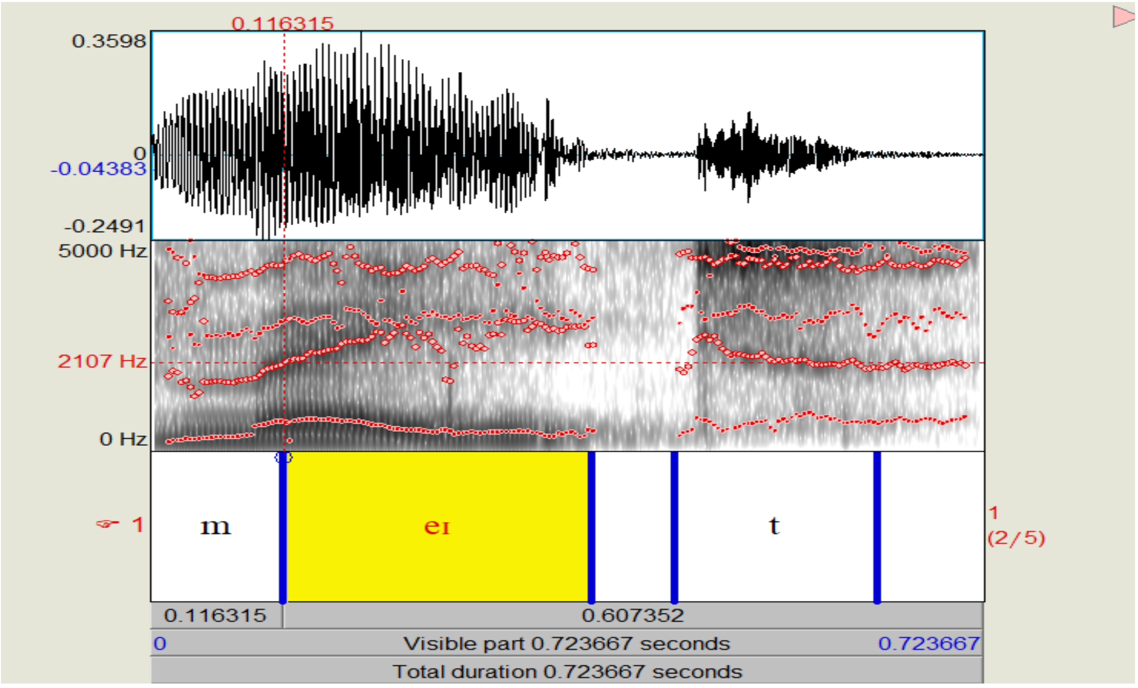

Figure 22: Waveform and spectrographic representation for the word “mate” produced by speaker RP1.

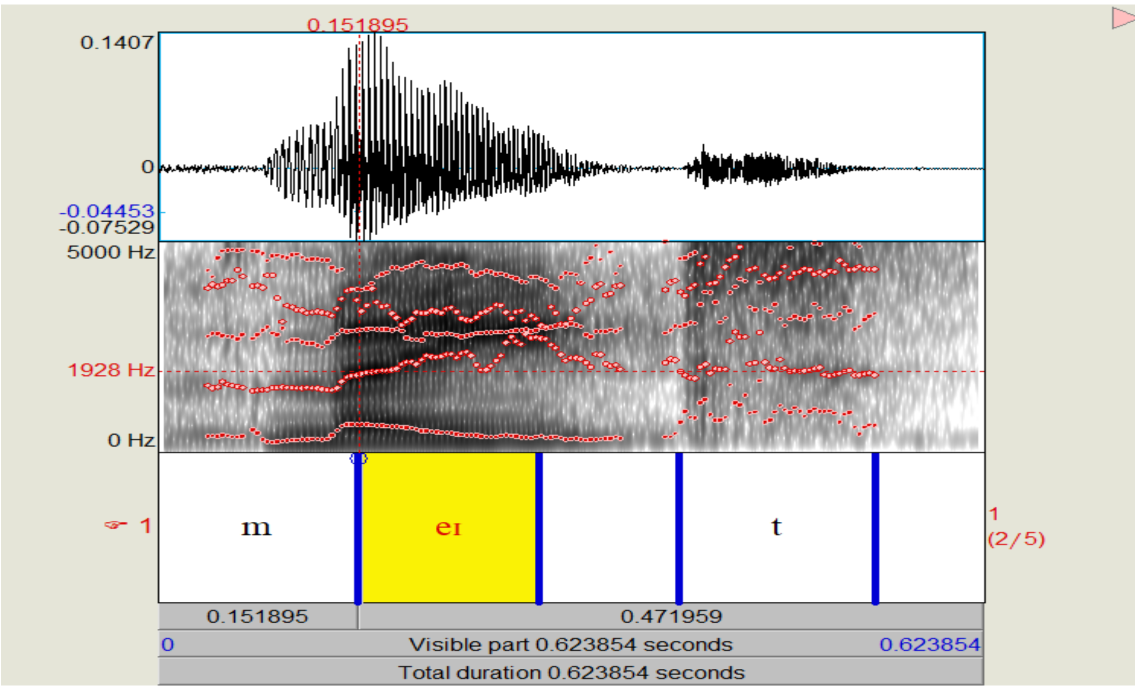

Figure 23: Waveform and spectrographic representation for the word “mate” produced by speaker RP2.

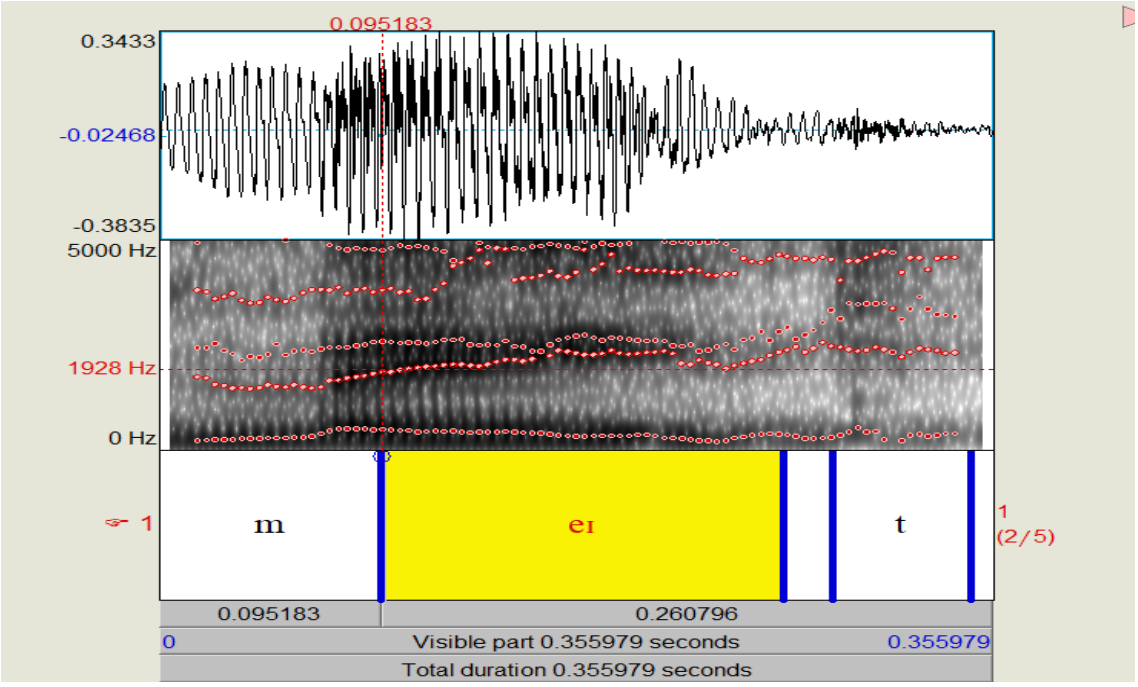

Figure 24: Waveform and spectrographic representation for the word “mate” produced by speaker RP3.

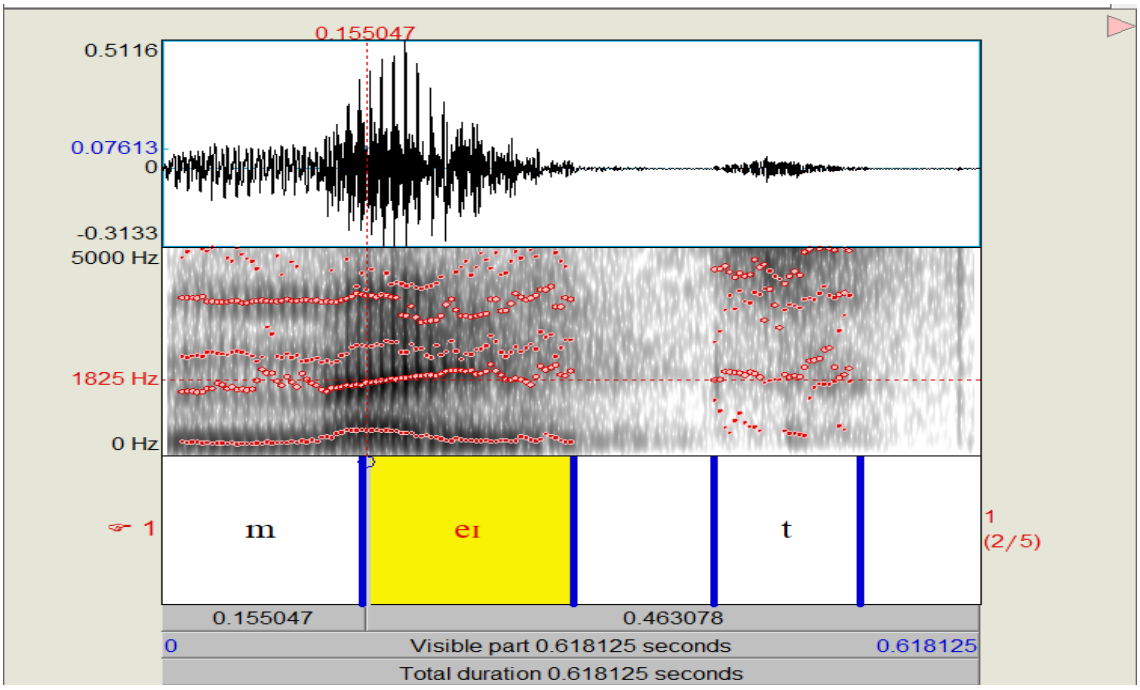

Figure 25: Waveform and spectrographic representation for the word “mate” produced by speaker RP4.

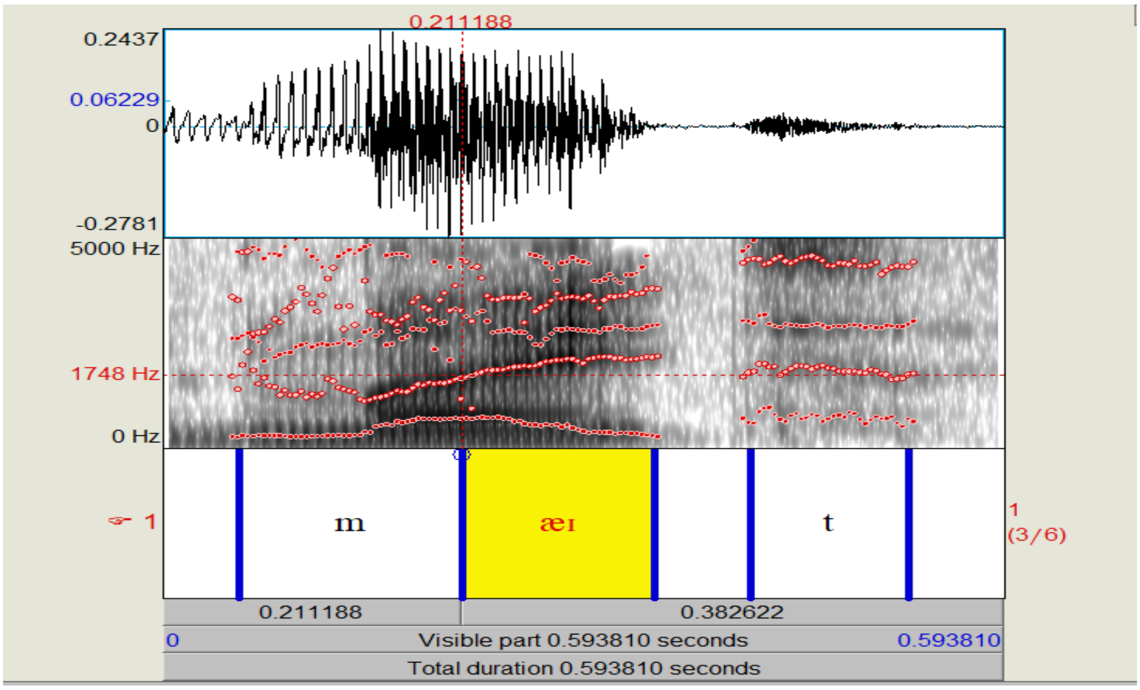

Figure 26: Waveform and spectrographic representation for the word "mate" produced by speaker AE1.

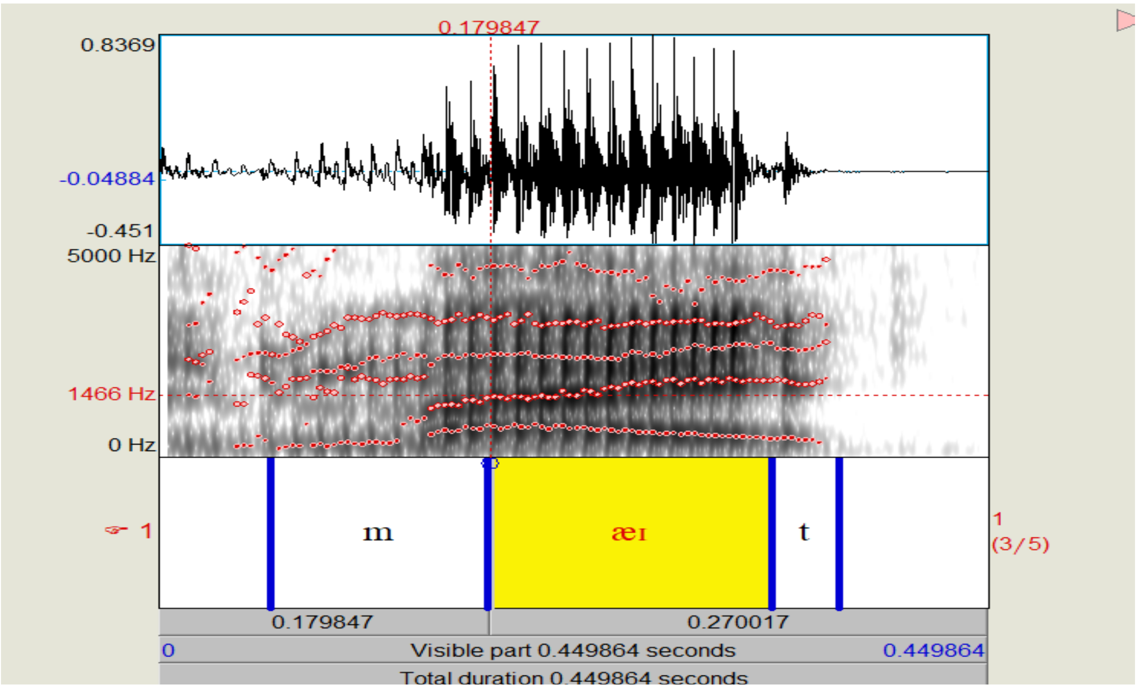

Figure 27: Waveform and spectrographic representation for the word "mate" produced by speaker AE2.

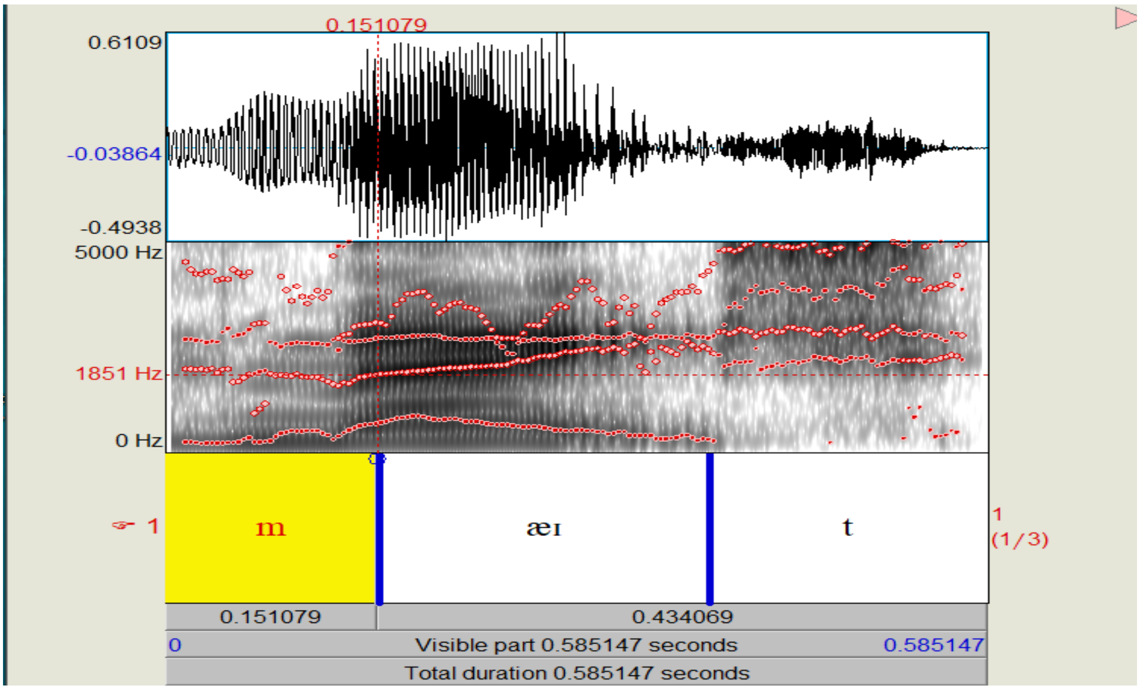

Figure 28: Waveform and spectrographic representation for the word “mate” produced by speaker AE3.

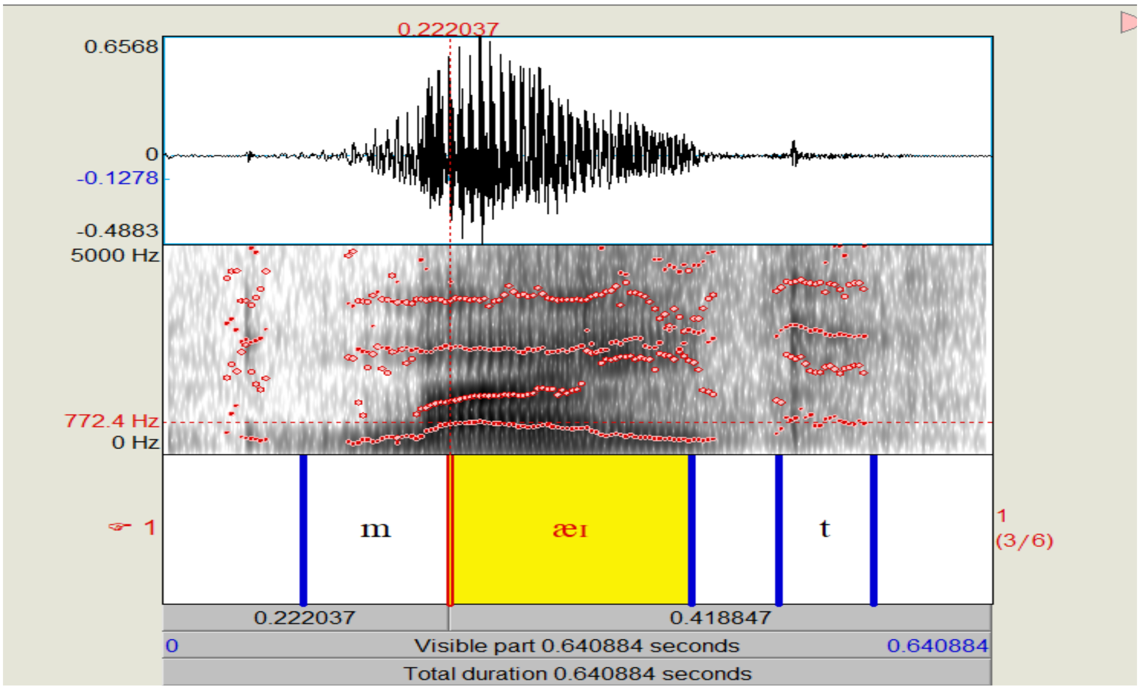

Figure 29: Waveform and spectrographic representation for the word “mate” produced by speaker AE4.

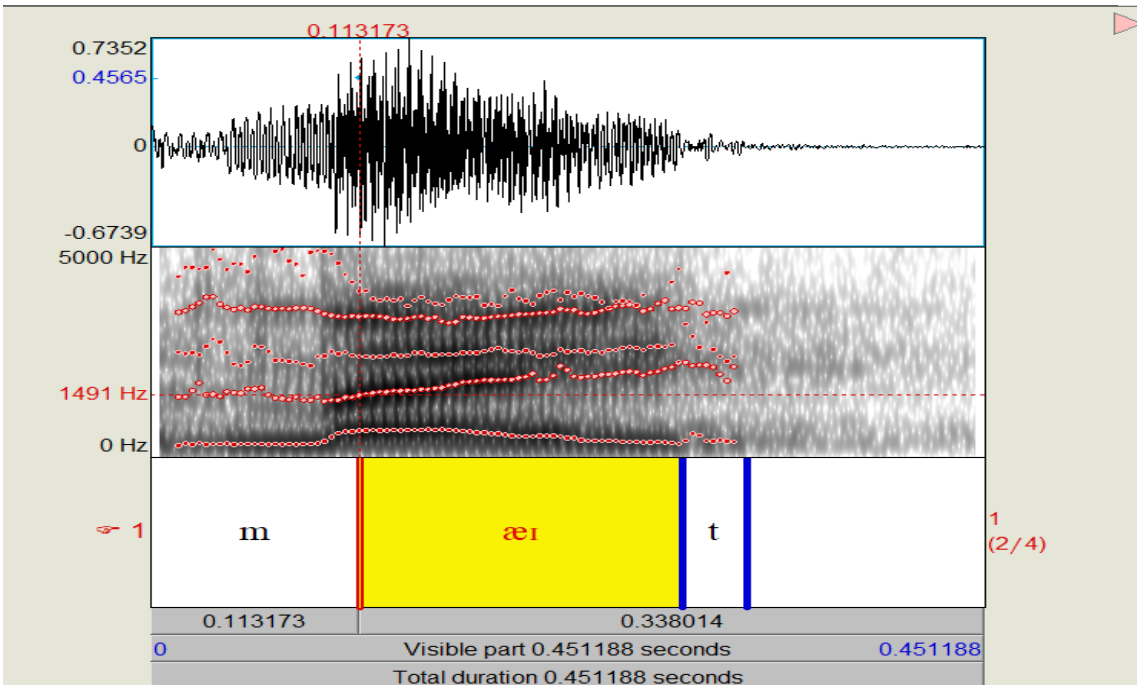

Figure 30: Waveform and spectrographic representation for the word “mate” produced by speaker AE6.

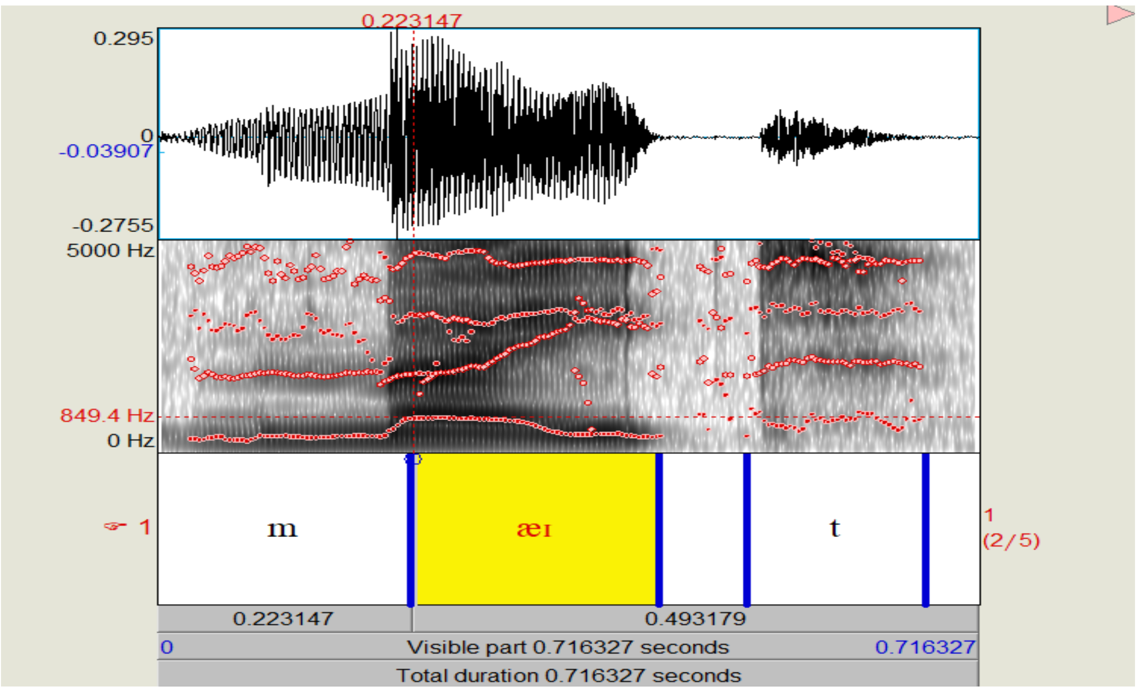

Figure 31: Waveform and spectrographic representation for the word “mate” produced by speaker AE7.

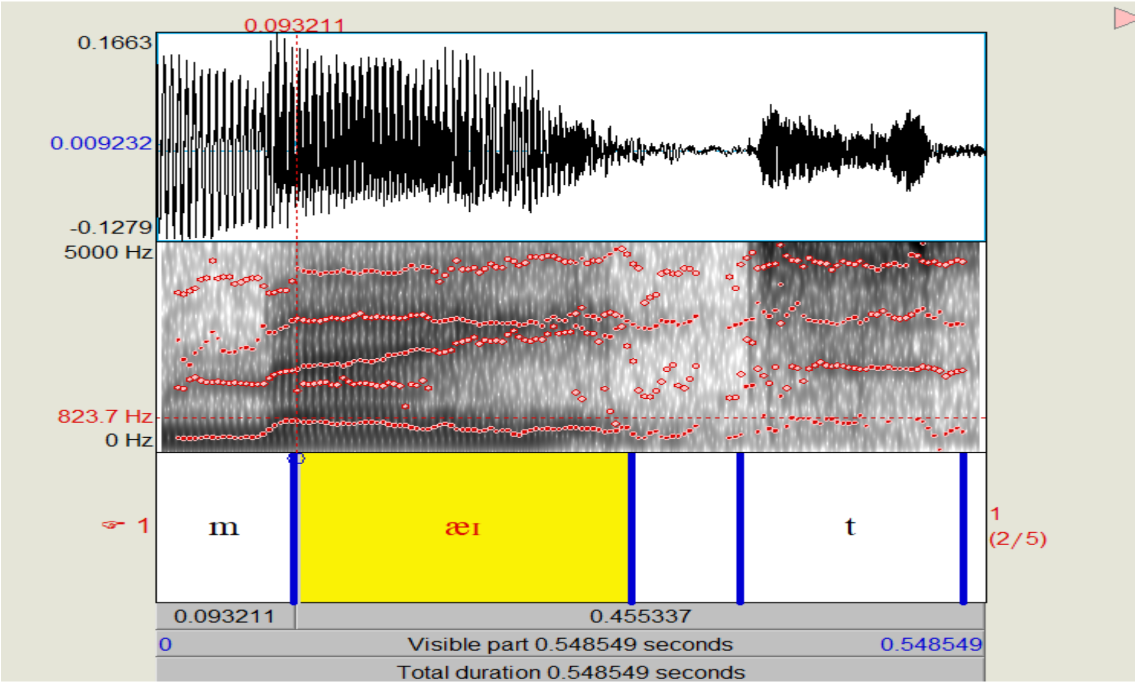

Figure 32: Waveform and spectrographic representation for the word "mate" produced by speaker AE8.

Face

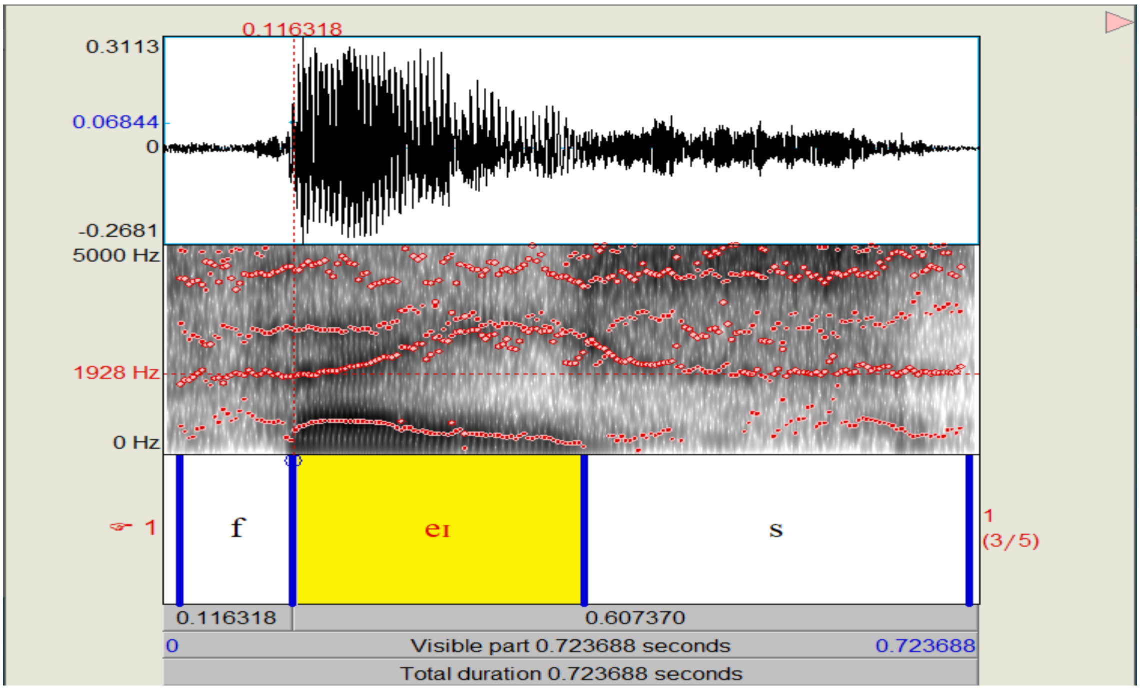

Figure 33: Waveform and spectrographic representation for the word "face" produced by speaker RP1.

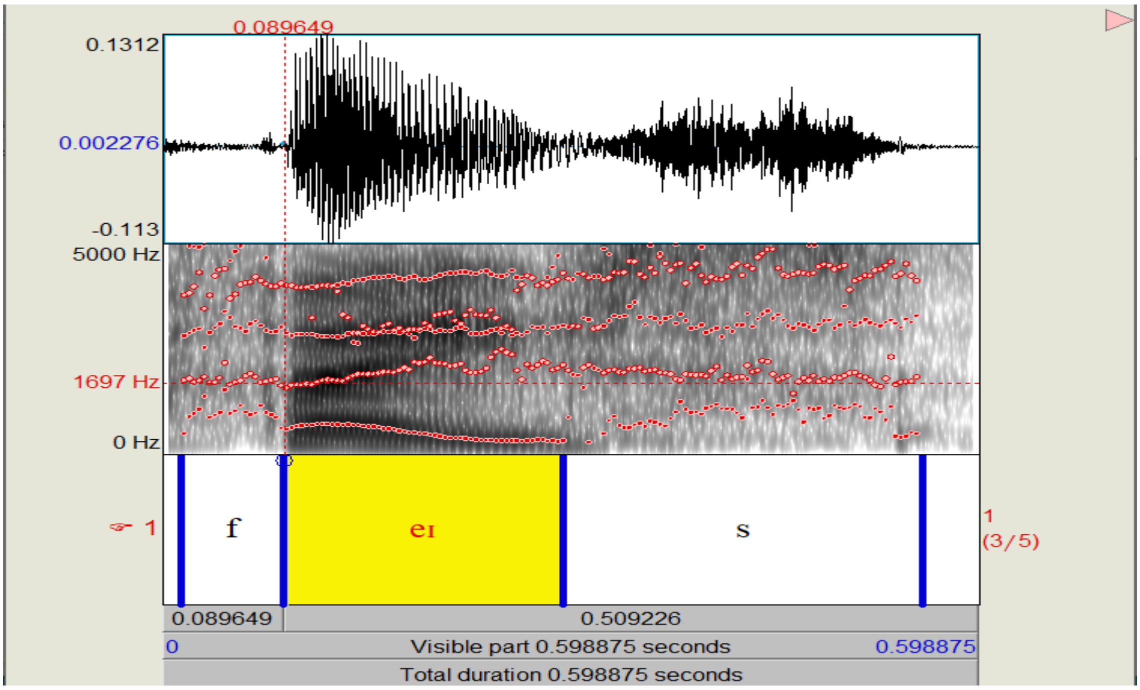

Figure 34: Waveform and spectrographic representation for the word "face" produced by speaker RP2.

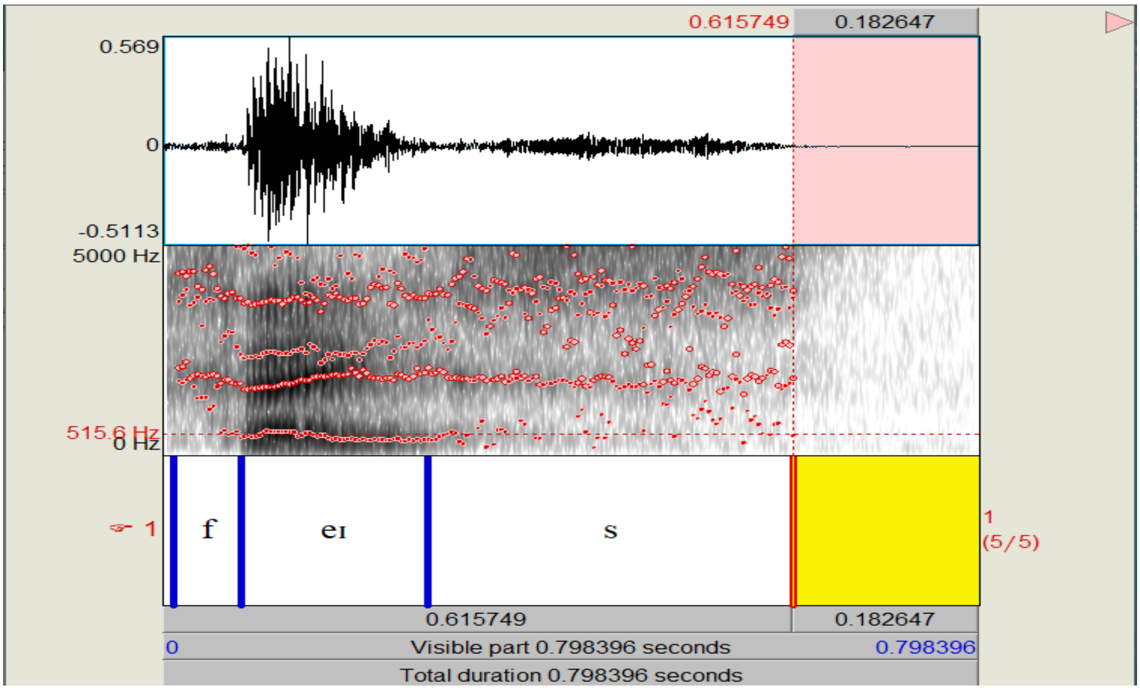

Figure 35: Waveform and spectrographic representation for the word "face" produced by speaker RP4.

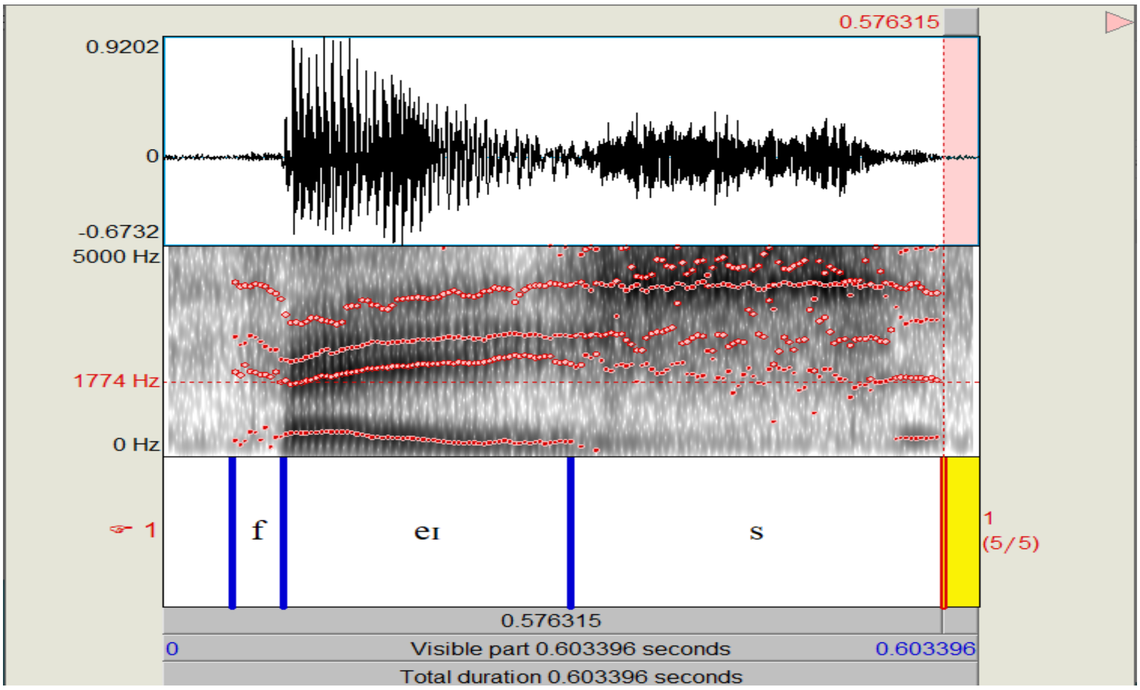

Figure 36: Waveform and spectrographic representation for the word "face" produced by speaker RP5.

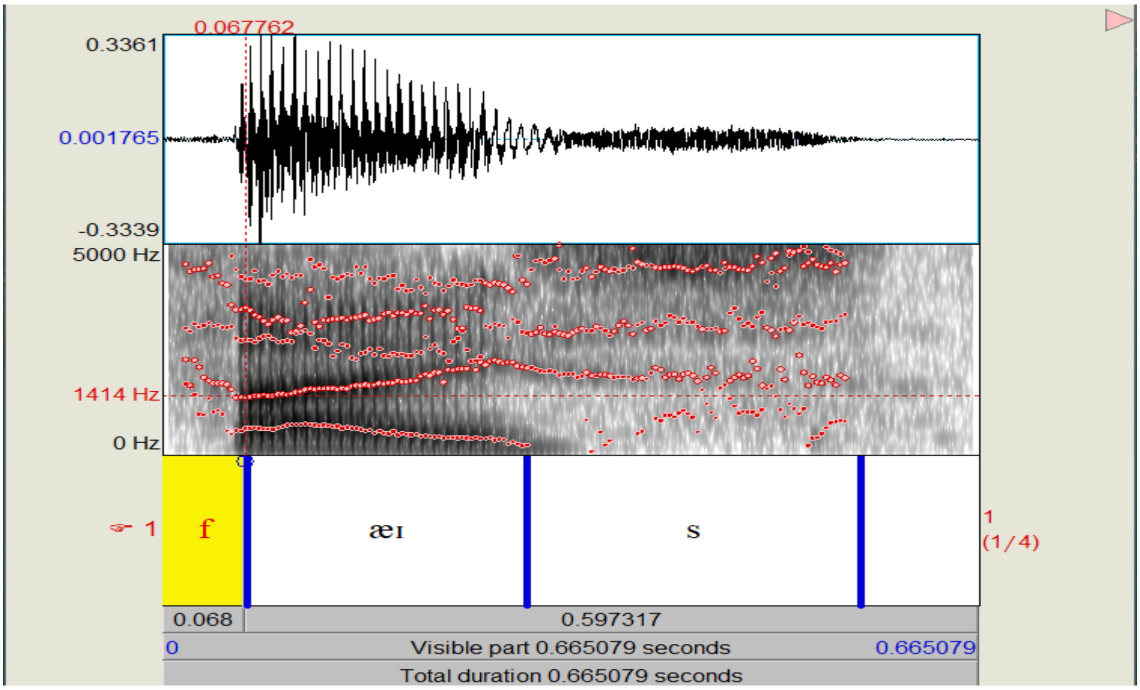

Figure 37: Waveform and spectrographic representation for the word "face" produced by speaker AE1.

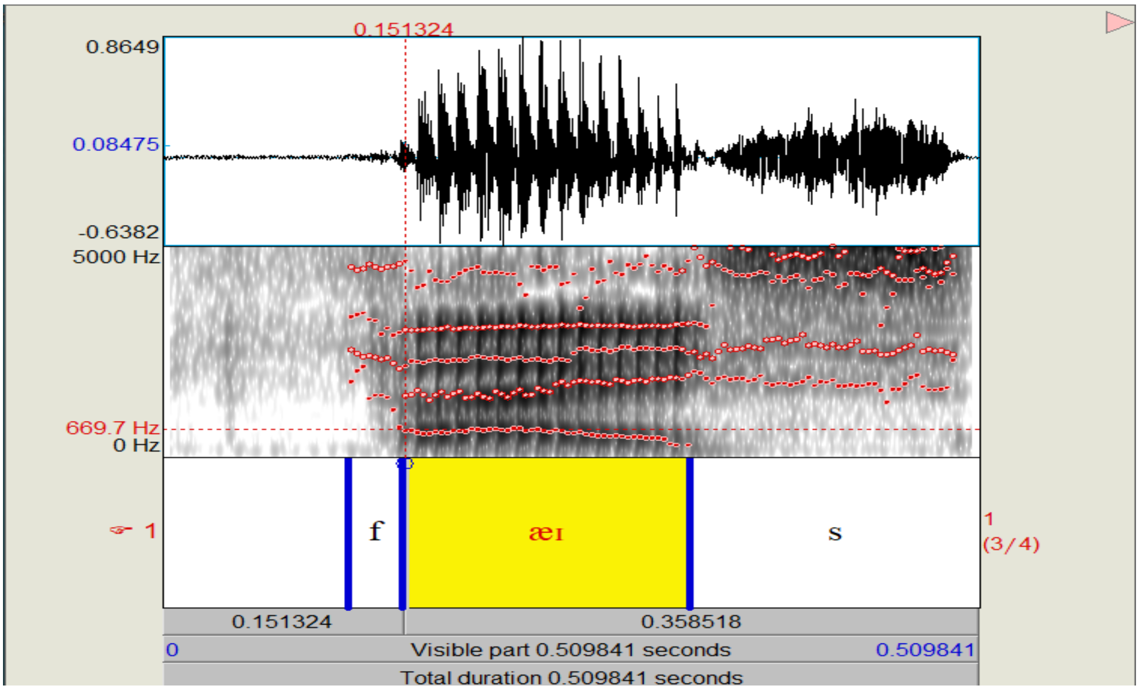

Figure 38: Waveform and spectrographic representation for the word "face" produced by speaker AE2.

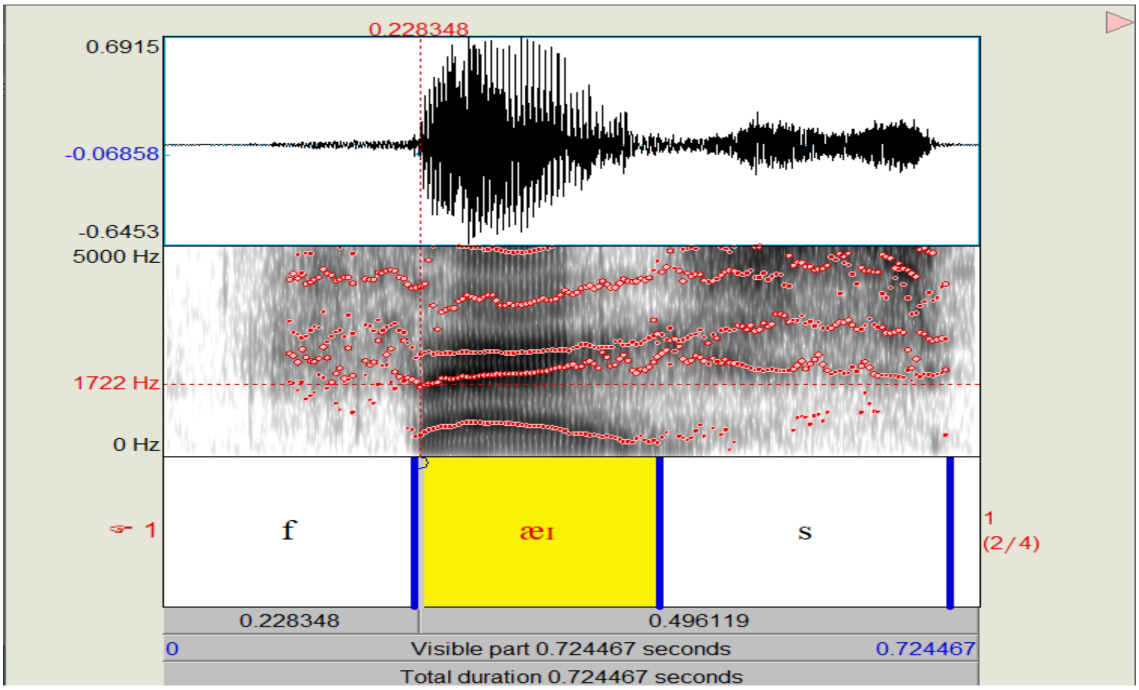

Figure 39: Waveform and spectrographic representation for the word "face" produced by speaker AE3.

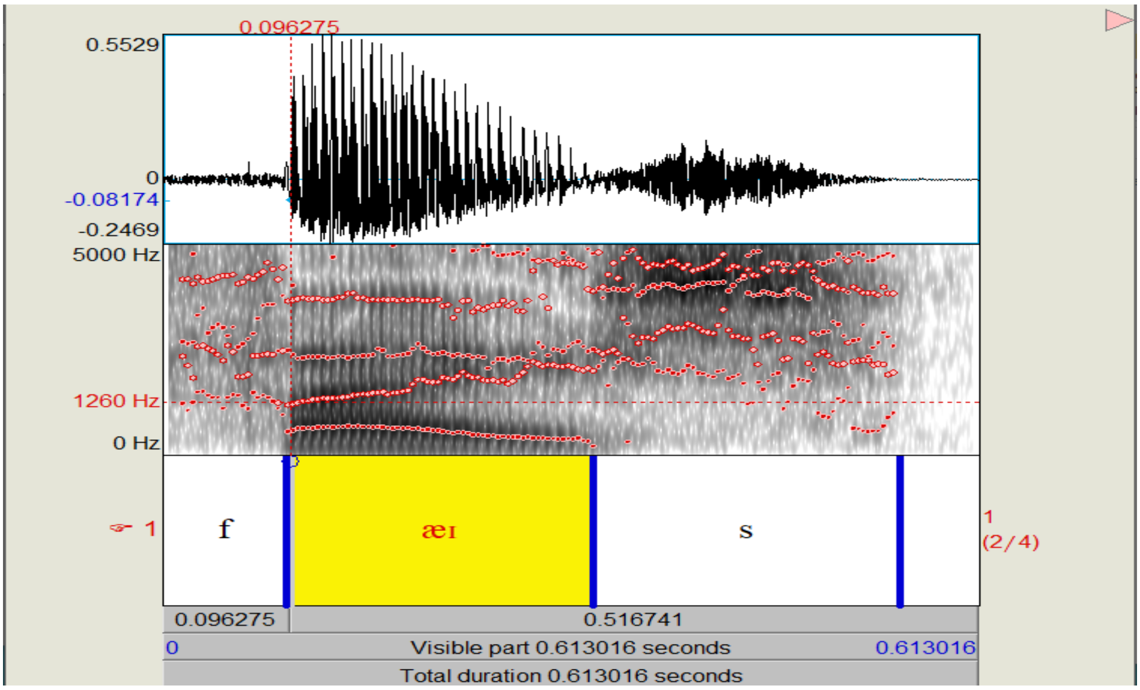

Figure 40: Waveform and spectrographic representation for the word "face" produced by speaker AE4.

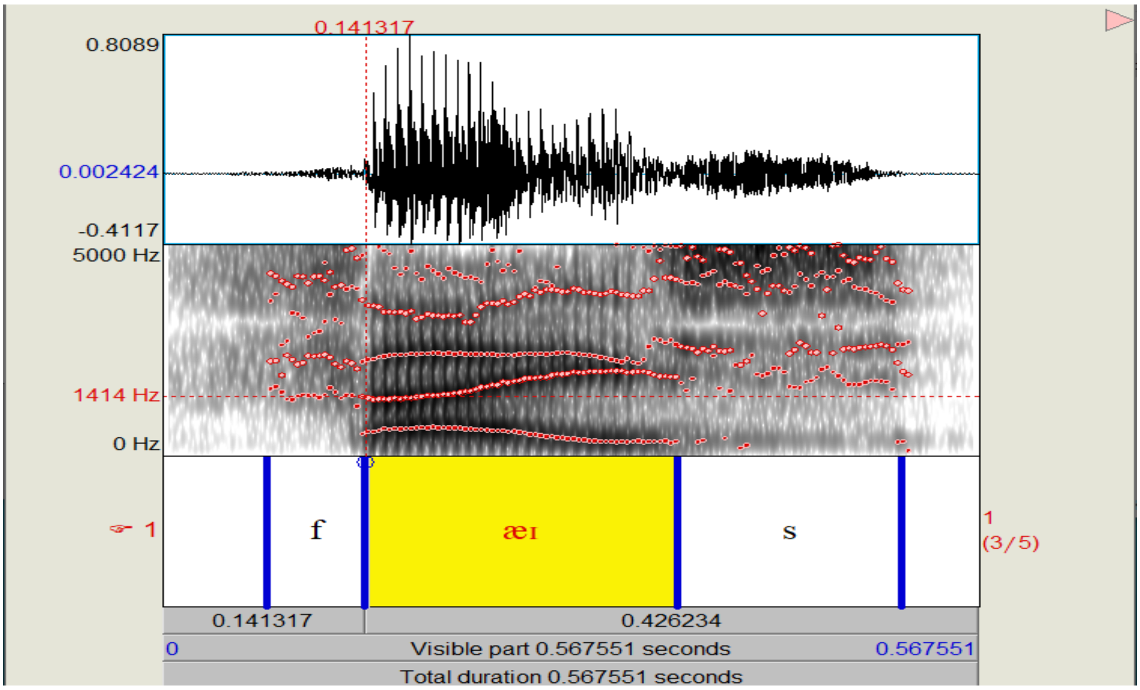

Figure 41: Waveform and spectrographic representation for the word "face" produced by speaker AE5.

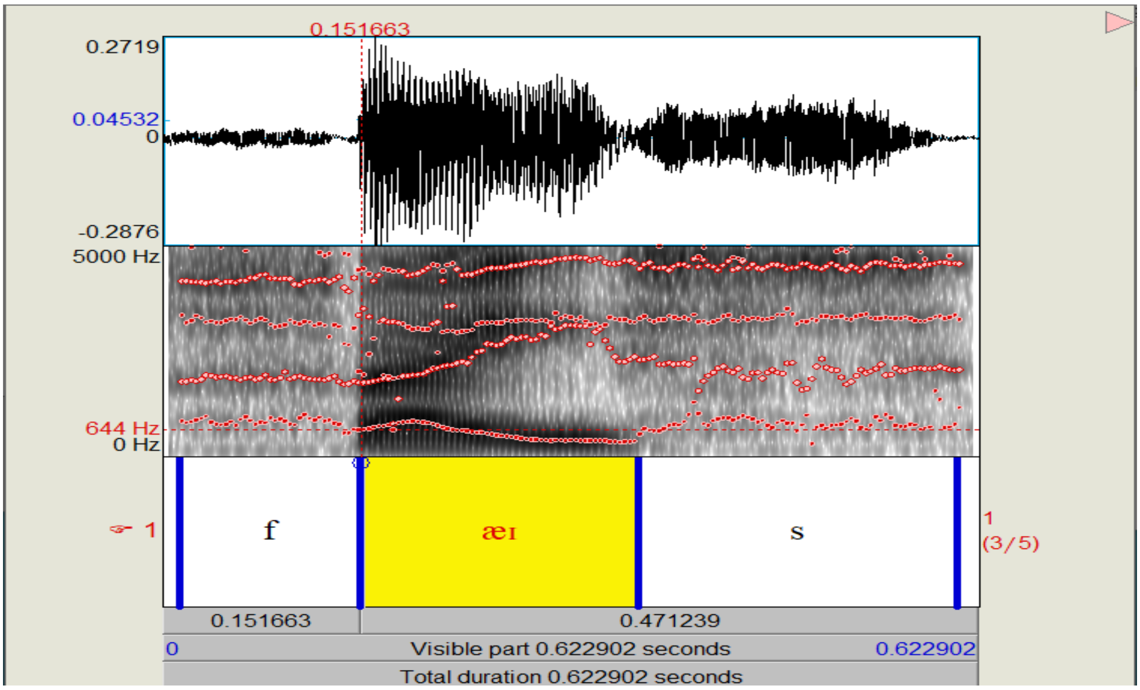

Figure 42: Waveform and spectrographic representation for the word "face" produced by speaker AE7.

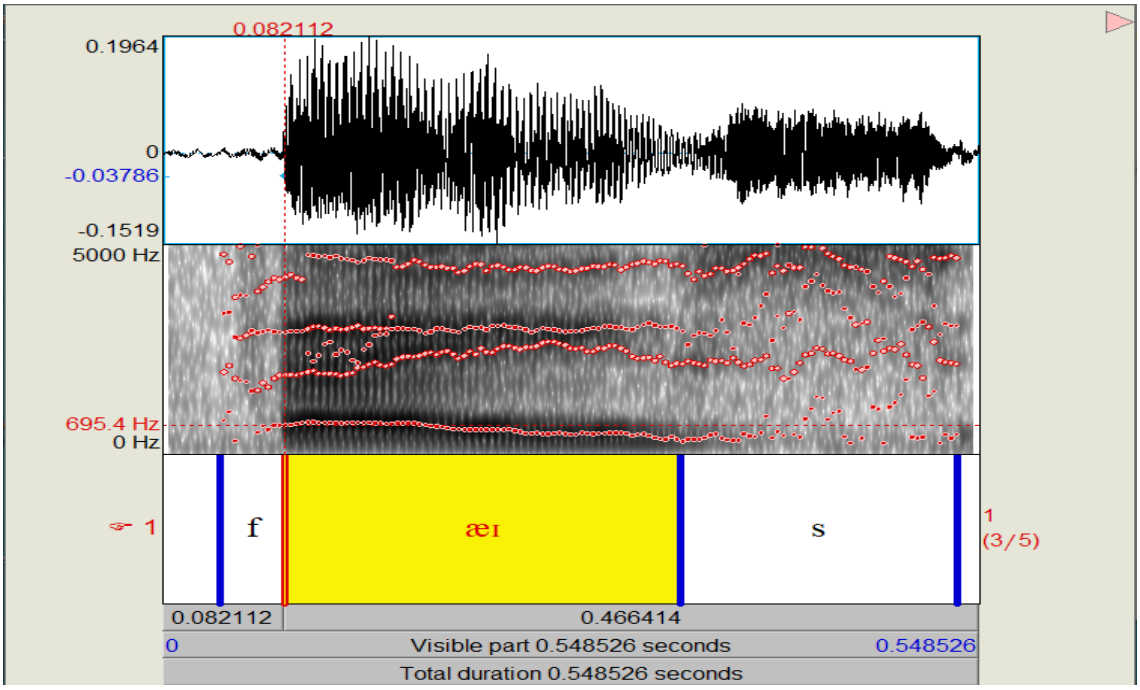

Figure 43: Waveform and spectrographic representation for the word "face" produced by speaker AE8.

Light

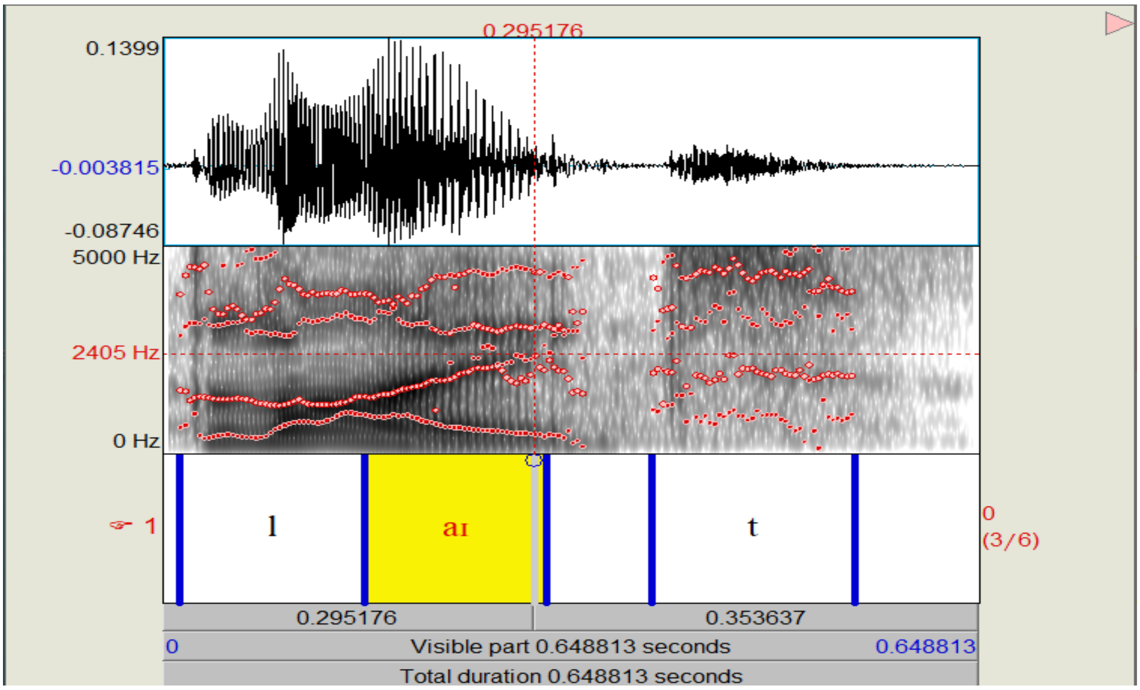

Figure 44: Waveform and spectrographic representation for the word “light” produced by speaker RP2.

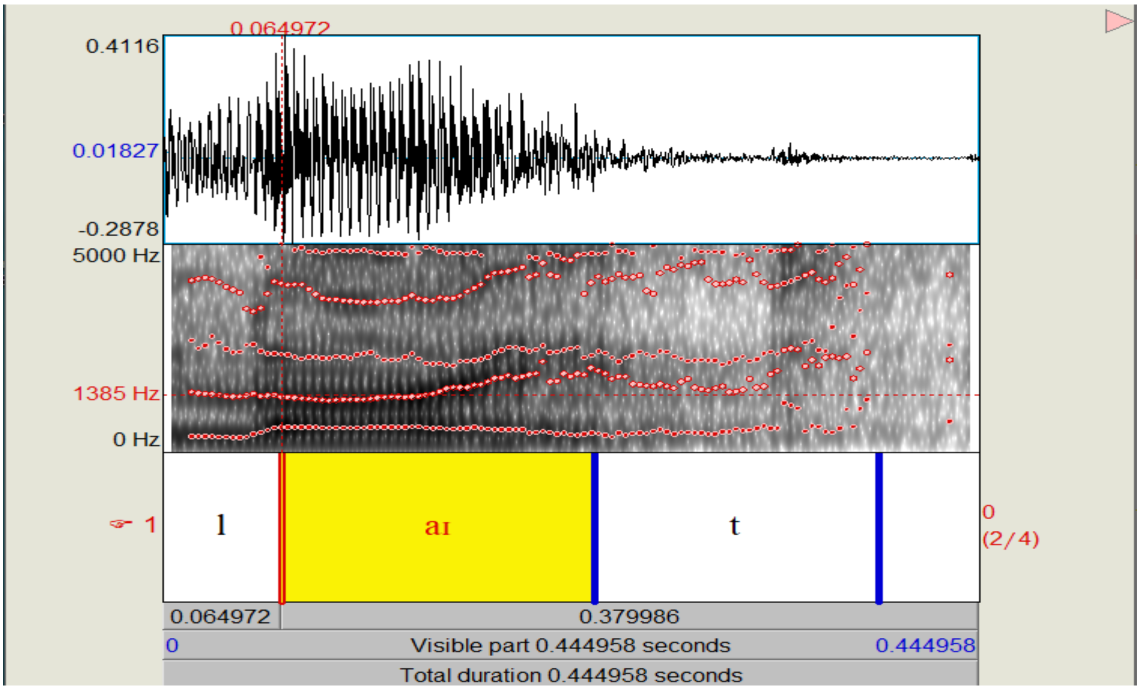

Figure 45: Waveform and spectrographic representation for the word “light” produced by speaker RP3.

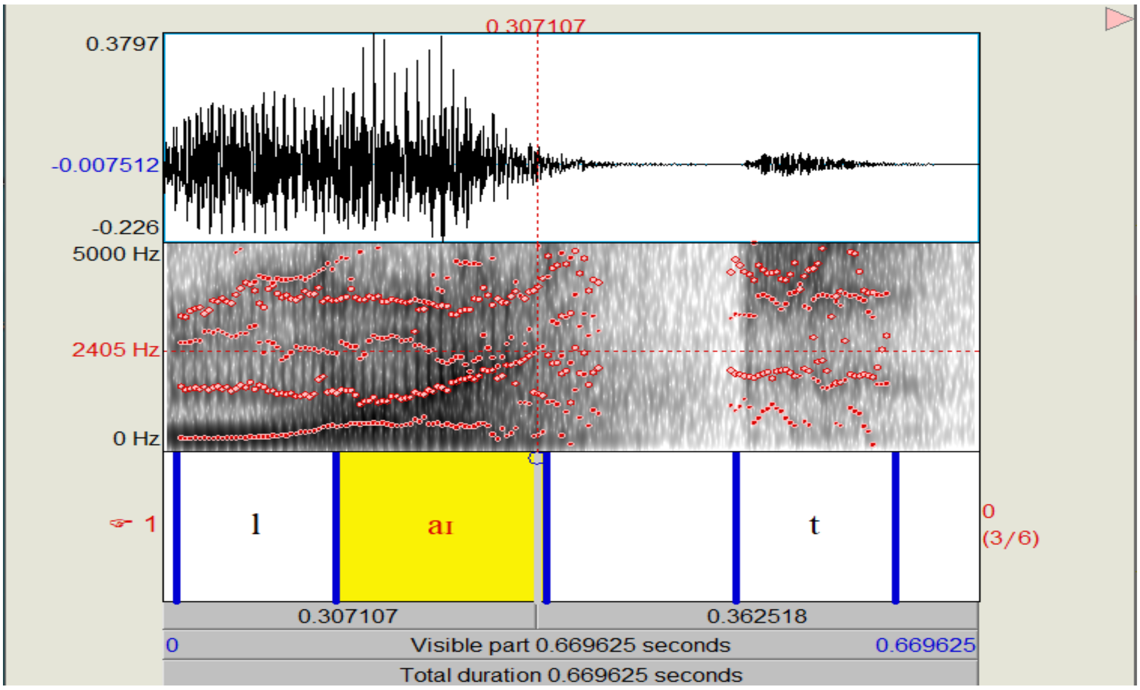

Figure 46: Waveform and spectrographic representation for the word "light" produced by speaker RP4.

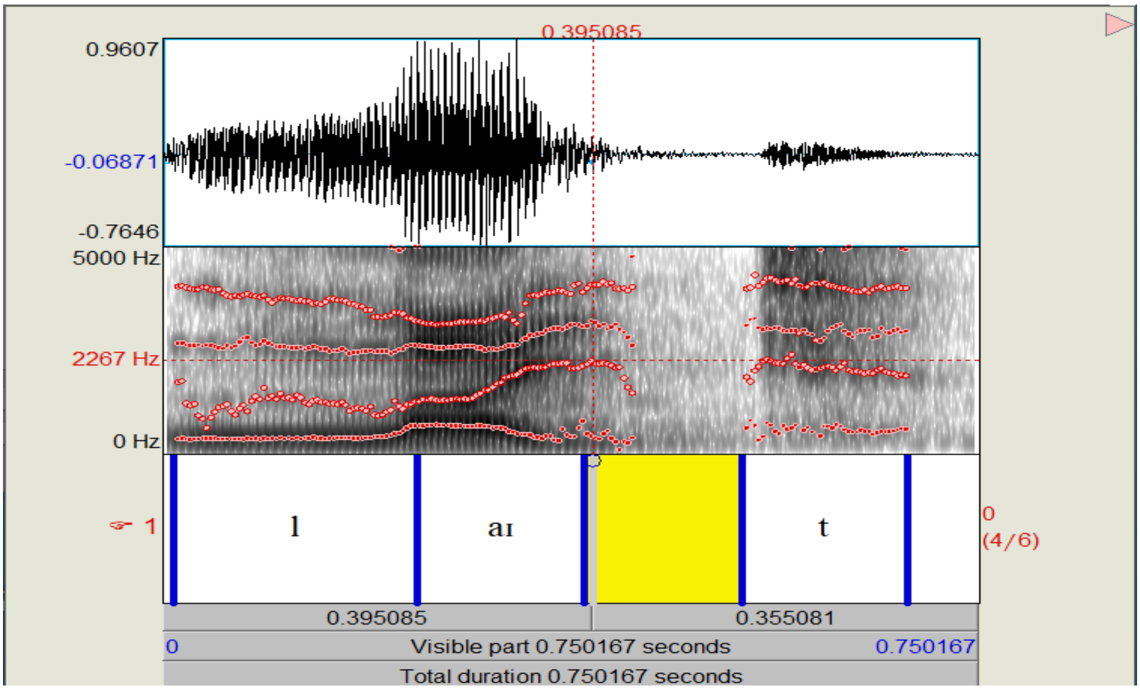

Figure 47: Waveform and spectrographic representation for the word "light" produced by speaker RP5.

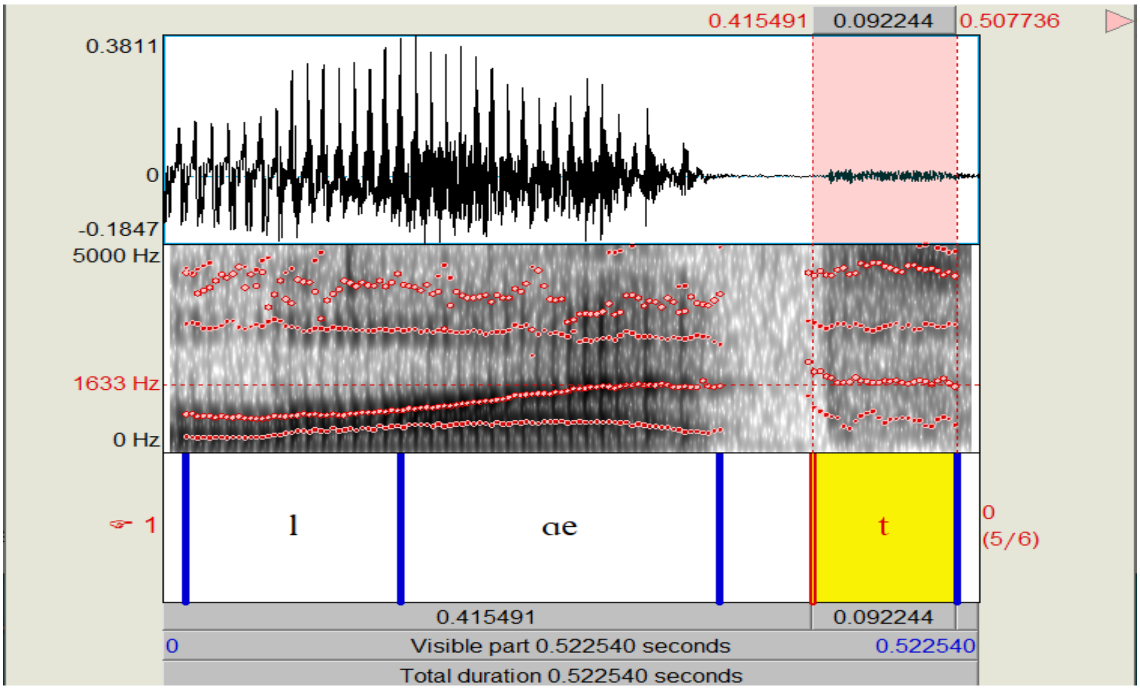

Figure 48: Waveform and spectrographic representation for the word "light" produced by speaker AE1.

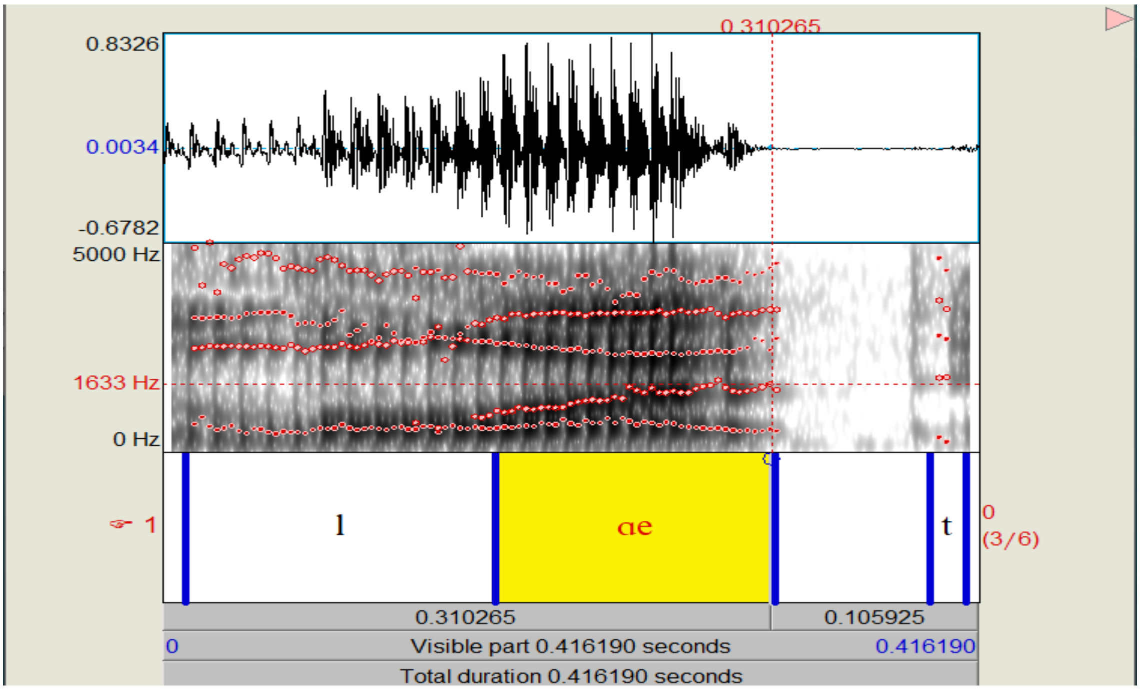

Figure 49: Waveform and spectrographic representation for the word "light" produced by speaker AE2.

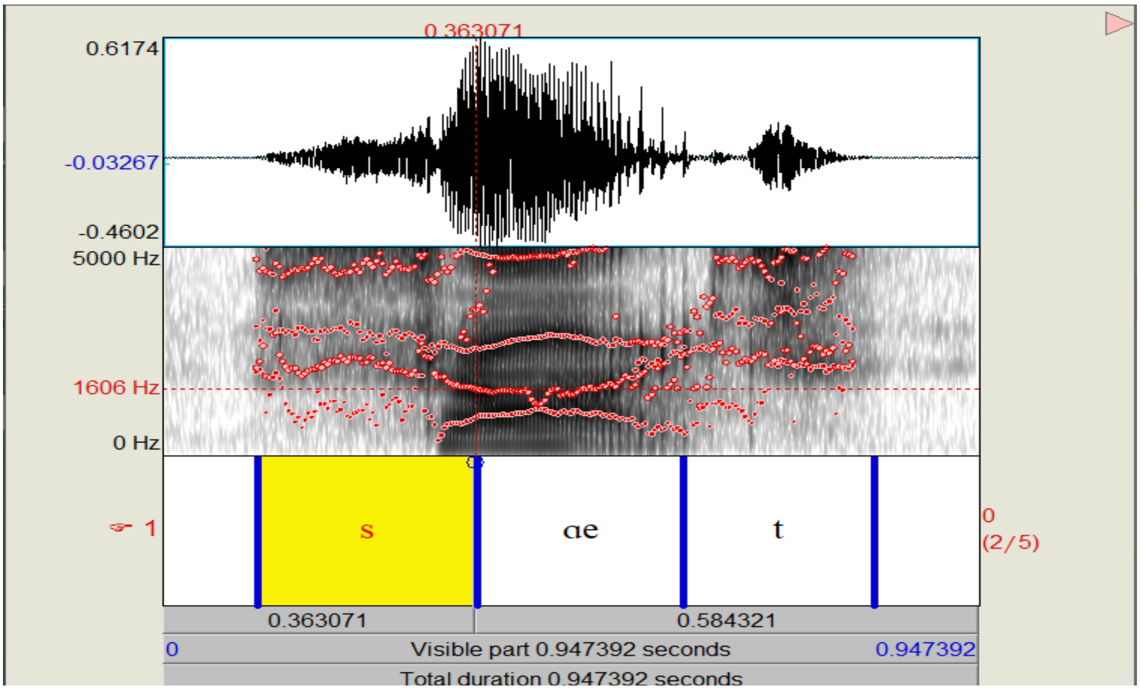

Figure 45: Waveform and spectrographic representation for the word “sight” produced by speaker AE3 by mistake.

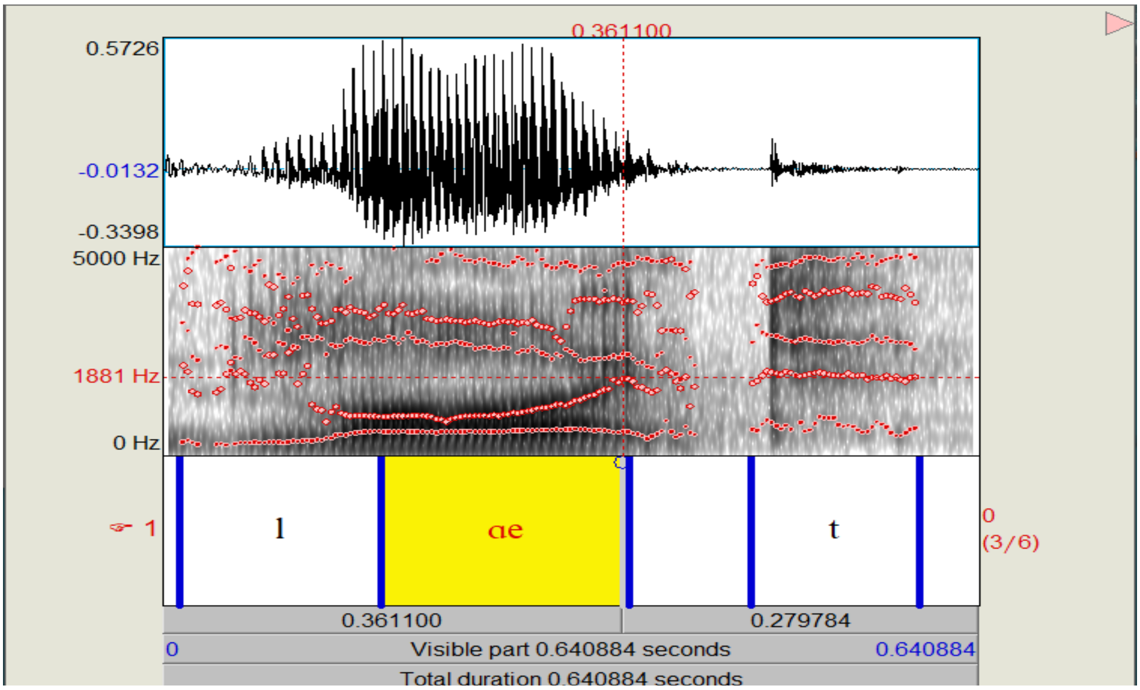

Figure 46: Waveform and spectrographic representation for the word “sight” produced by speaker AE4.

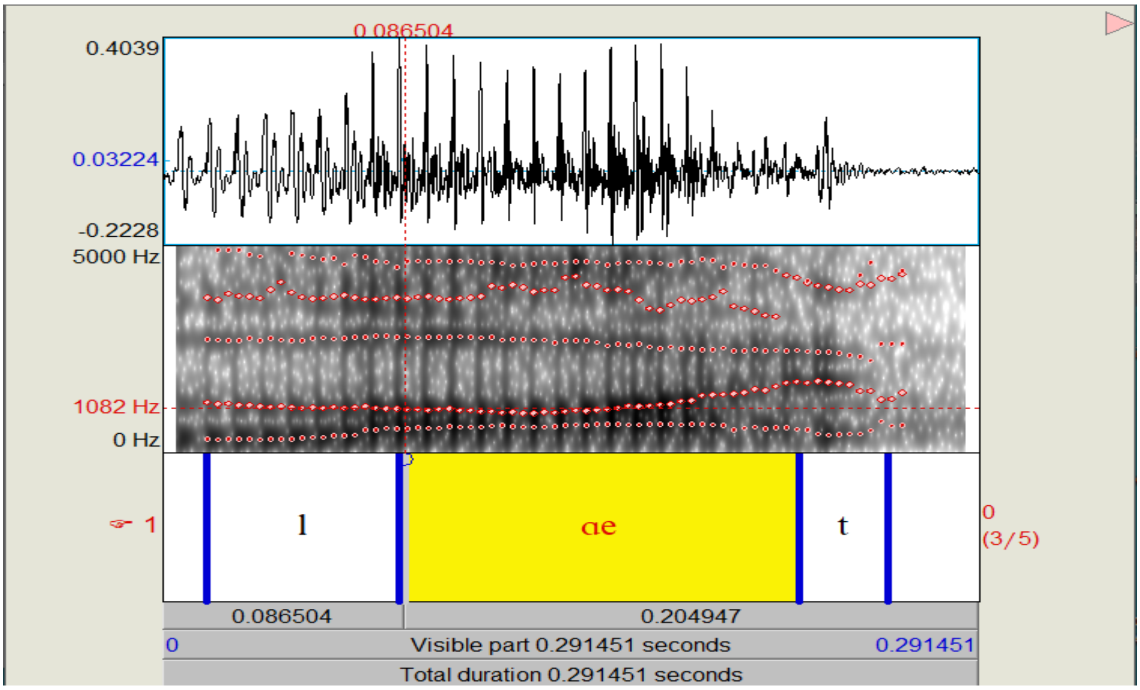

Figure 47: Waveform and spectrographic representation for the word “sight” produced by speaker AE5.

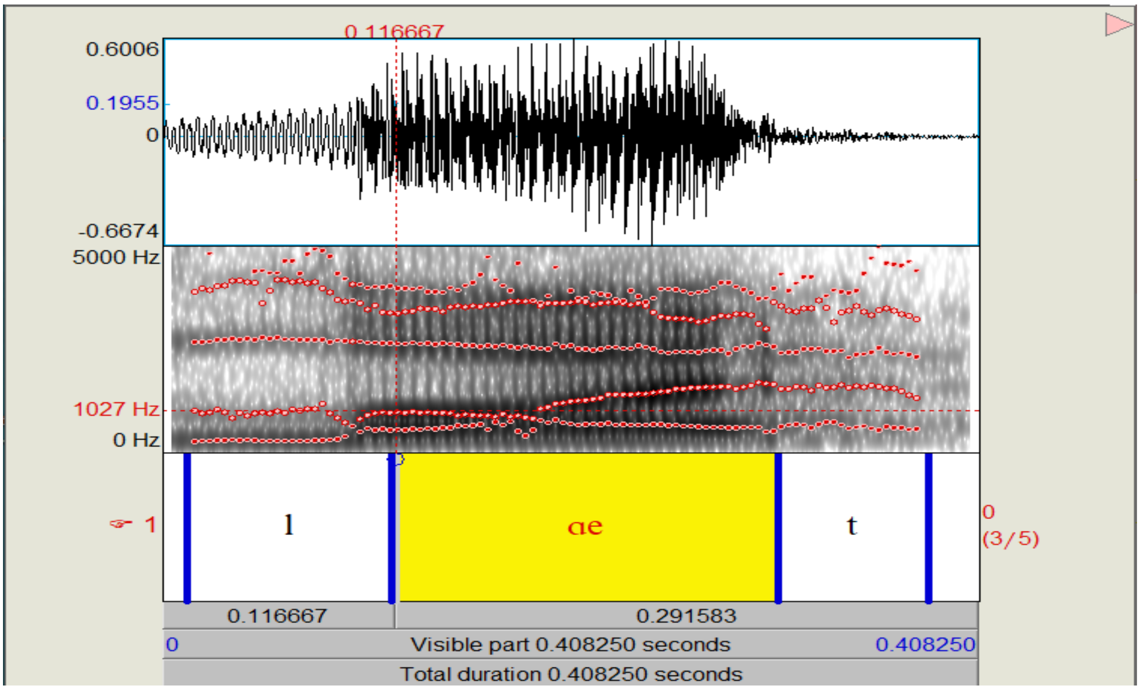

Figure 48: Waveform and spectrographic representation for the word “sight” produced by speaker AE6.

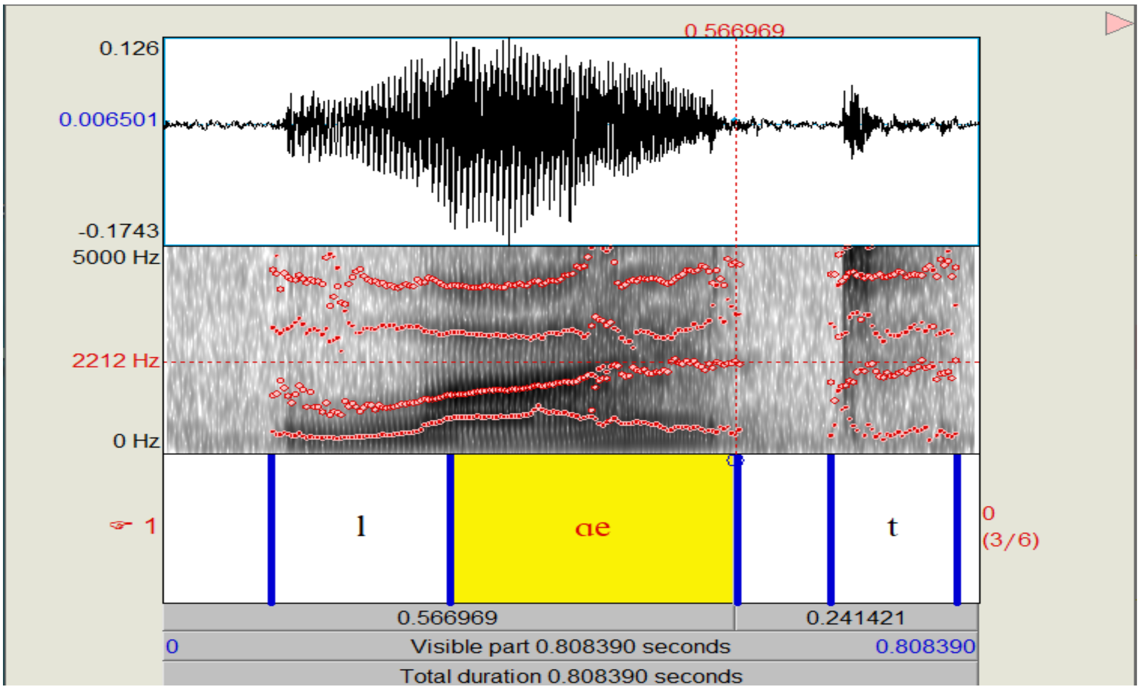

Figure 49: Waveform and spectrographic representation for the word “sight” produced by speaker AE8.

Hide

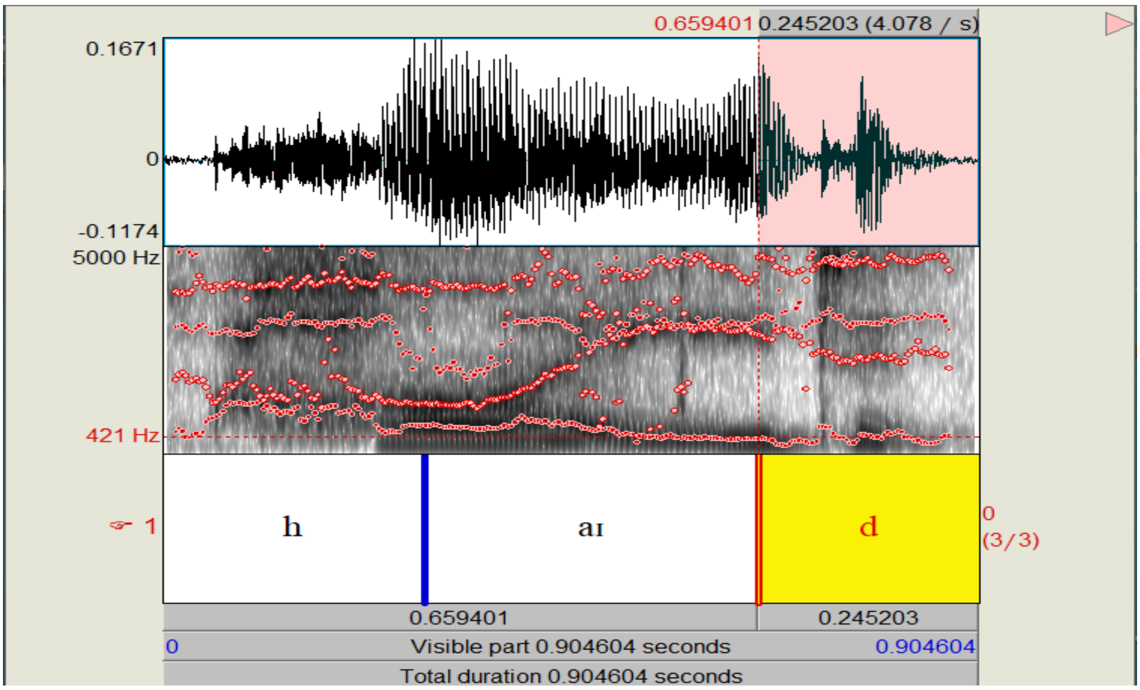

Figure 49: Waveform and spectrographic representation for the word “hide” produced by speaker RP1.

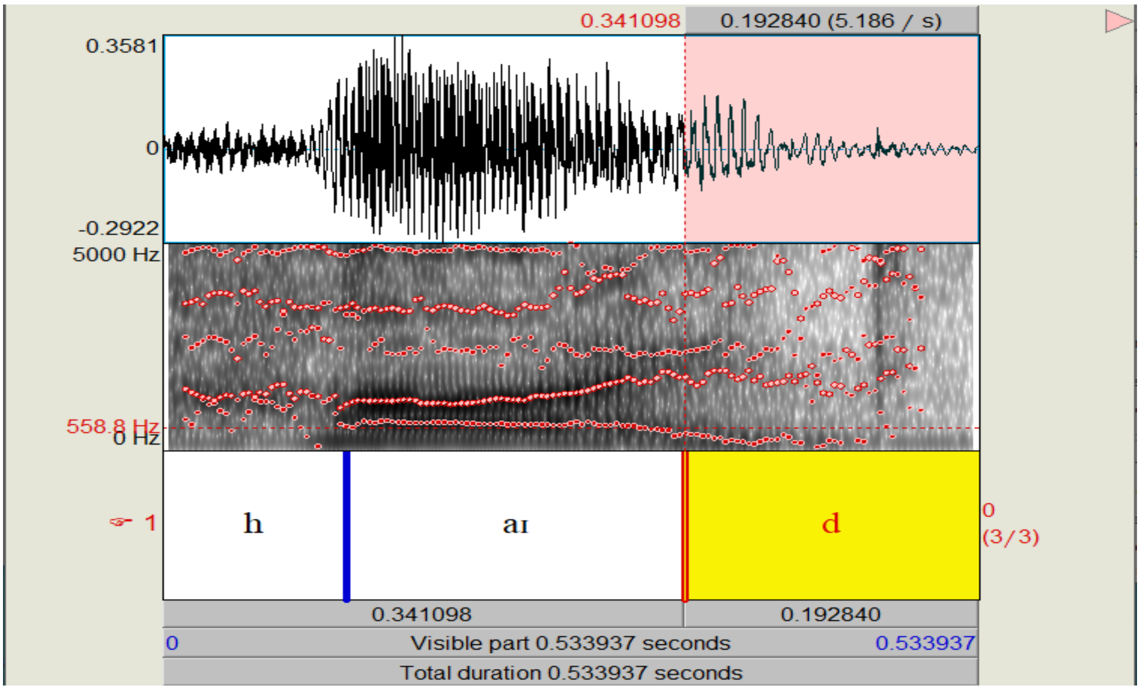

Figure 50: Waveform and spectrographic representation for the word "hide" produced by speaker RP3.

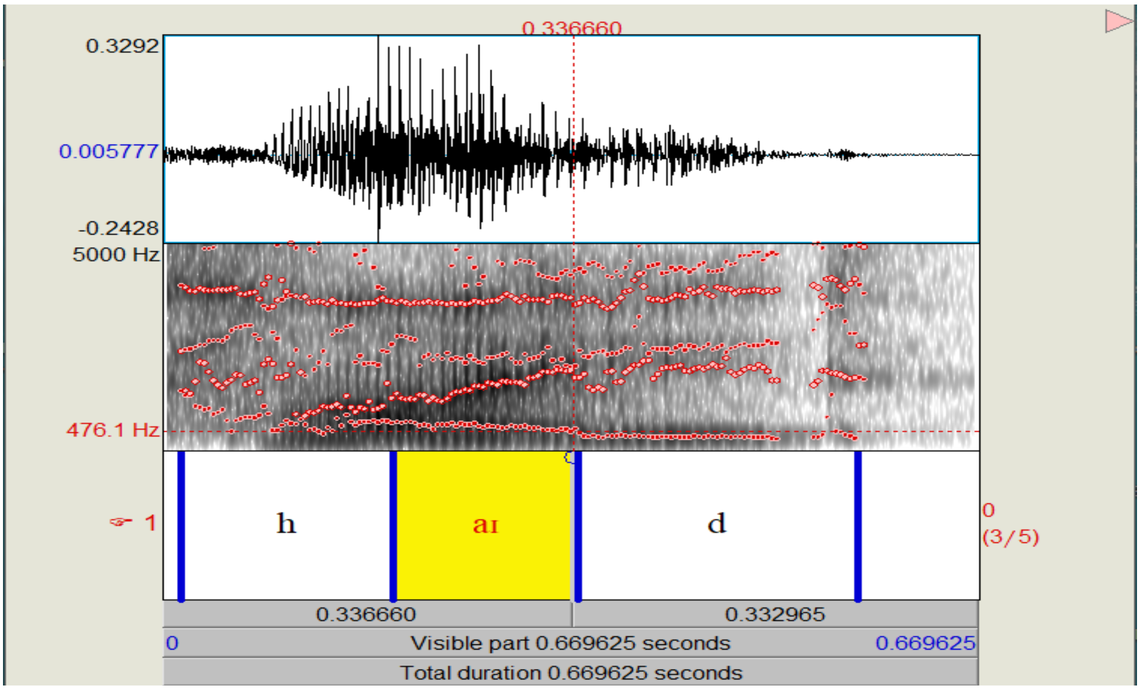

Figure 51: Waveform and spectrographic representation for the word "hide" produced by speaker RP4.

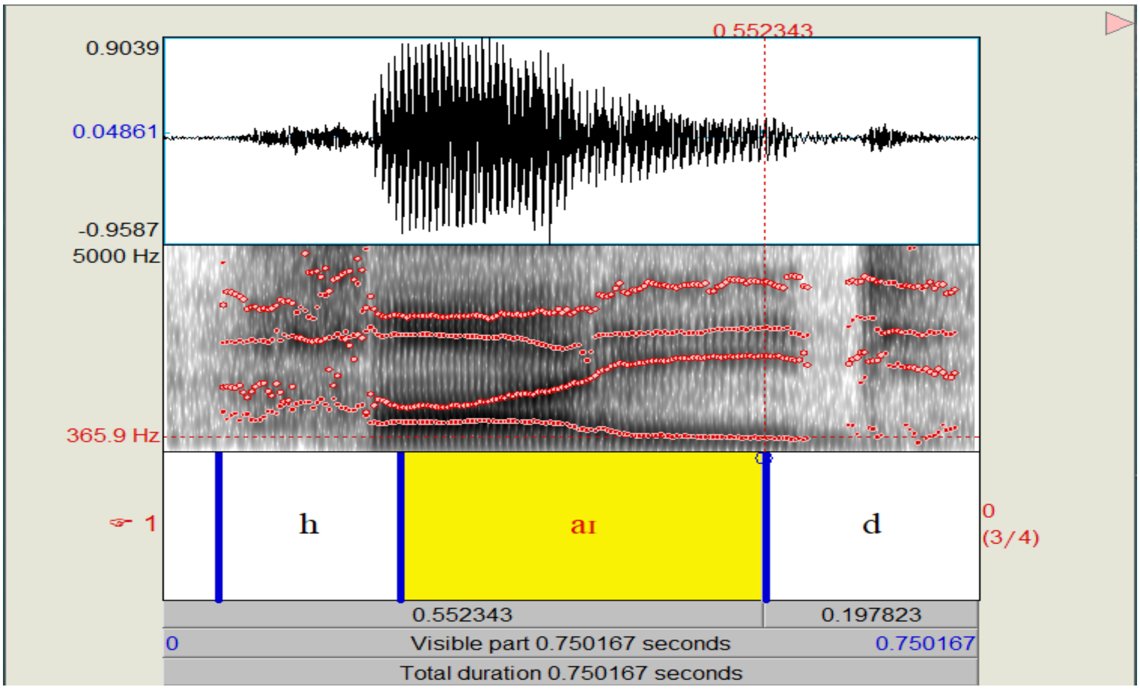

Figure 52: Waveform and spectrographic representation for the word “hide” produced by speaker RP5.

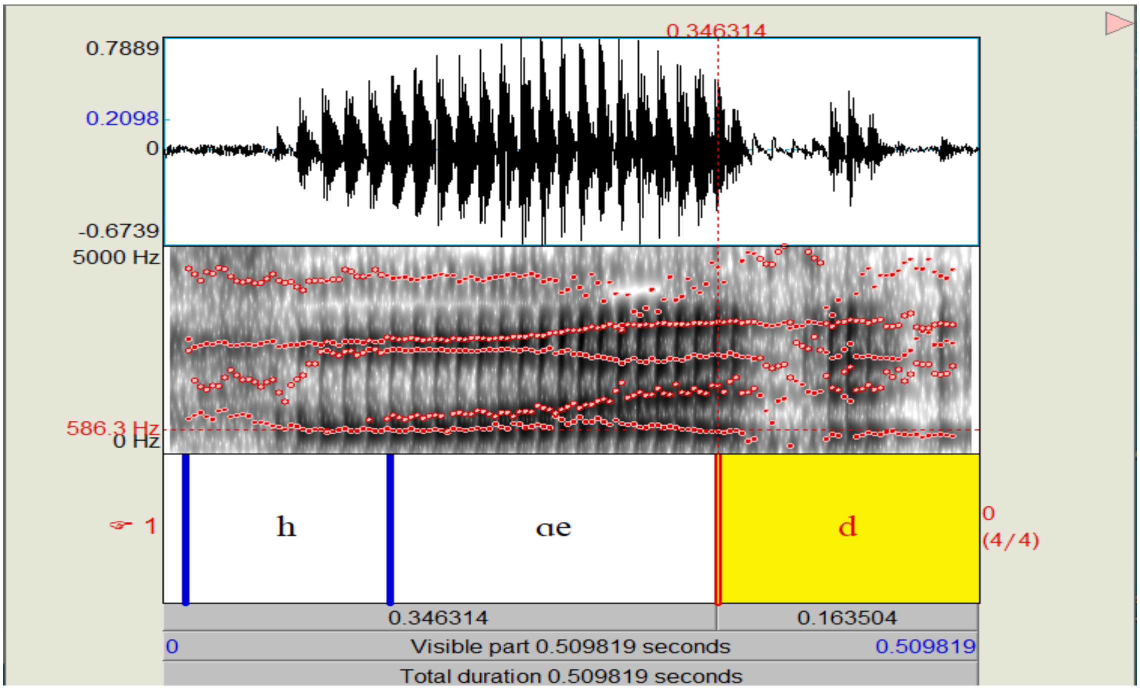

Figure 53: Waveform and spectrographic representation for the word “hide” produced by speaker AE2.

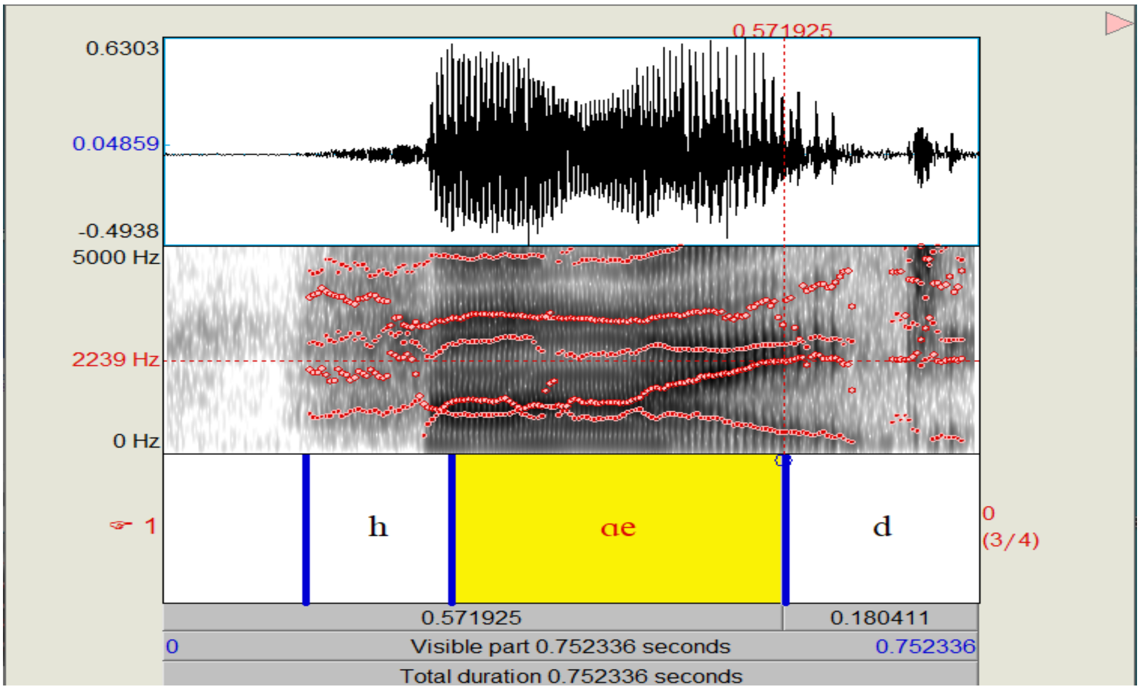

Figure 54: Waveform and spectrographic representation for the word "hide" produced by speaker AE3.

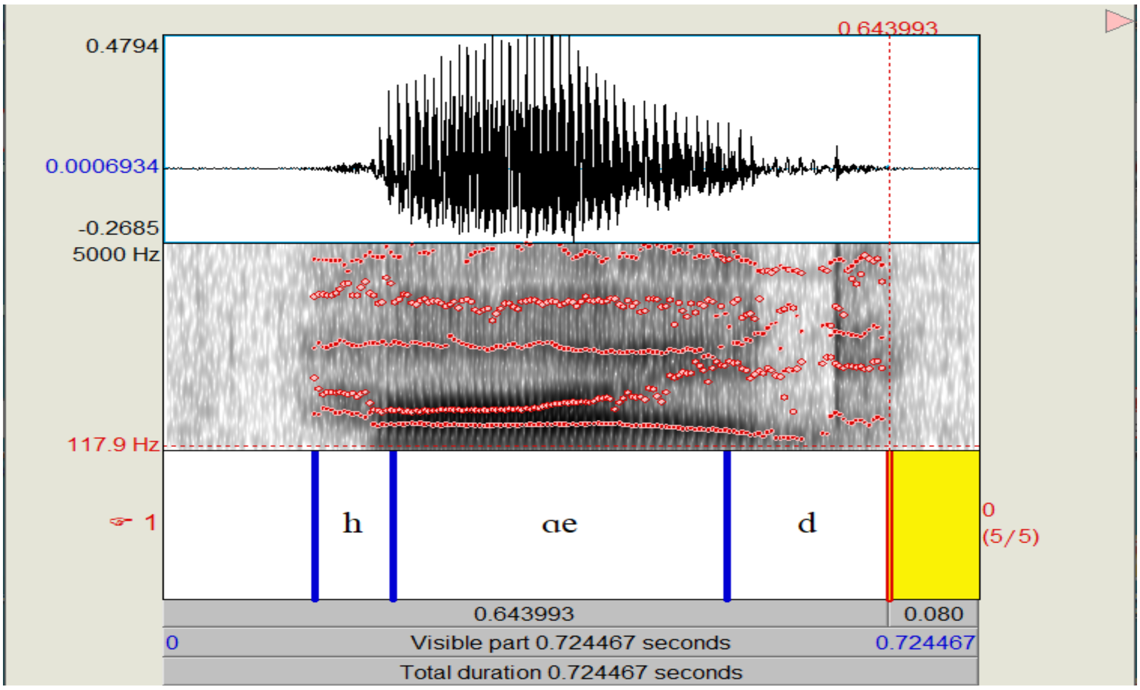

Figure 55: Waveform and spectrographic representation for the word "hide" produced by speaker AE4.

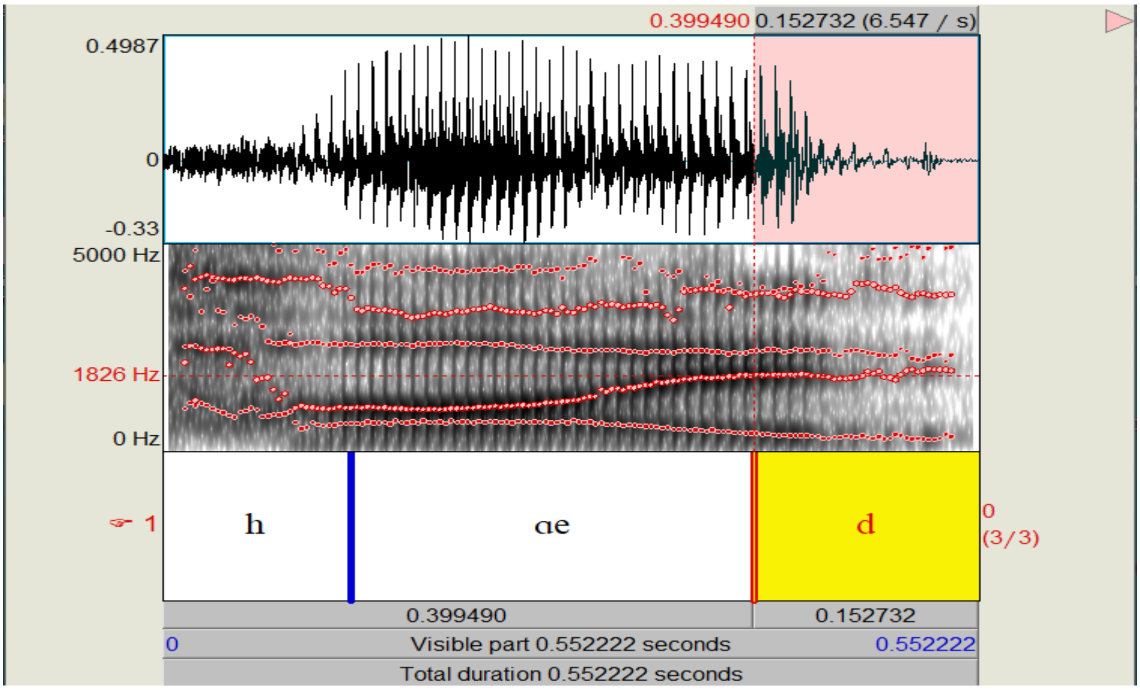

Figure 56: Waveform and spectrographic representation for the word "hide" produced by speaker AE5.

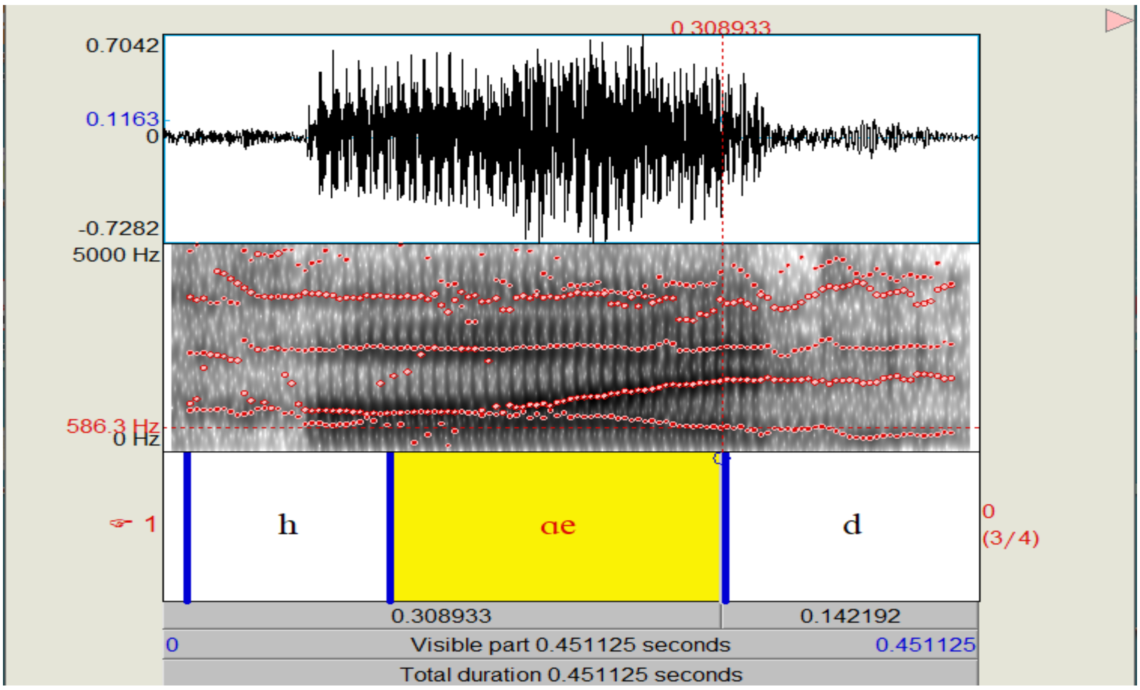

Figure 57: Waveform and spectrographic representation for the word "hide" produced by speaker AE6.

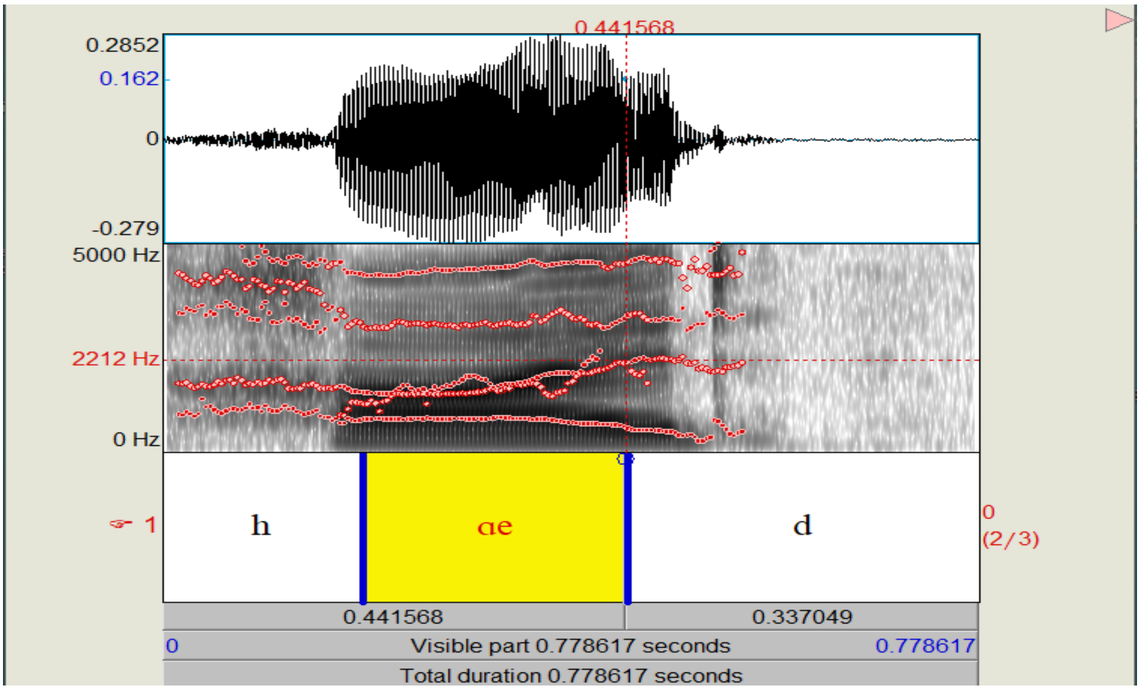

Figure 58: Waveform and spectrographic representation for the word "hide" produced by speaker AE7.

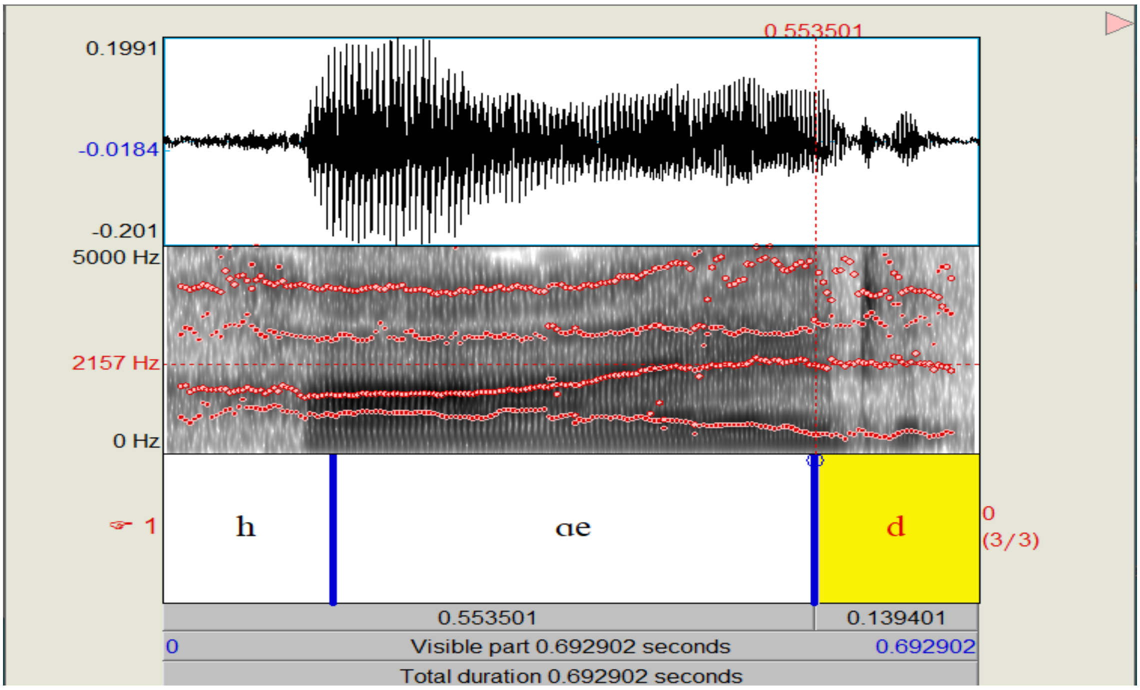

Figure 59: Waveform and spectrographic representation for the word "hide" produced by speaker AE8.

Cold

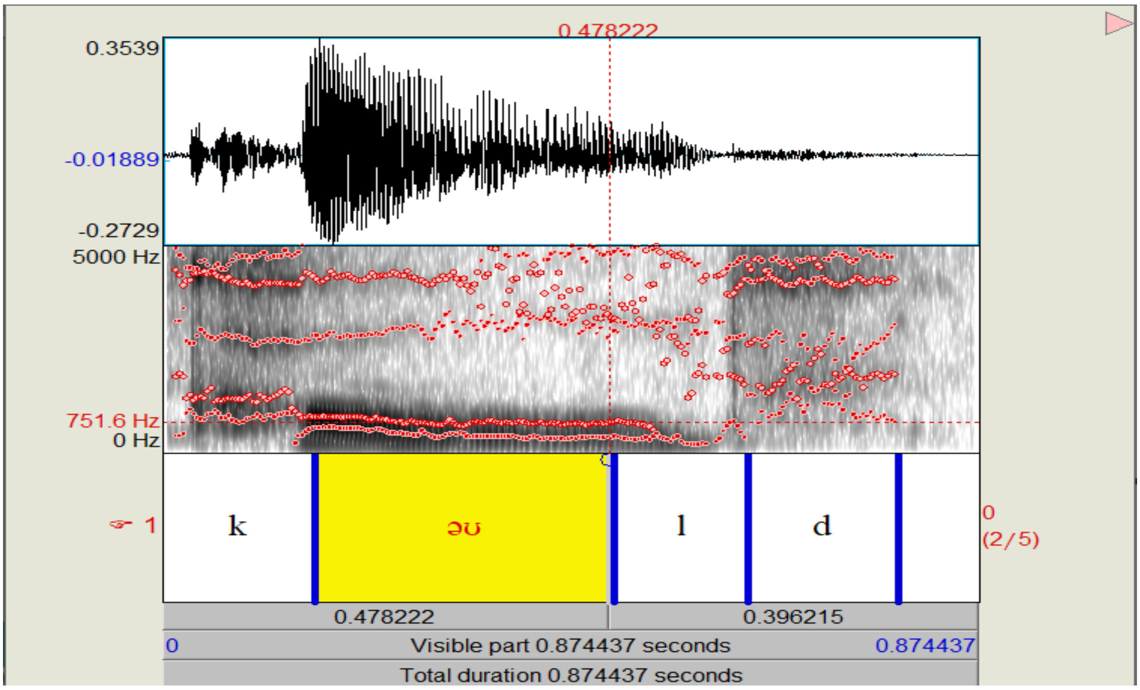

Figure 60: Waveform and spectrographic representation for the word "cold" produced by speaker RP1.

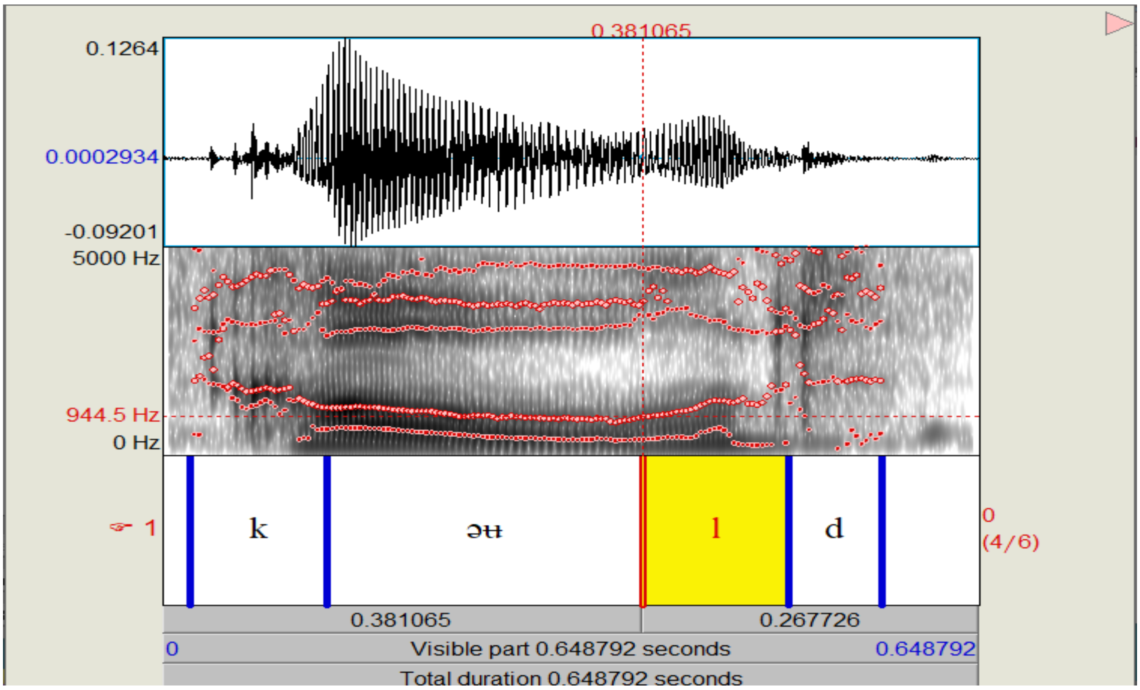

Figure 61: Waveform and spectrographic representation for the word "cold" produced by speaker RP2.

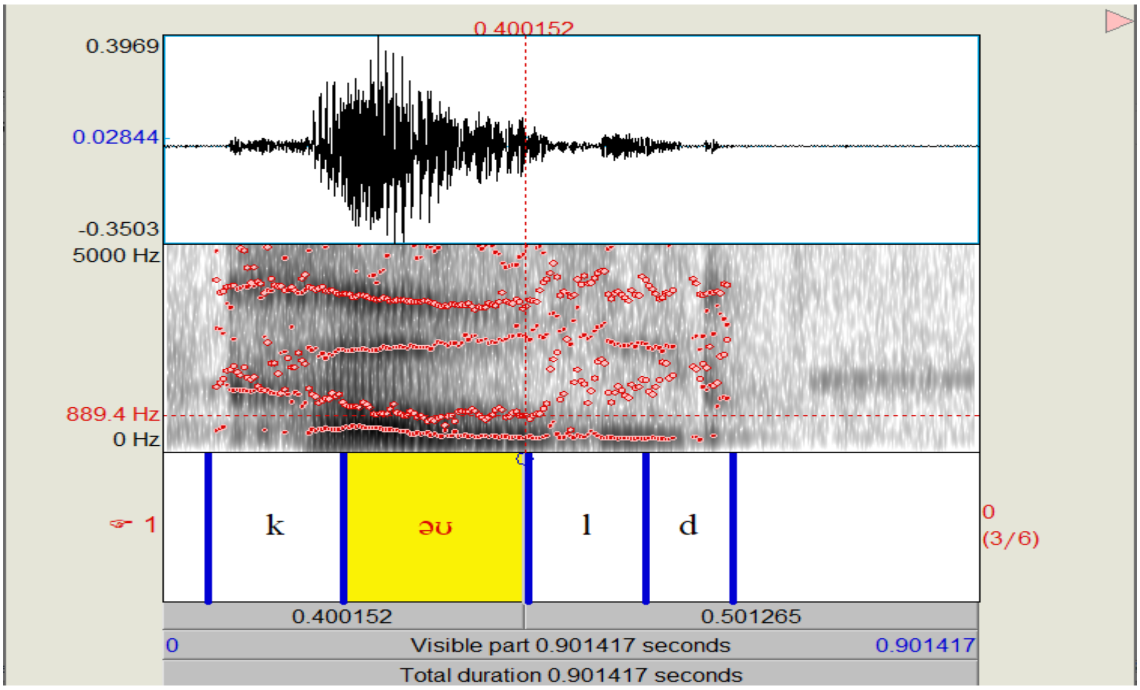

Figure 62: Waveform and spectrographic representation for the word “cold” produced by speaker RP4.

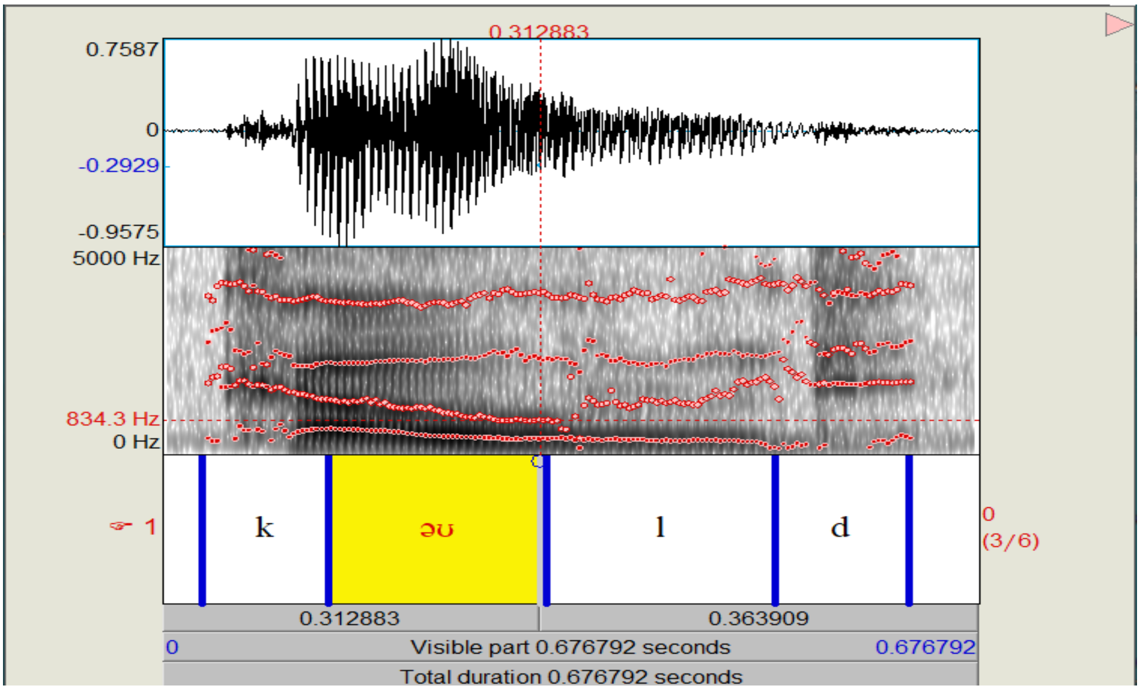

Figure 63: Waveform and spectrographic representation for the word “cold” produced by speaker RP5.

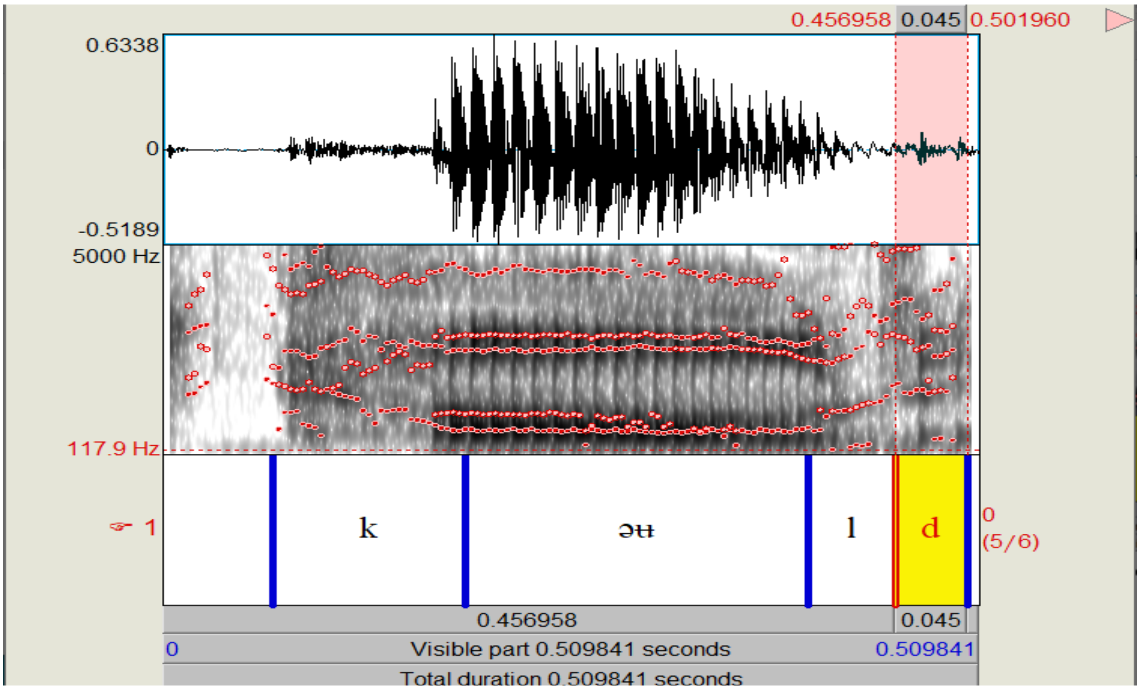

Figure 64: Waveform and spectrographic representation for the word "cold" produced by speaker AE2.

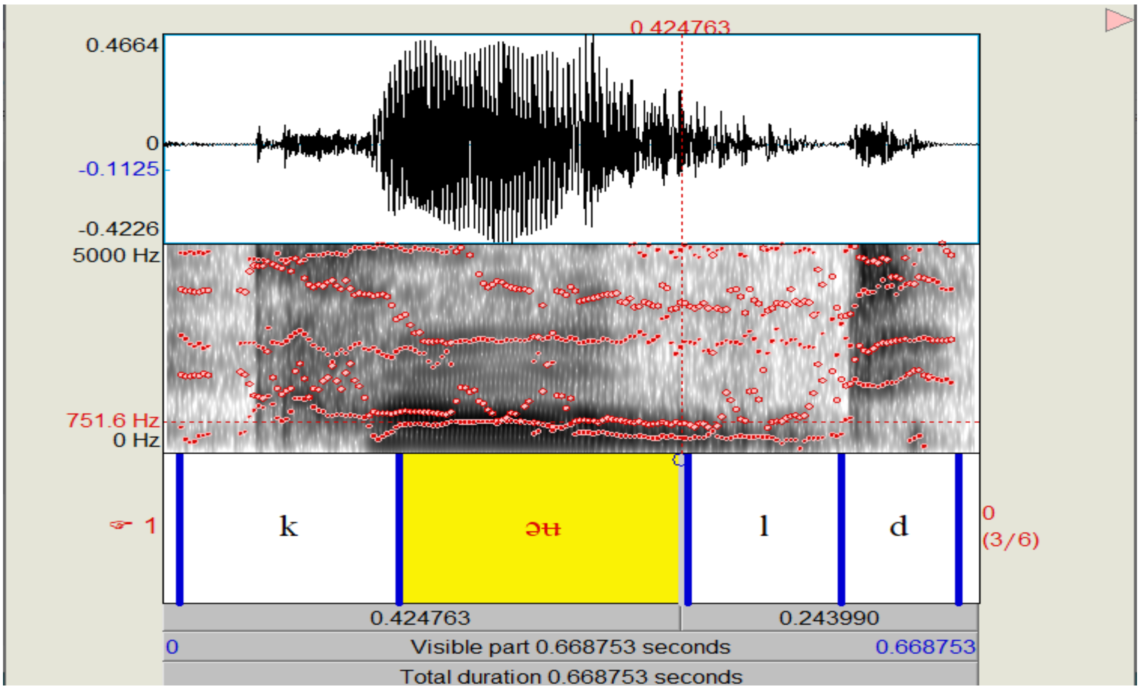

Figure 65: Waveform and spectrographic representation for the word "cold" produced by speaker AE3.

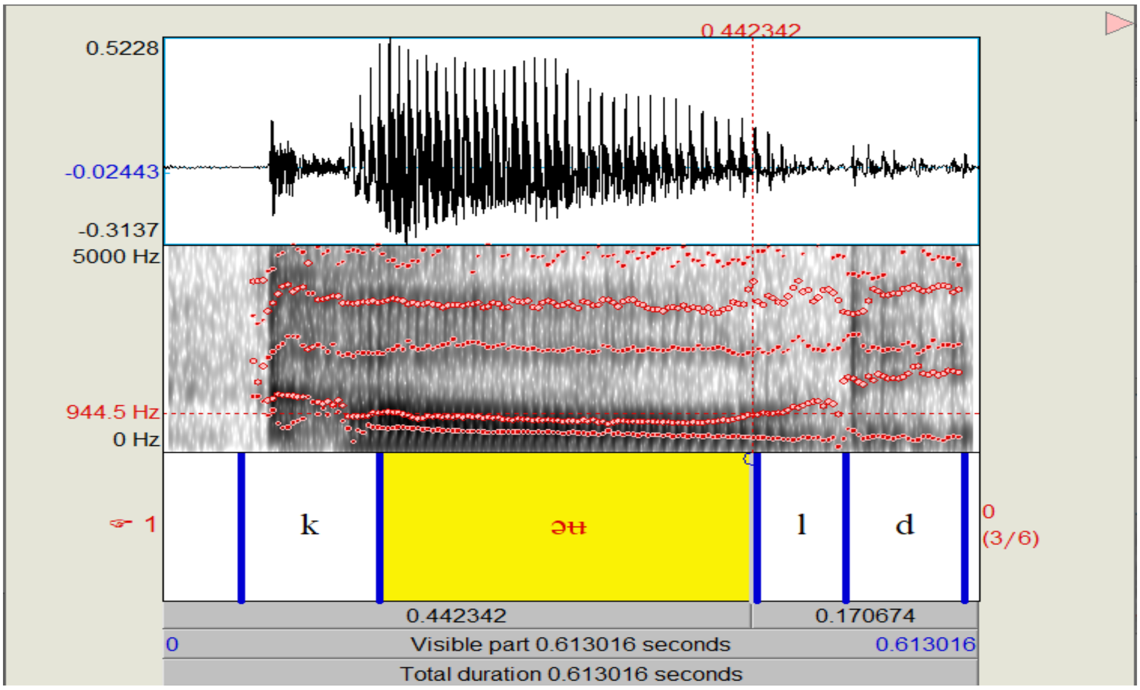

Figure 66: Waveform and spectrographic representation for the word "cold" produced by speaker AE4.

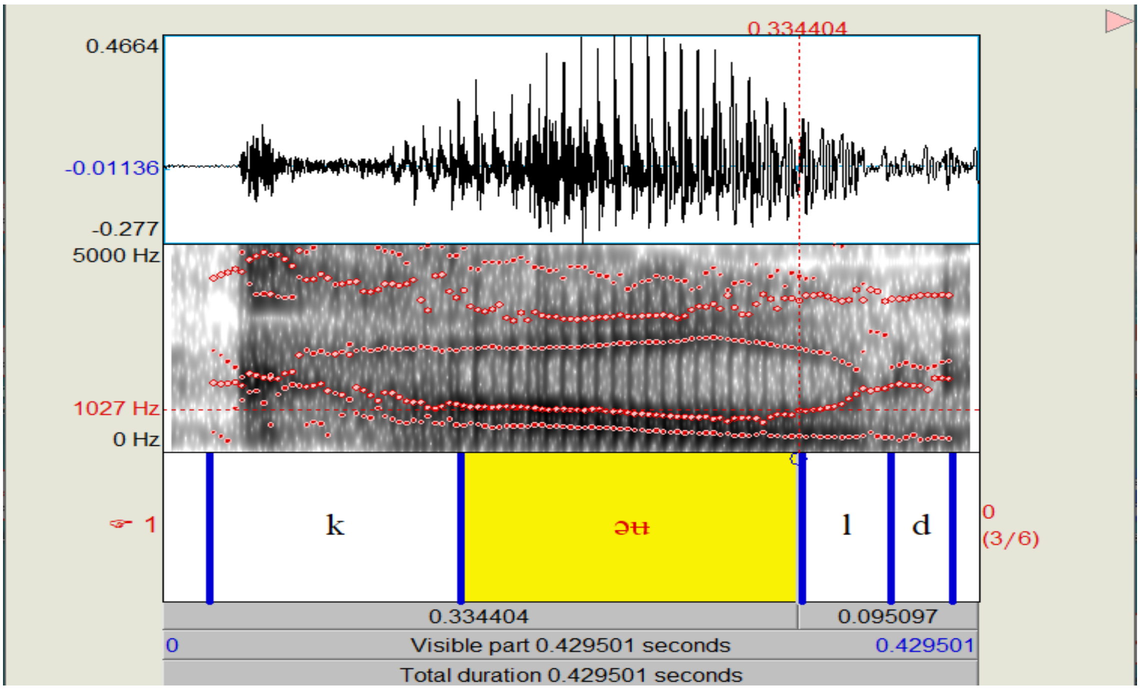

Figure 67: Waveform and spectrographic representation for the word "cold" produced by speaker AE5.

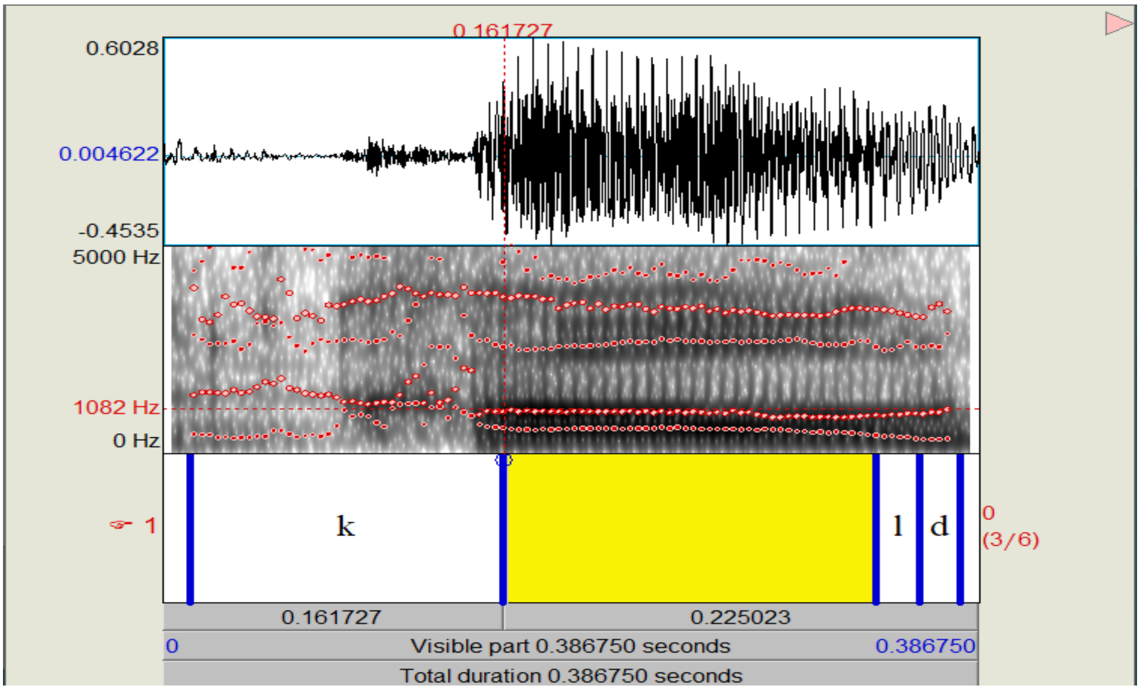

Figure 68: Waveform and spectrographic representation for the word “cold” produced by speaker AE6.

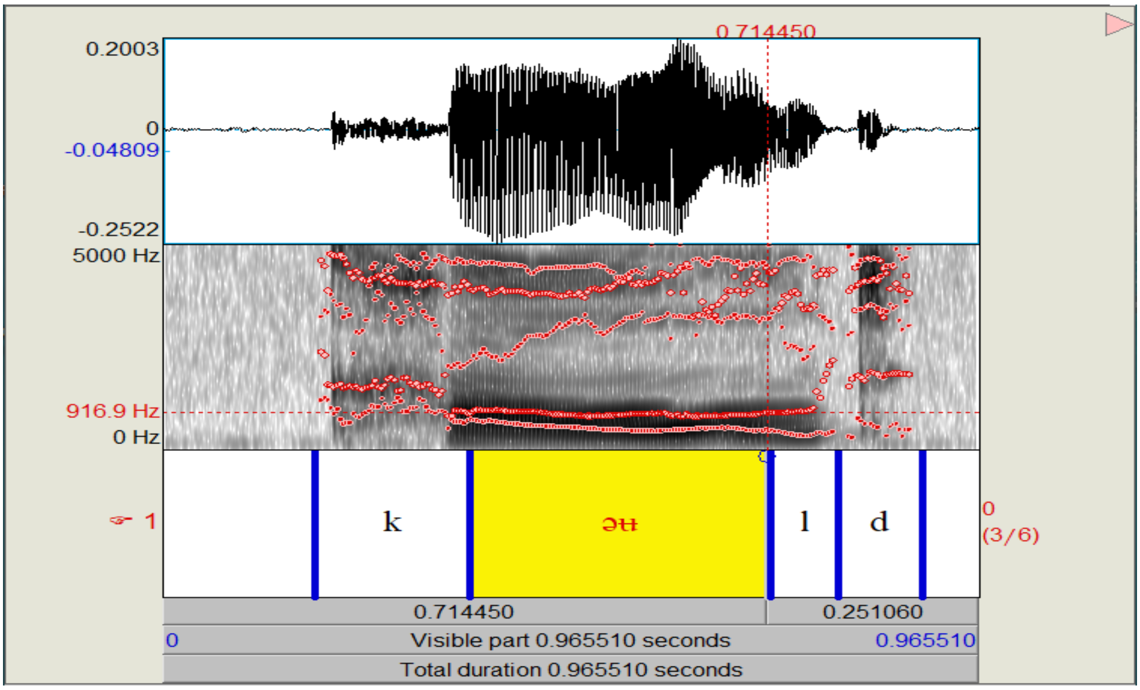

Figure 69: Waveform and spectrographic representation for the word “cold” produced by speaker AE7.

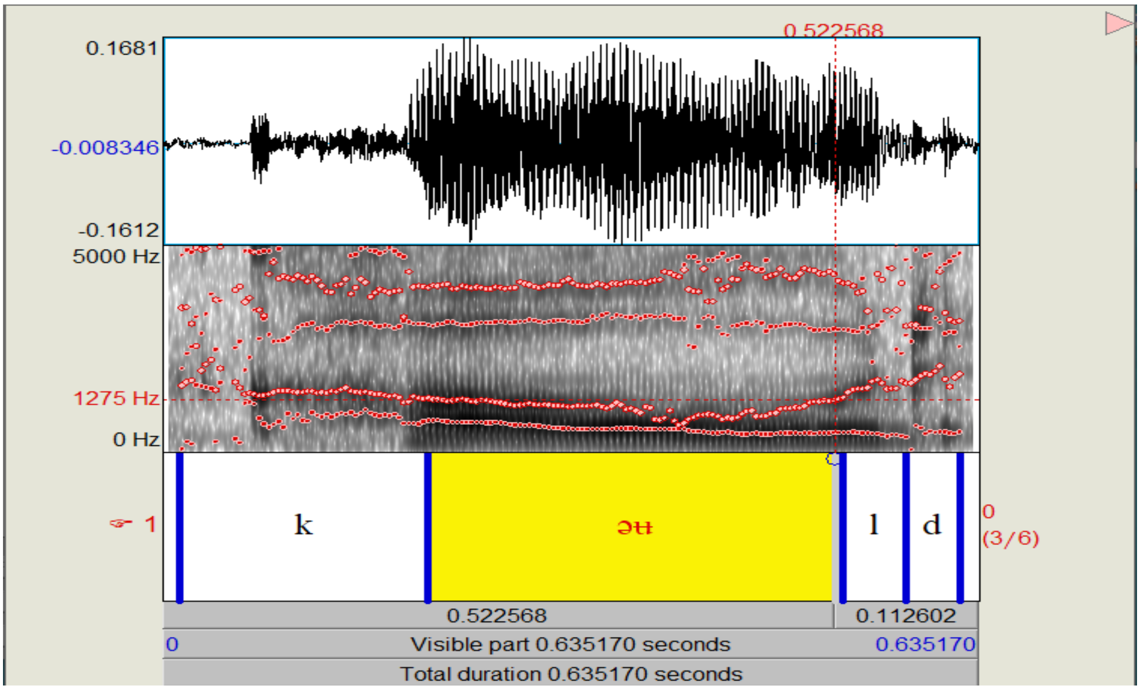

Figure 70: Waveform and spectrographic representation for the word “cold” produced by speaker AE8.

Boat

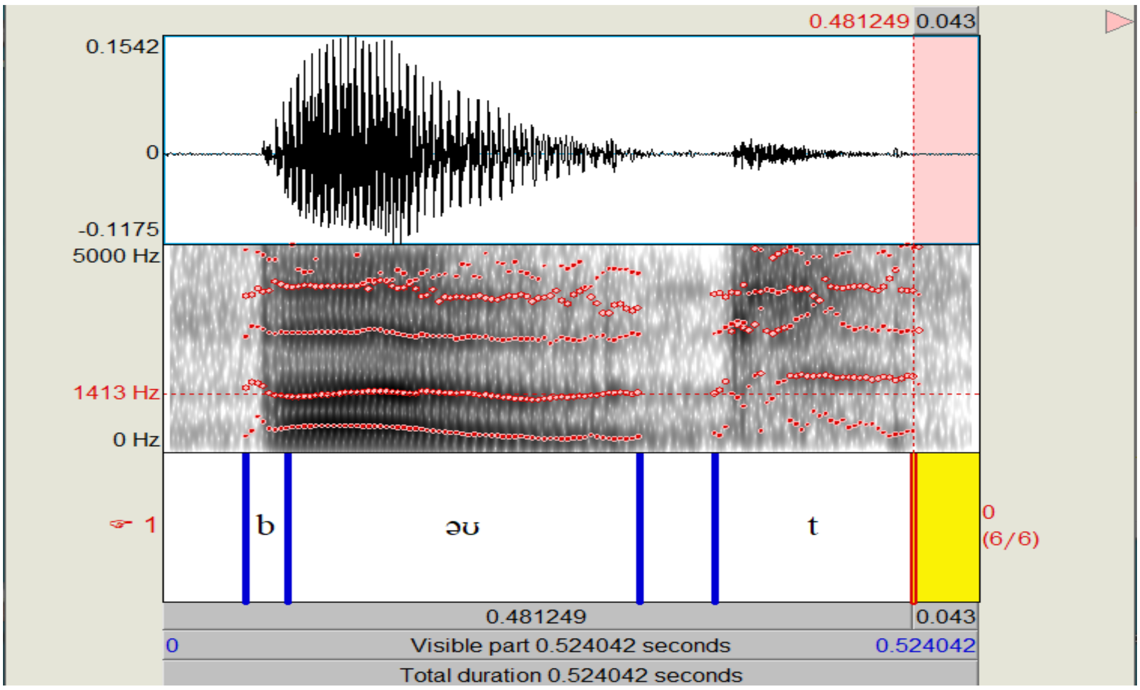

Figure 71: Waveform and spectrographic representation for the word “boat” produced by speaker RP2.

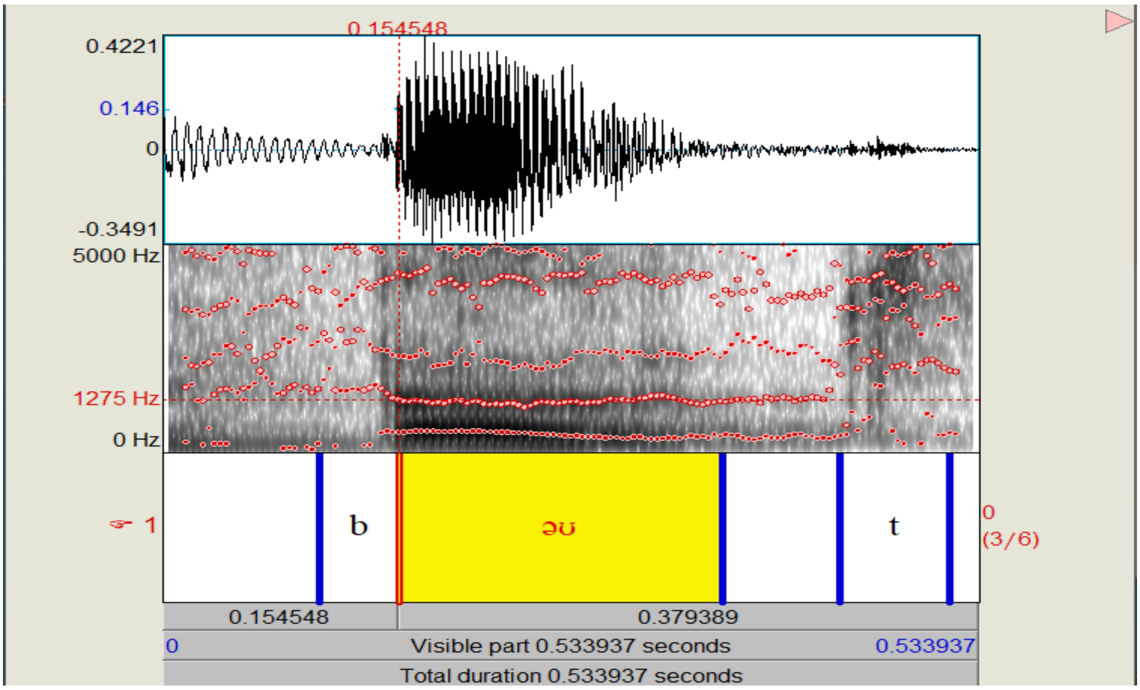

Figure 72: Waveform and spectrographic representation for the word "boat" produced by speaker RP3.

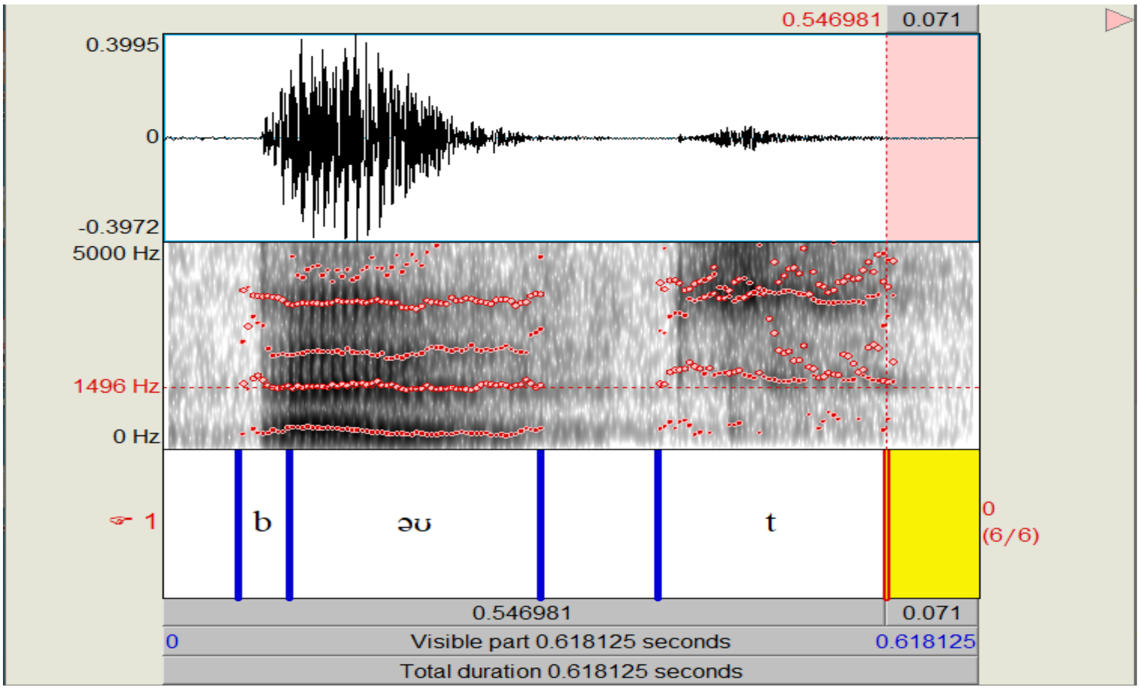

Figure 73: Waveform and spectrographic representation for the word "boat" produced by speaker RP4.

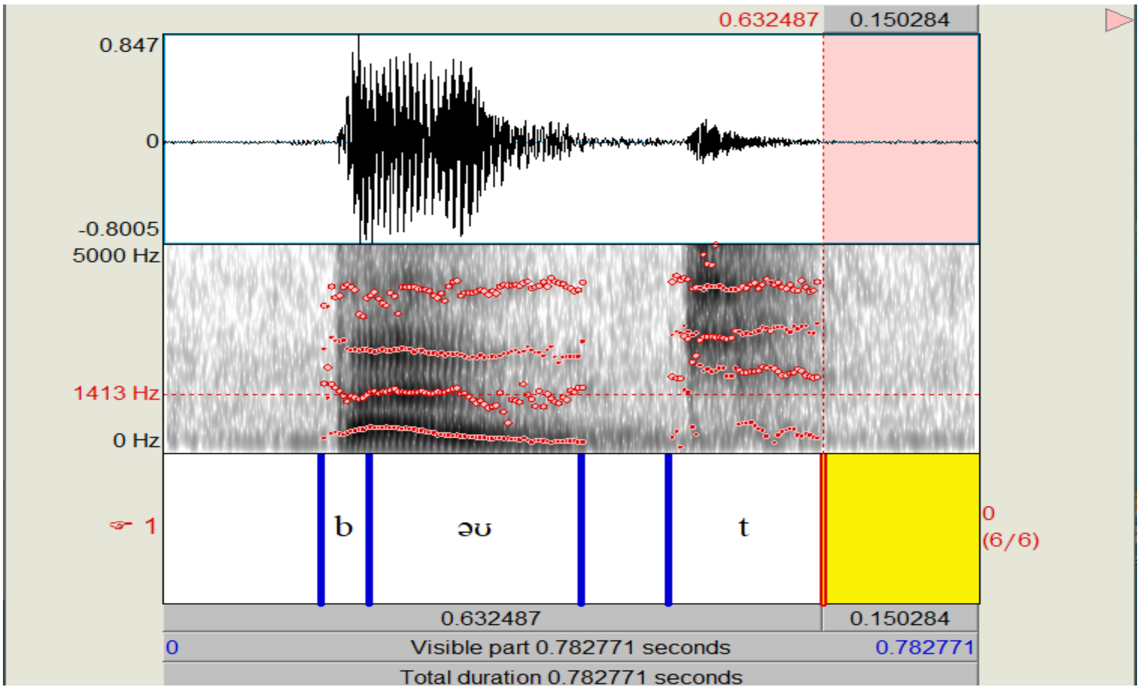

Figure 74: Waveform and spectrographic representation for the word “boat” produced by speaker RP5.

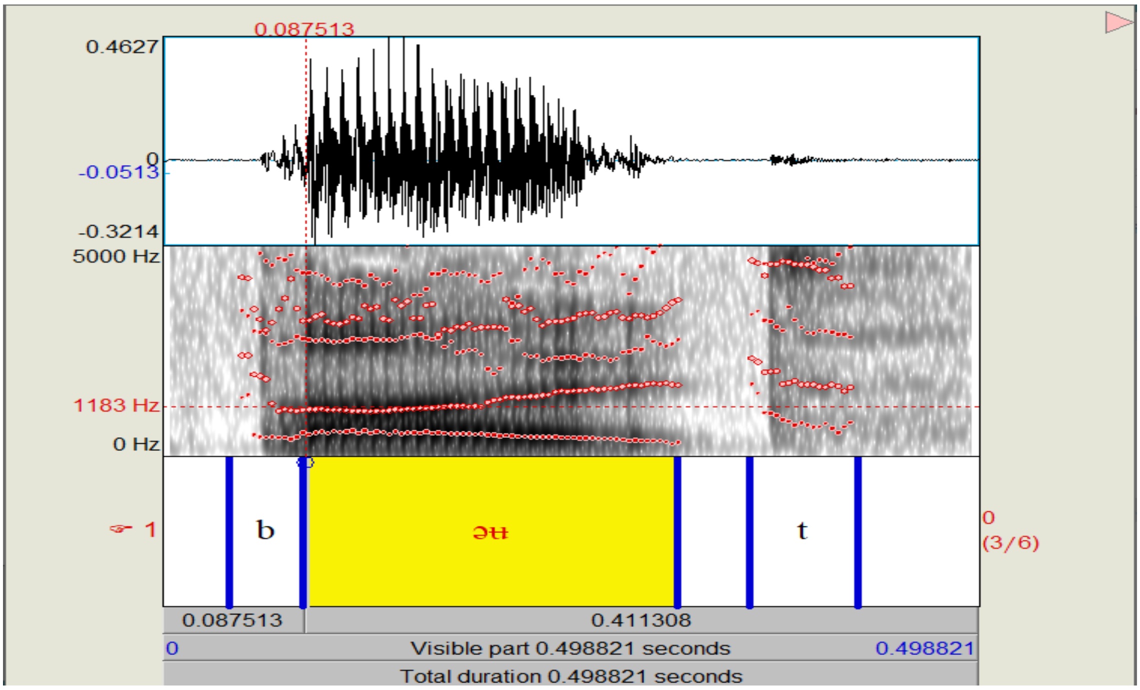

Figure 75: Waveform and spectrographic representation for the word “boat” produced by speaker AE1.

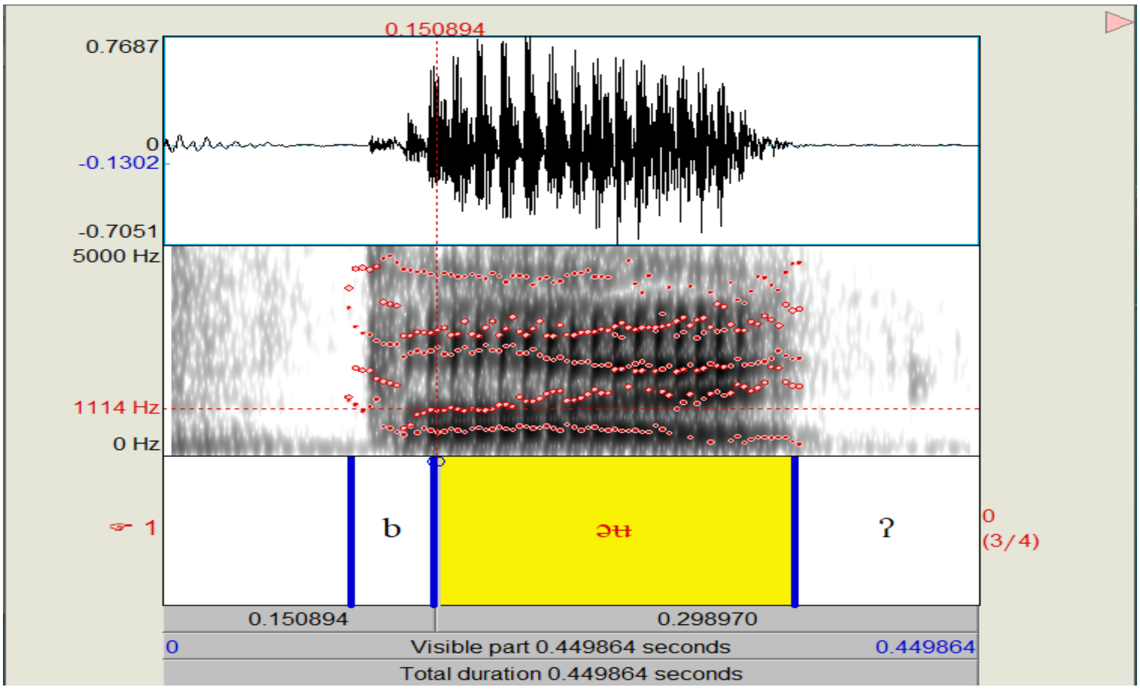

Figure 76: Waveform and spectrographic representation for the word "boat" produced by speaker AE2.

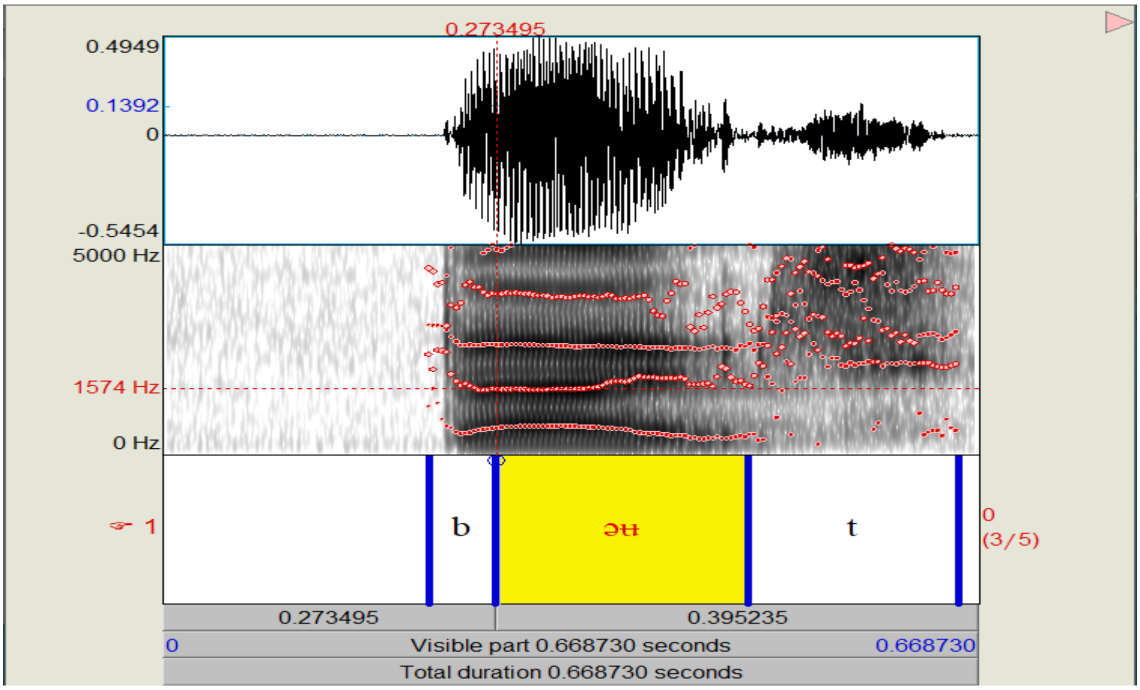

Figure 77: Waveform and spectrographic representation for the word "boat" produced by speaker AE3.

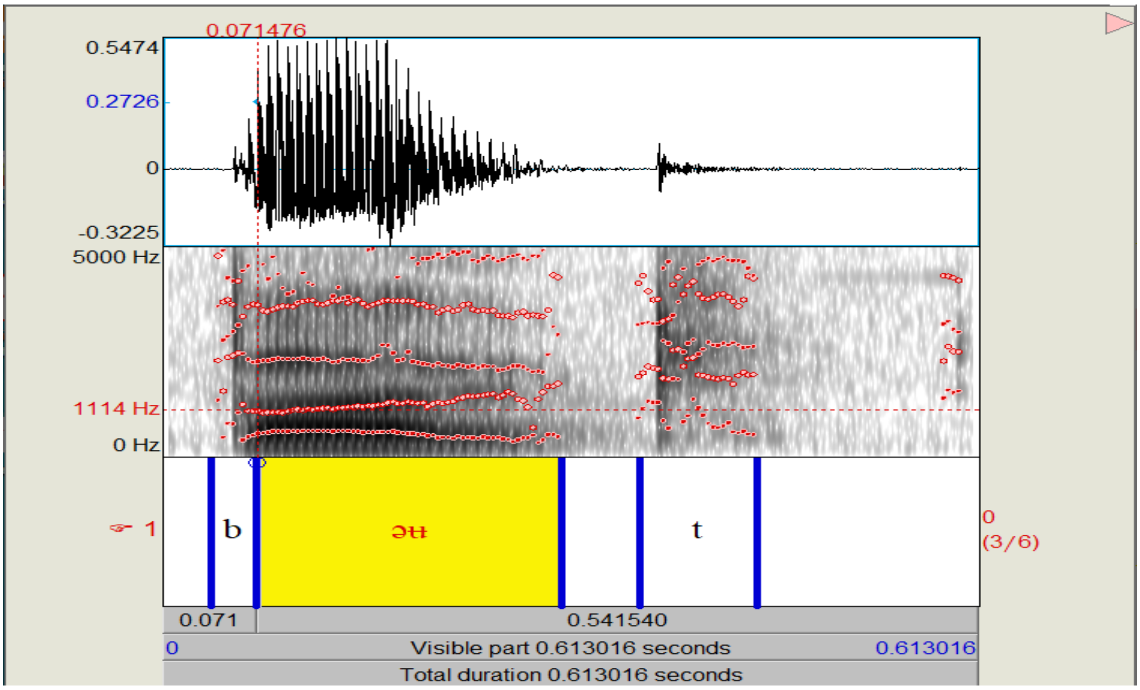

Figure 78: Waveform and spectrographic representation for the word “boat” produced by speaker AE4.

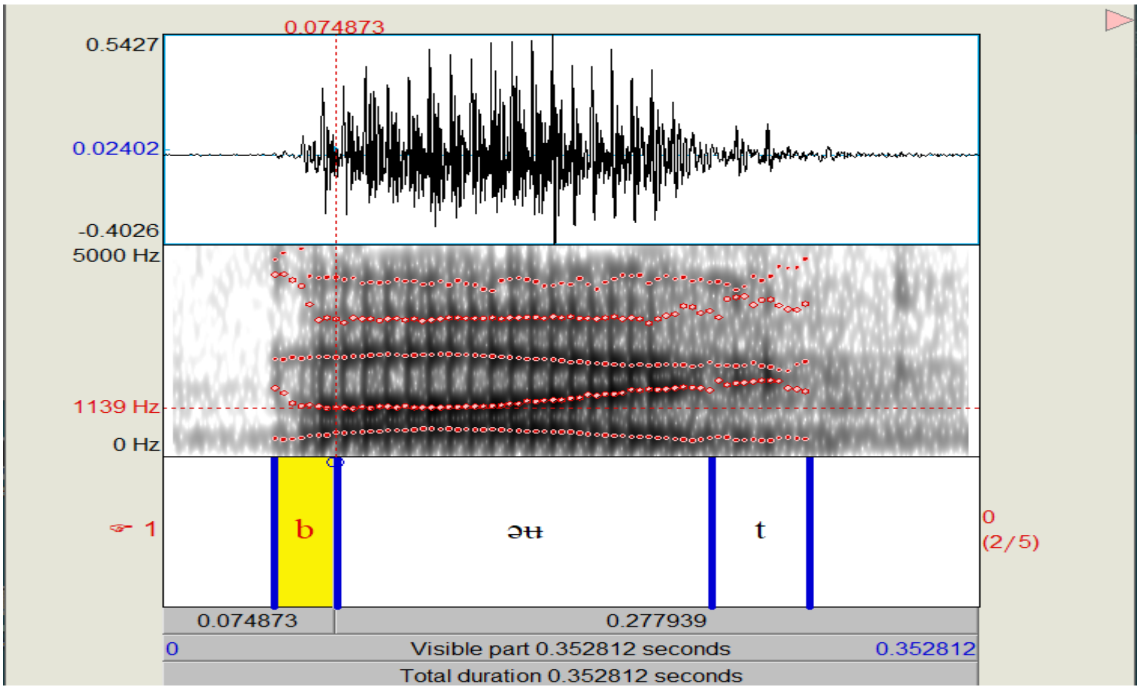

Figure 79: Waveform and spectrographic representation for the word “boat” produced by speaker AE5.

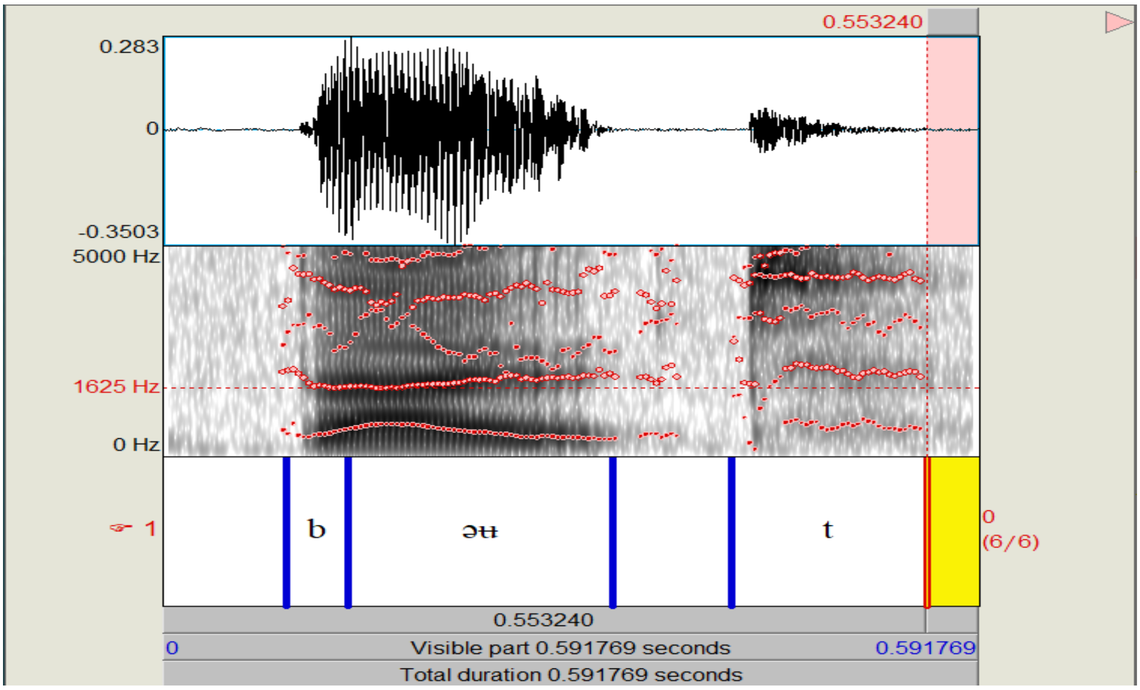

Figure 80: Waveform and spectrographic representation for the word "boat" produced by speaker AE7.

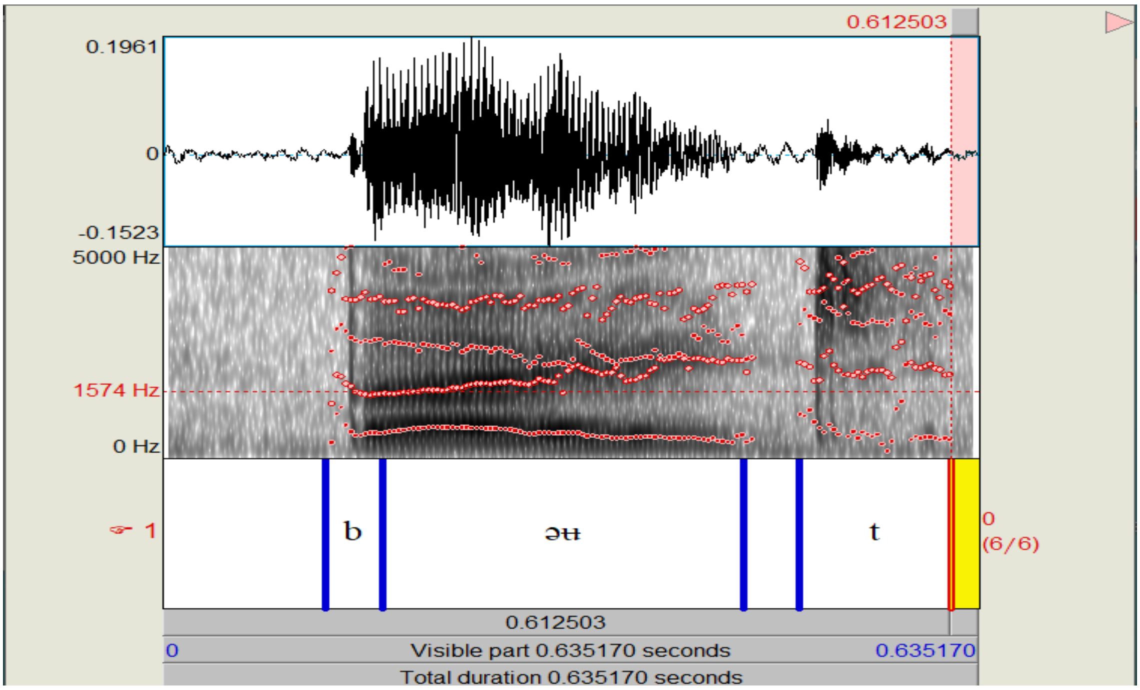

Figure 81: Waveform and spectrographic representation for the word "boat" produced by speaker AE8.

Mouth

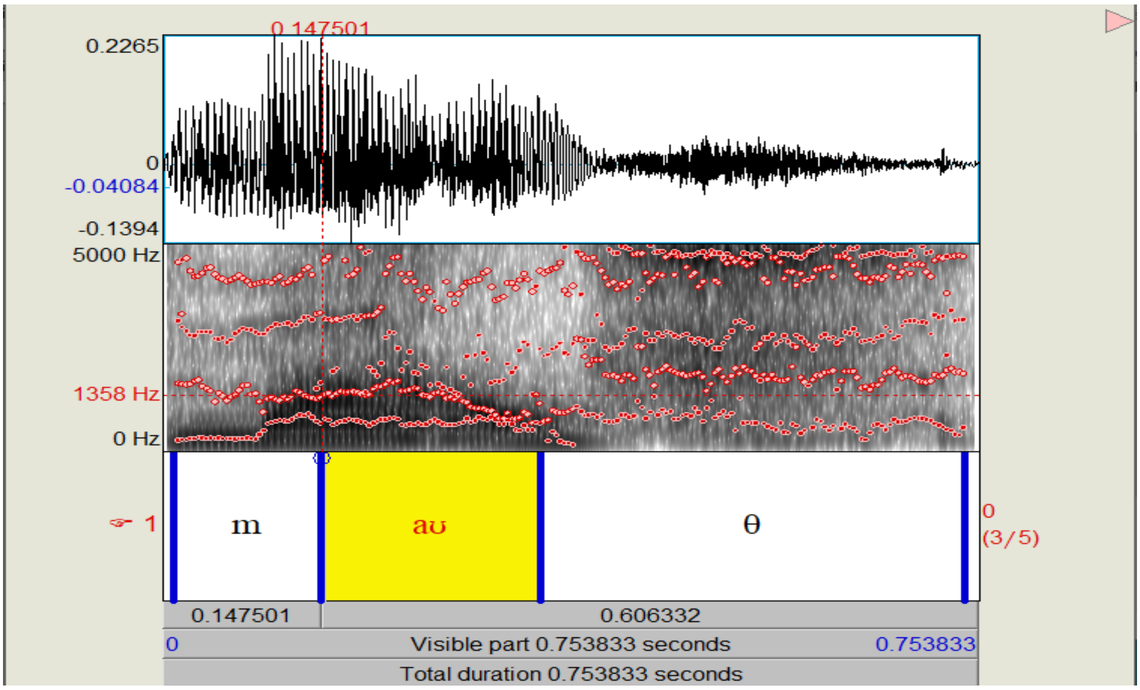

Figure 82: Waveform and spectrographic representation for the word “mouth” produced by speaker RP1.

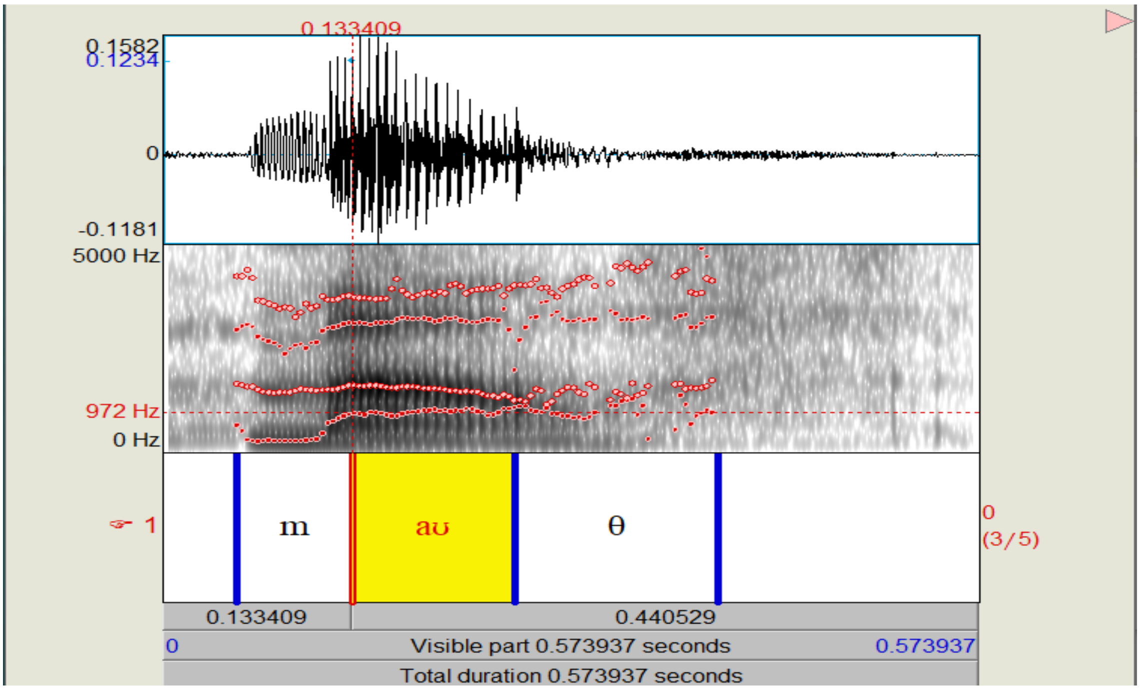

Figure 83: Waveform and spectrographic representation for the word “mouth” produced by speaker RP2.

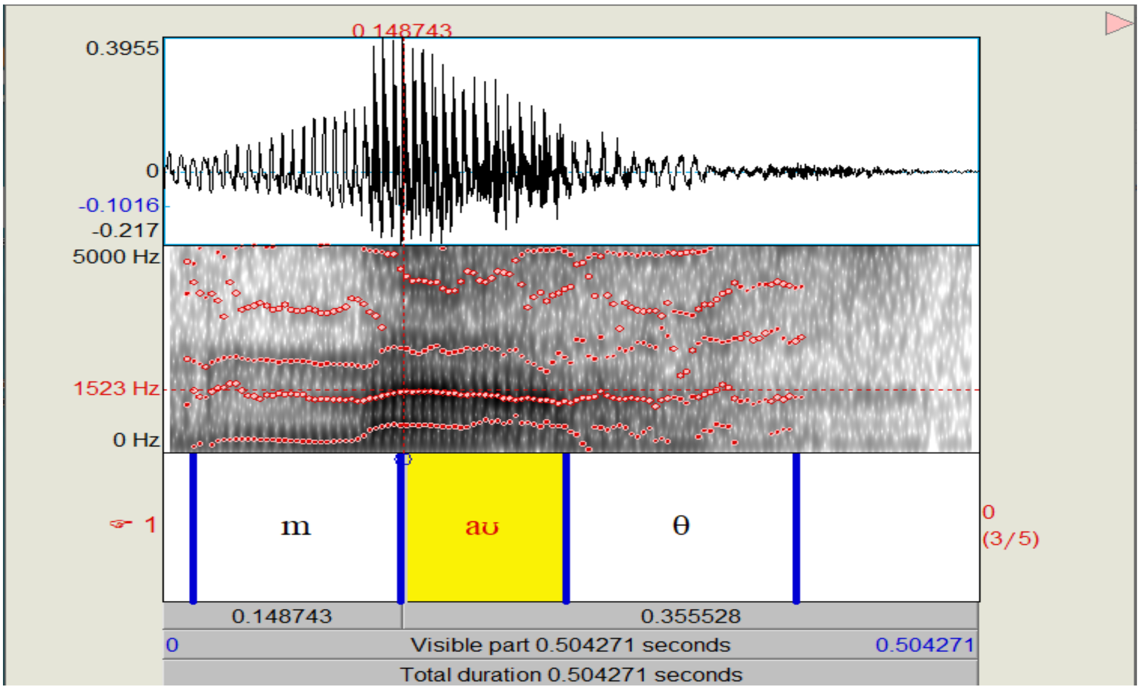

Figure 84: Waveform and spectrographic representation for the word "mouth" produced by speaker RP3.

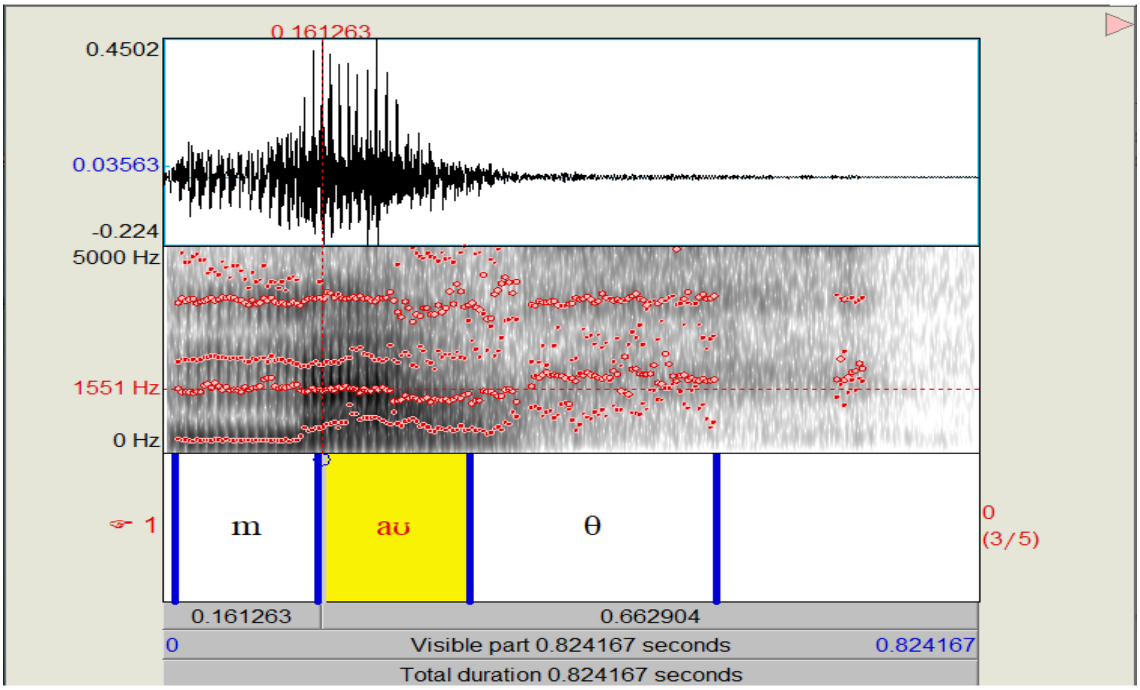

Figure 85: Waveform and spectrographic representation for the word "mouth" produced by speaker RP4.

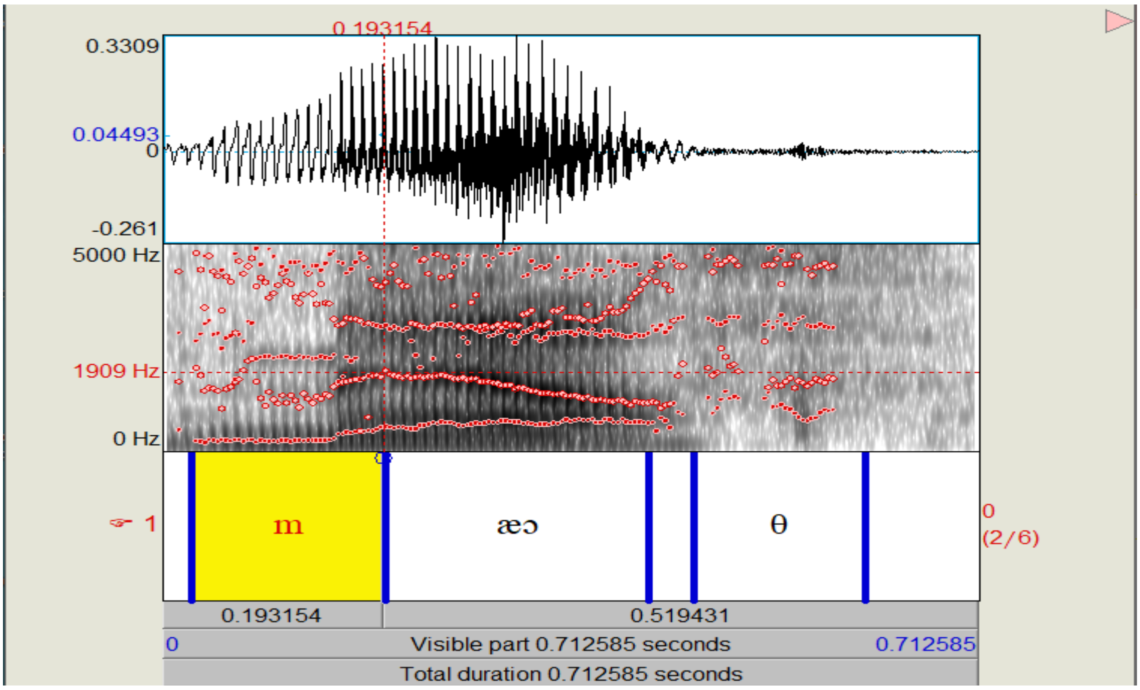

Figure 86: Waveform and spectrographic representation for the word “mouth” produced by speaker AE1.

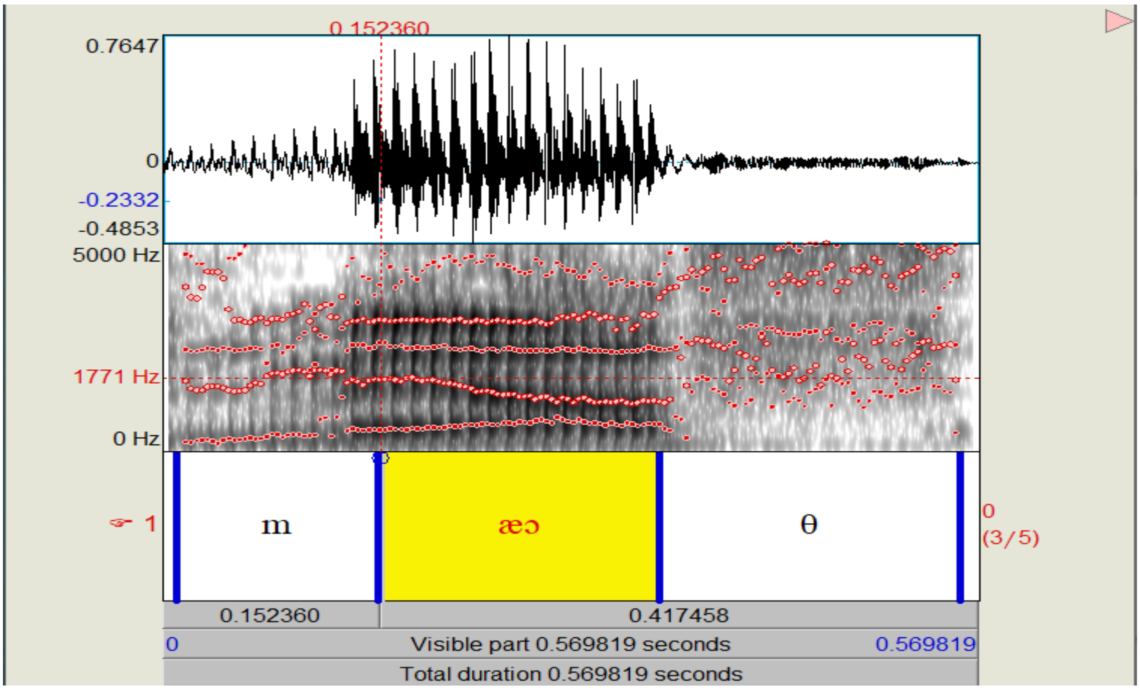

Figure 87: Waveform and spectrographic representation for the word “mouth” produced by speaker AE2.

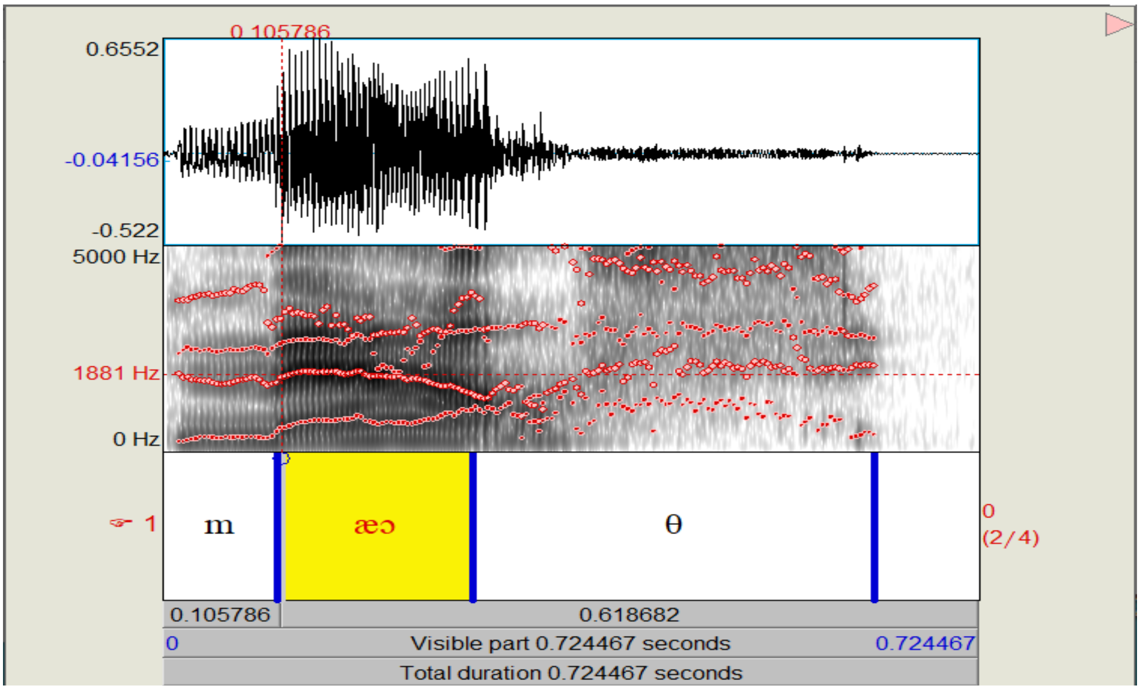

Figure 88: Waveform and spectrographic representation for the word “mouth” produced by speaker AE3.

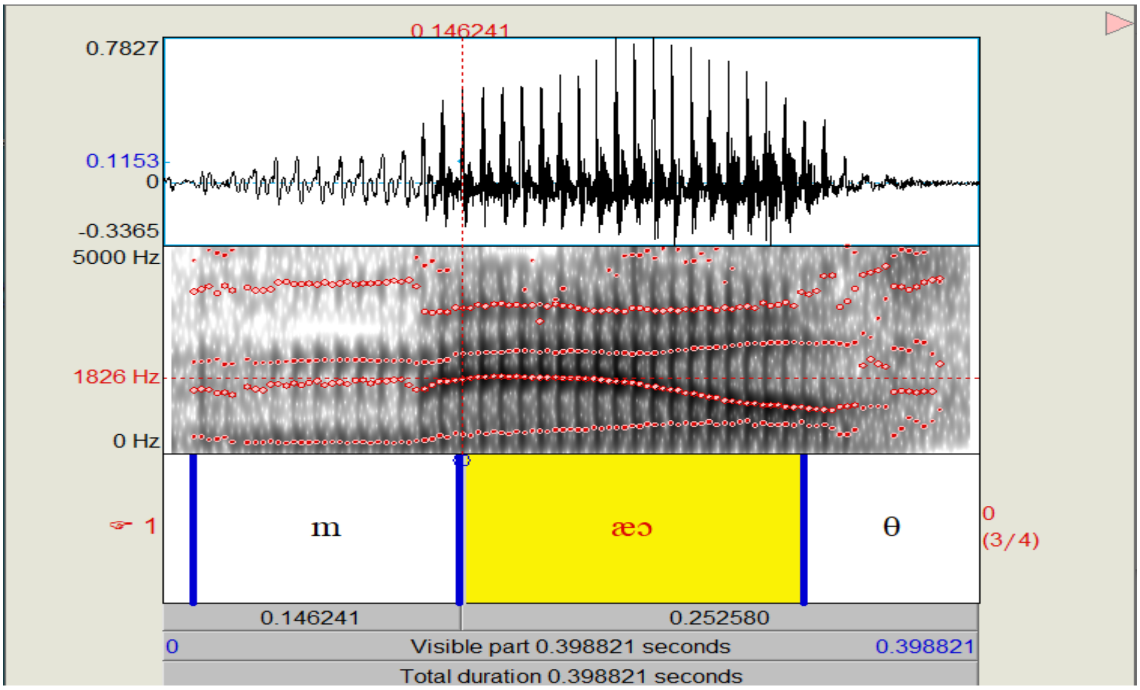

Figure 89: Waveform and spectrographic representation for the word “mouth” produced by speaker AE5.

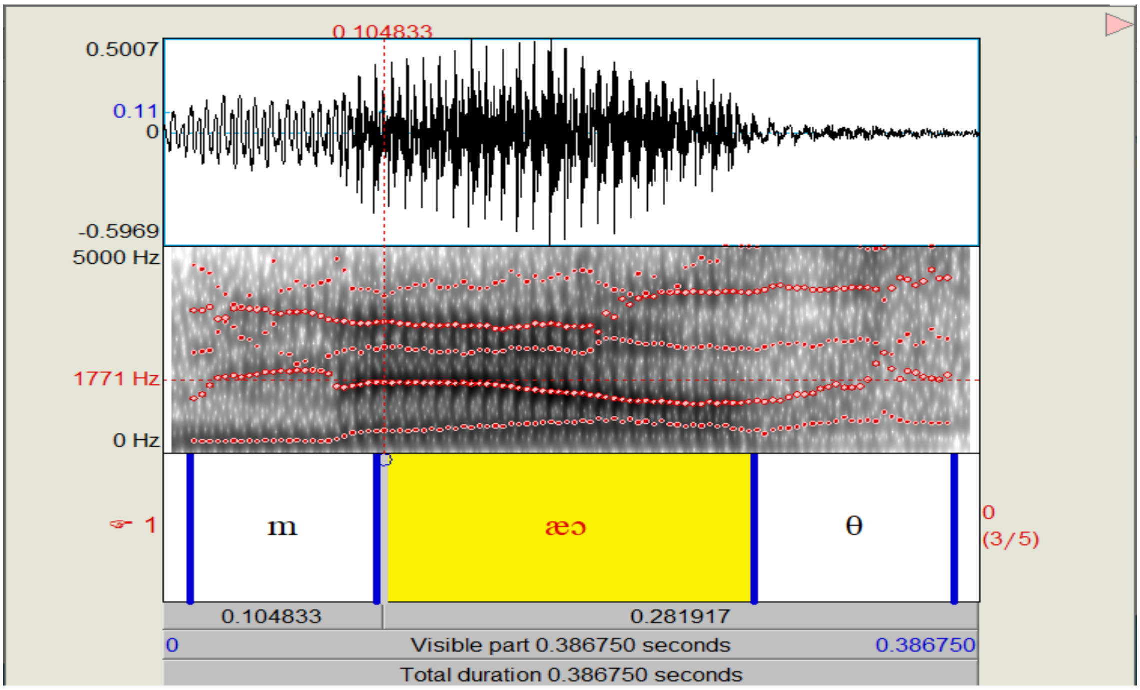

Figure 90: Waveform and spectrographic representation for the word "mouth" produced by speaker AE6.

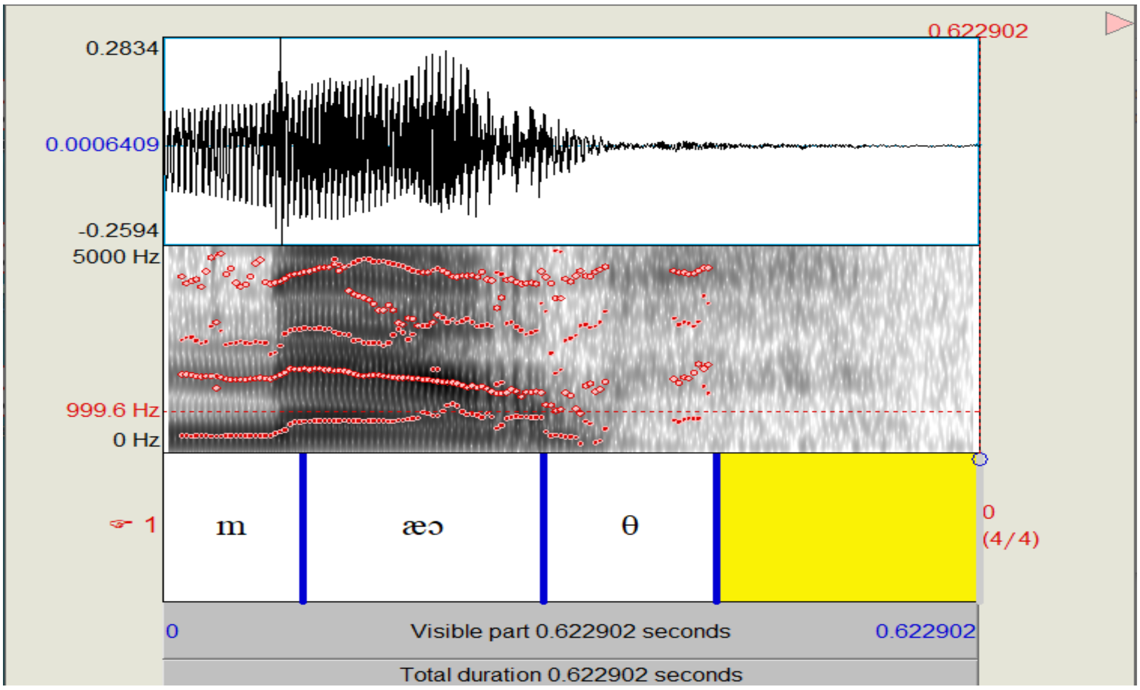

Figure 91: Waveform and spectrographic representation for the word "mouth" produced by speaker AE7.

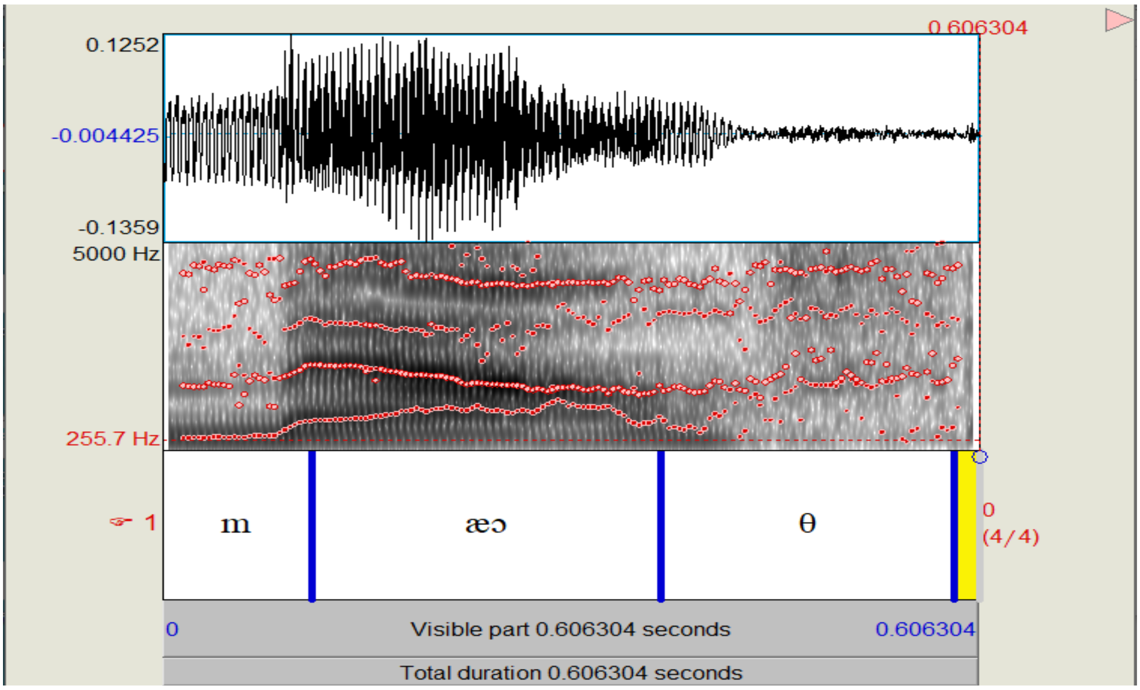

Figure 92: Waveform and spectrographic representation for the word “mouth” produced by speaker AE8.

Howl

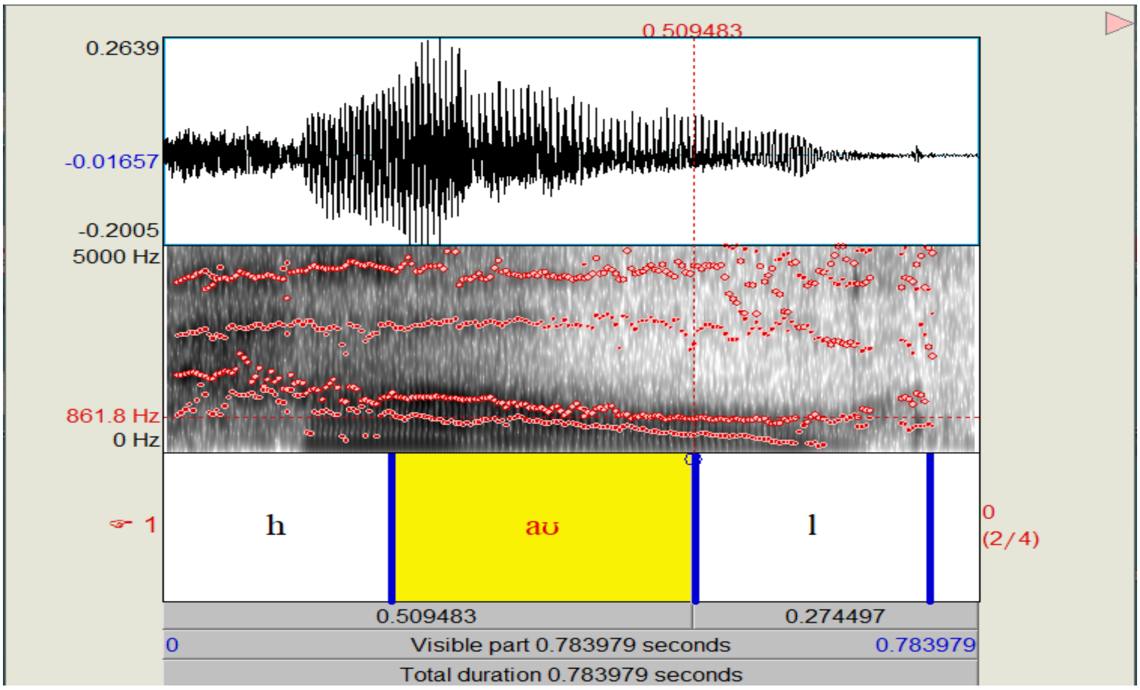

Figure 93: Waveform and spectrographic representation for the word “howl” produced by speaker RP1.

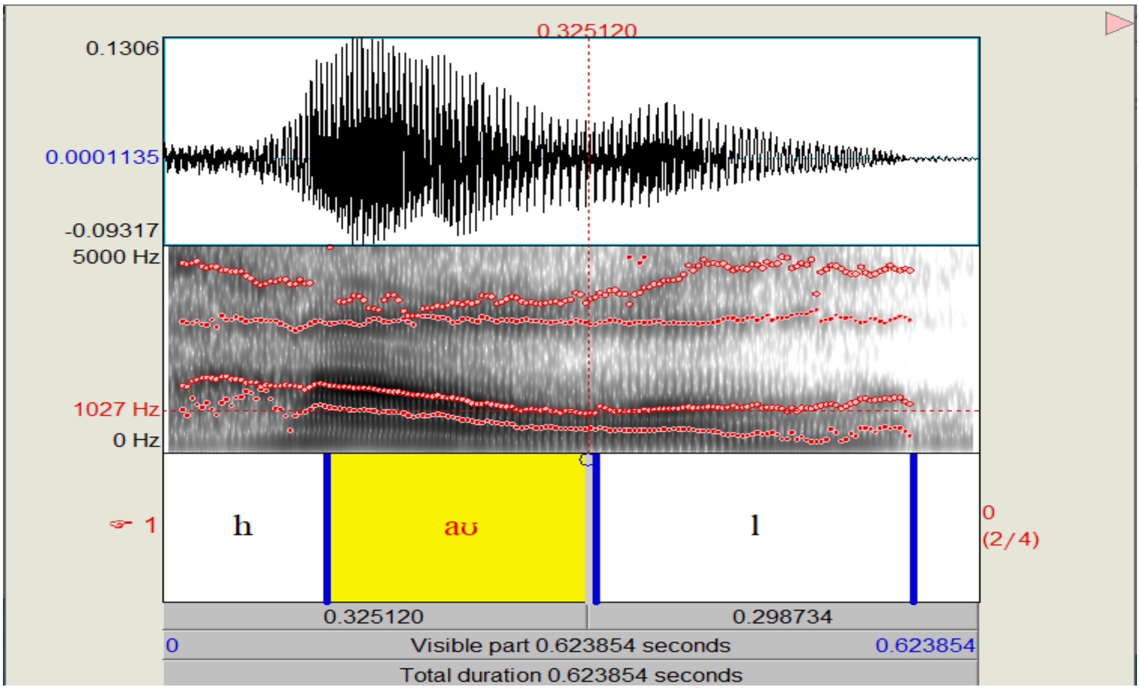

Figure 94: Waveform and spectrographic representation for the word “howl” produced by speaker RP2.

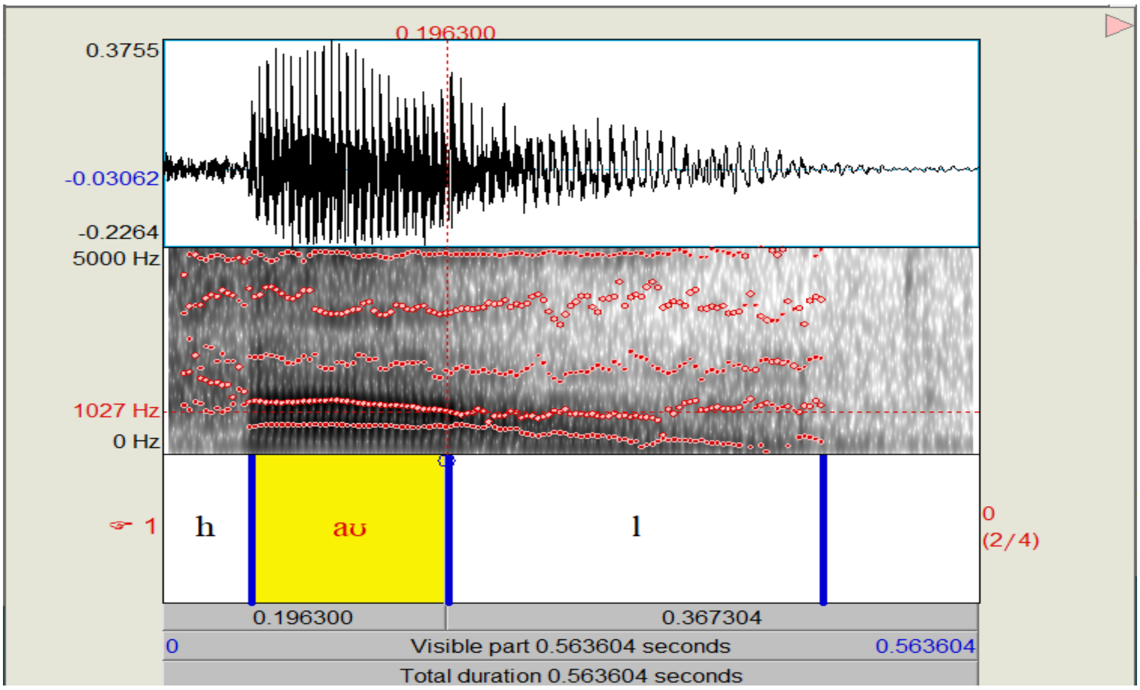

Figure 95: Waveform and spectrographic representation for the word “howl” produced by speaker RP3.

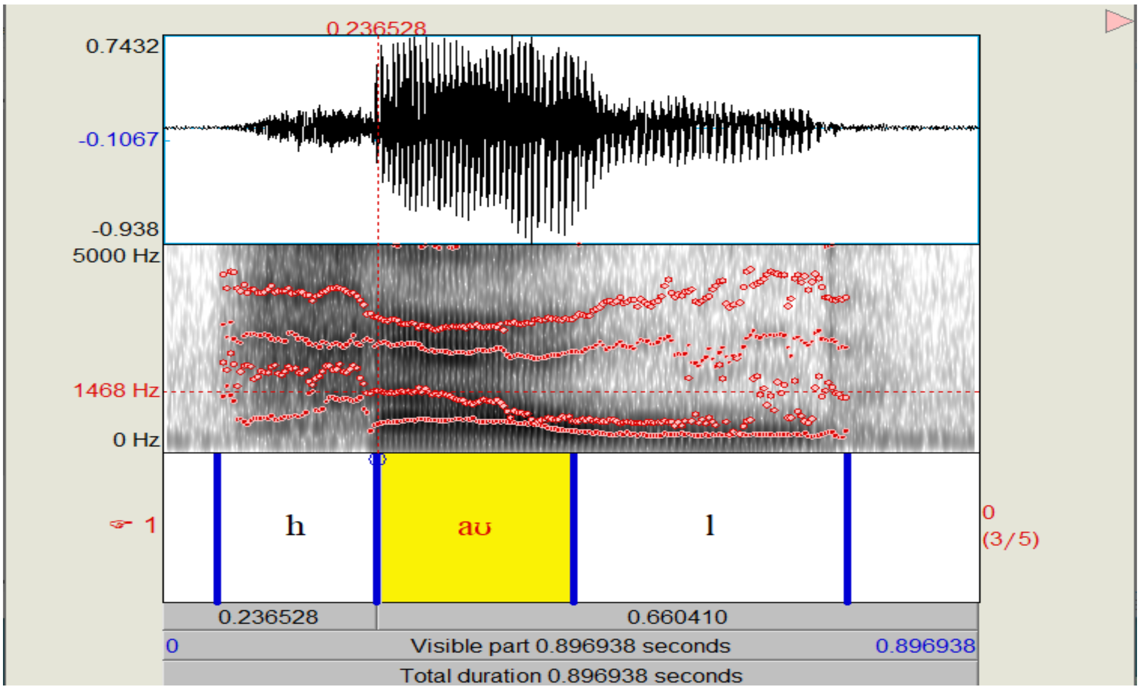

Figure 96: Waveform and spectrographic representation for the word “howl” produced by speaker RP5.

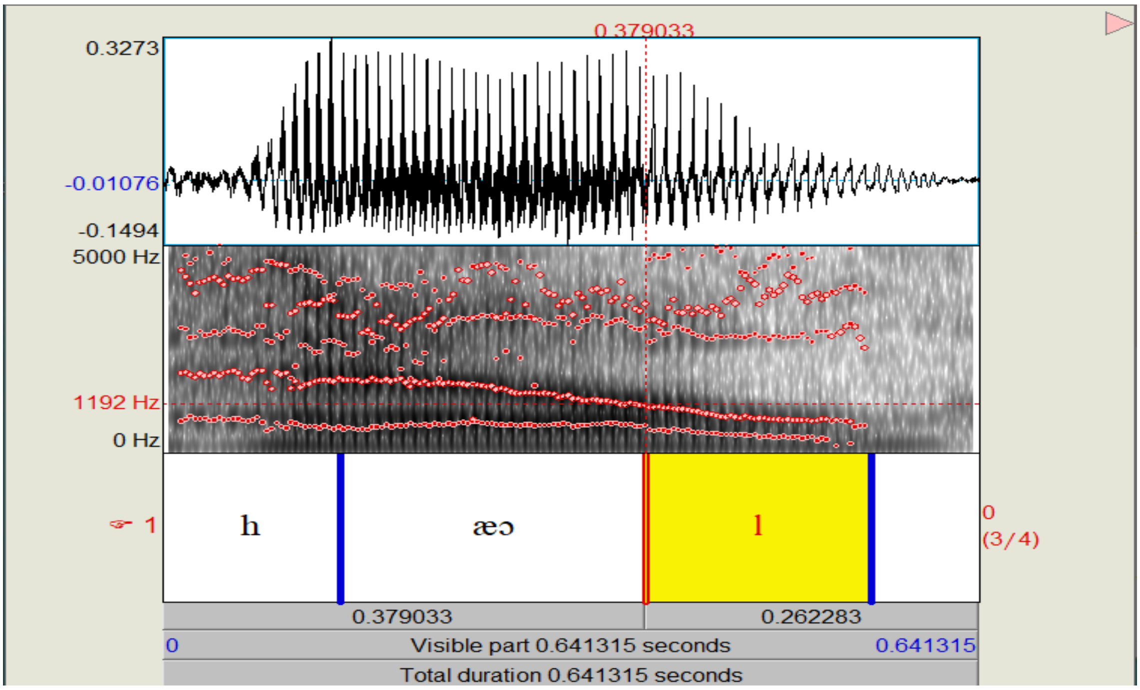

Figure 97: Waveform and spectrographic representation for the word “howl” produced by speaker AE1.

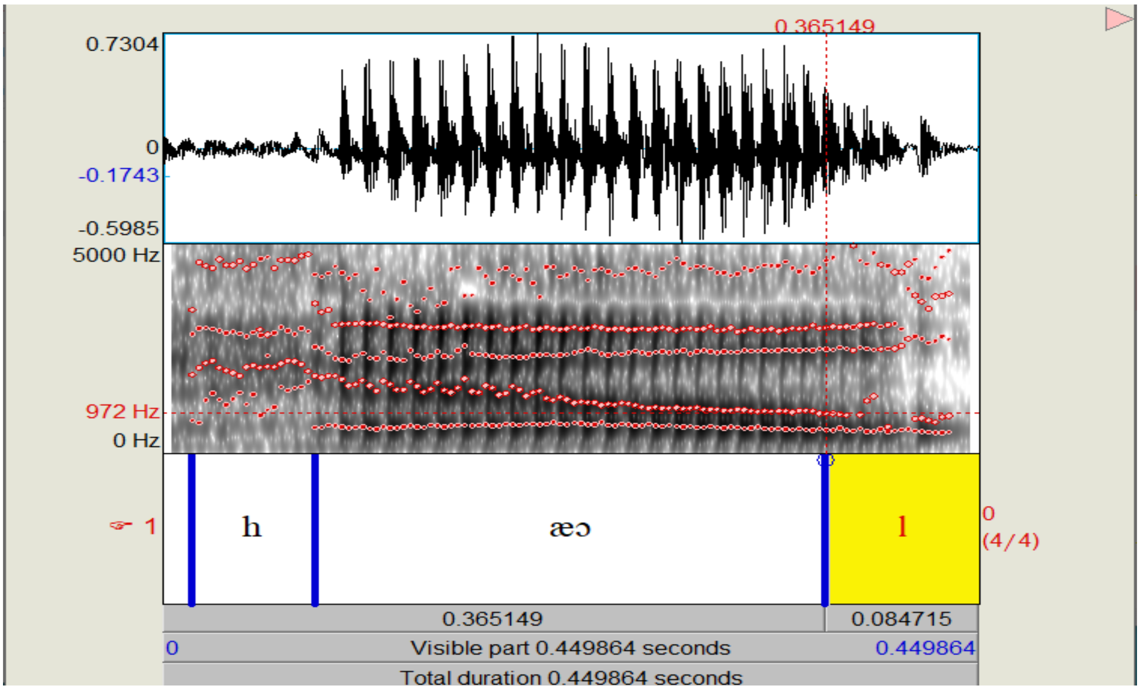

Figure 98: Waveform and spectrographic representation for the word "howl" produced by speaker AE2.

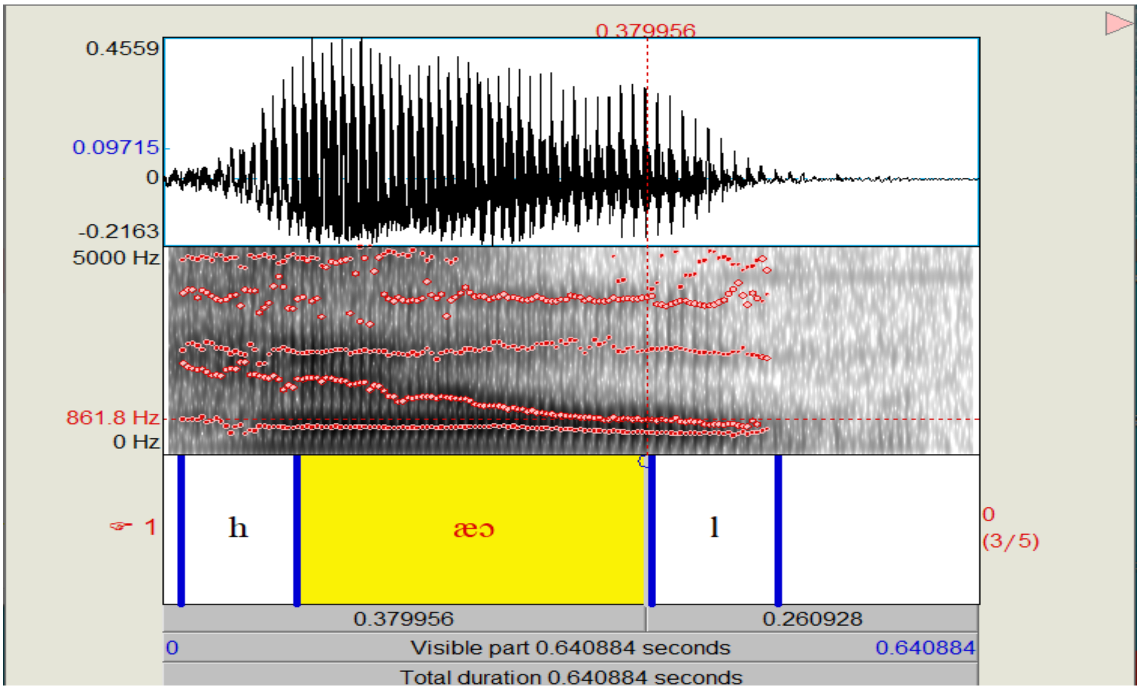

Figure 99: Waveform and spectrographic representation for the word "howl" produced by speaker AE4.

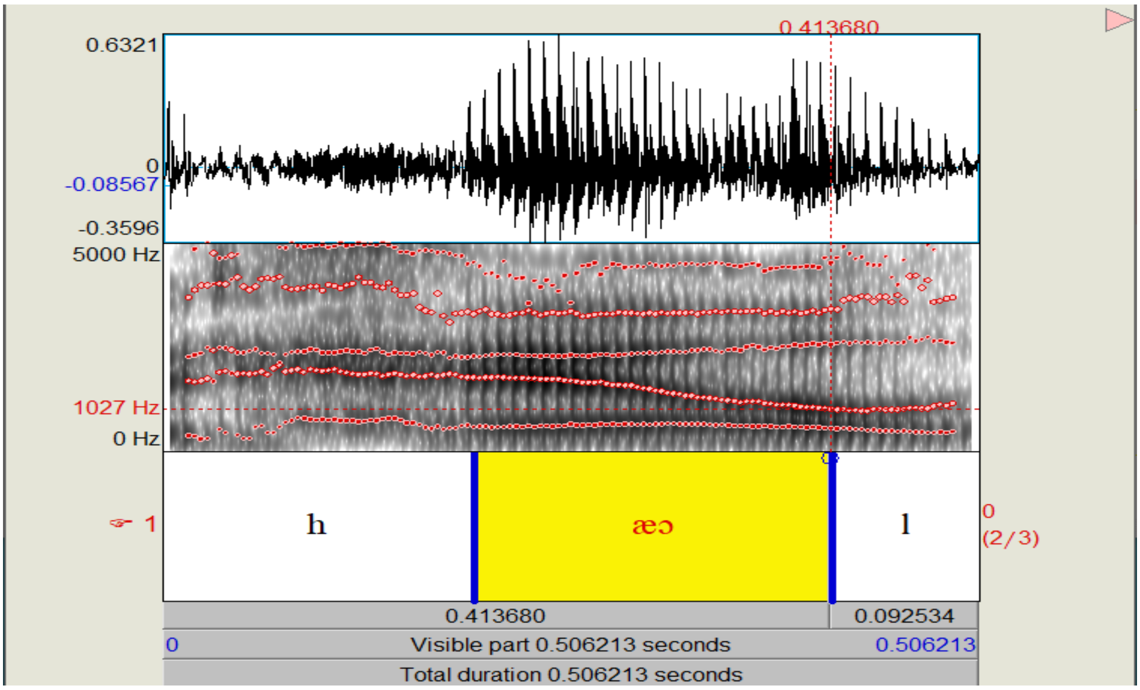

Figure 100: Waveform and spectrographic representation for the word "howl" produced by speaker AE5.

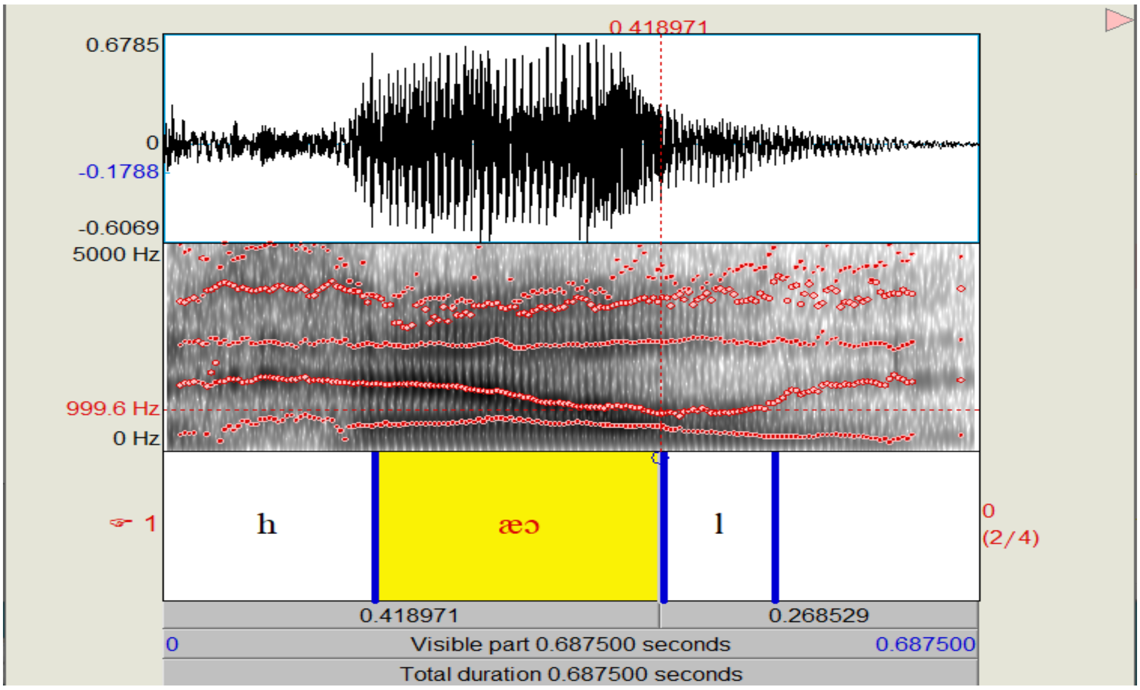

Figure 101: Waveform and spectrographic representation for the word "howl" produced by speaker AE6.

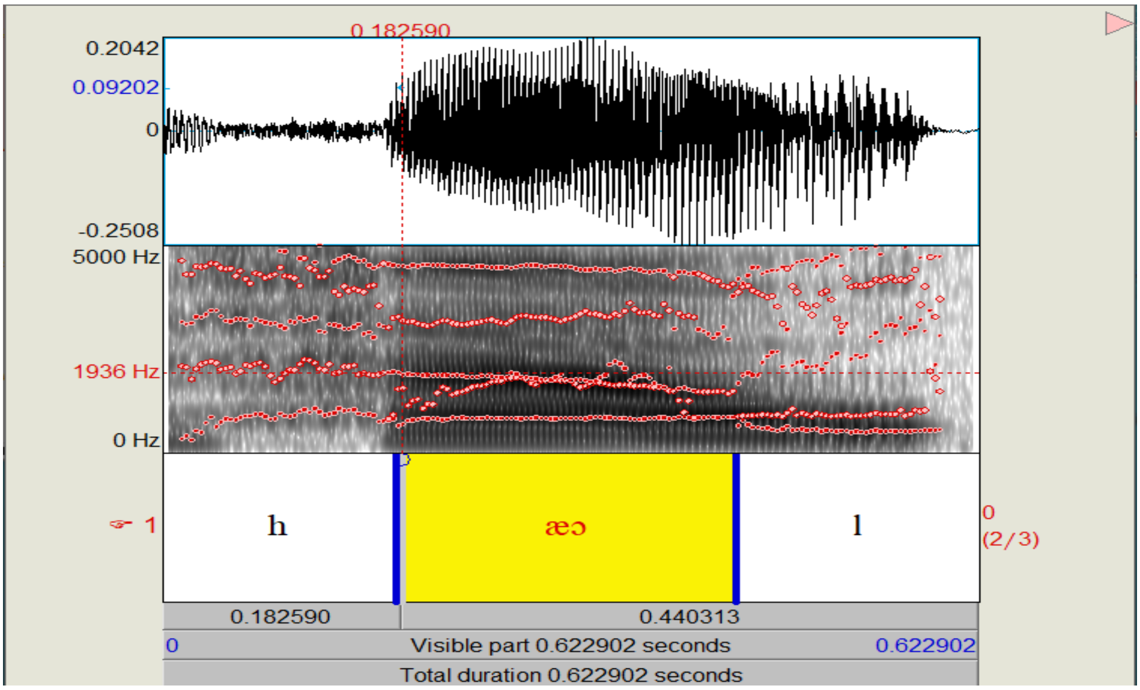

Figure 102: Waveform and spectrographic representation for the word “howl” produced by speaker AE7.

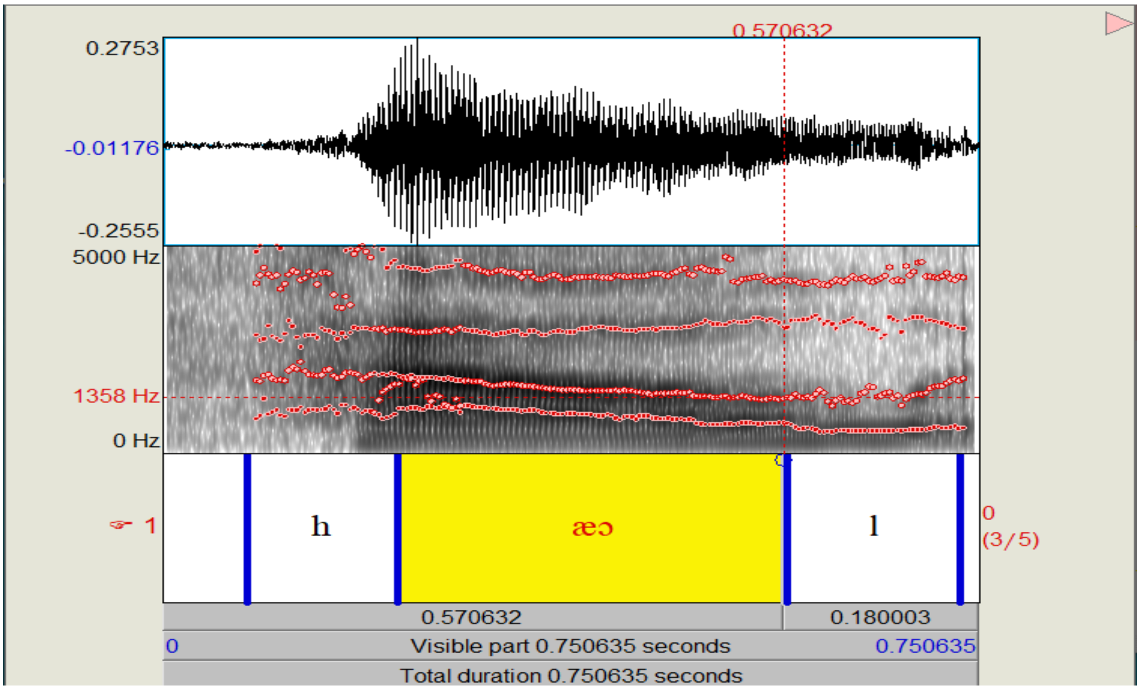

Figure 103: Waveform and spectrographic representation for the word “howl” produced by speaker AE8.

6.3. Recordings

rp1.wav

rp2.wav

rp3.wav

rp4.wav

rp5.wav

au1.wav

au2.wav

au3.wav

au4.wav

au5 permission.wav

au5 sentences.wav

au6.wav

au7.wav

au8.wav