Forum for Linguistic Studies 2023; 5(3): 1787.
Original Article

Instrumental analysis of loss of assimilation of English nasal stops in adult Saudi EFL learners' speech

Pradeep Sharma[1], Mahboobeh Khaleghi[2], Mohammad Saleem[3,*]

[1] *Independent Researcher, Vancouver, B. C., V6K 2E8, Canada*
[2] *Independent Researcher, Rabigh 25754, Saudi Arabia*
[3] *King Abdulaziz University, Rabigh 21911, Saudi Arabia*
*** Corresponding author:** Mohammad Saleem, msshareef@kau.edu.sa

ARTICLE INFO

Received: 3 June 2023
Accepted: 31 August 2023
Available online: 8 October 2023

doi: 10.59400/fls.v5i3.1787

ABSTRACT: Assimilation of consonant sounds within words and at word boundaries is a common phenomenon in English. However, EFL (English as a Foreign Language) learners may fail to observe the rule which makes their English sound sub-standard. This paper investigates whether adult Saudi EFL learners follow assimilation rules in English. Selected words and phrases spoken by a group of participants, where assimilated nasal sounds were used, were recorded using Praat software. Drawing spectrograms of the sound signals, time taken by participants to pronounce the assimilated sound segments was calculated. The time taken by native speakers of English to pronounce the same sound segments was also checked. The mean values of time taken by participants and native speakers to pronounce each assimilated sound segment were compared, with the assumption that if the participants took more time to pronounce the sound segments, they missed assimilation. The findings revealed that although in comparison to native speakers, Saudi EFL learners clocked slightly higher time duration, for most sounds the difference in time was not significant from the statistical point of view. The conclusion is that Saudi English learners are making efforts to be as close to native speakers as possible in using assimilated nasal sounds in English.

KEYWORDS: assimilation of nasals; contextual predictability; Saudi EFL; pronunciation errors; phonology

1. Introduction

Assimilation is a very common phenomenon in every human language including English (Cruttenden, 2014; Dawood and Atawneh, 2015; Heselwood and Watson, 2013; Jones, 1976; Ladefoged, 2003; Ladefoged and Johnson, 2011; Laver, 1994). Assimilation is particularly common in the case of nasal sounds (McMahon, 2002). A common observation is that assimilation varies in extent according to speaking rate and style. Assimilation is more likely to be found in rapid, casual speech and less likely in slow, careful speech (Roach, 2012, p. 110). Since fluent English is generally spoken at a faster rate, English accommodates more assimilation of sounds. Another observation is that assimilation takes place either across word boundaries or within a word. Assimilation within a word gives rise to what becomes the standard pronunciation of the given word (Skandera and Burleigh, 2011, p. 89). For instance, pens /penz/, impossible /ɪm'pɒs.ə.bəl/, illegitimate /ɪl.ɪ'dʒɪt.ə.mət/, illogical /ɪ'lɒdʒ.ɪ.kəl/, immortal

/ɪ'mɔː.təl/, records /rɪkóːdz/, and walked /wɔːkt/, are examples of this phenomenon (Lecumberri and Maidment, 2000).

1.1. Research problem

It has been observed that the normal speech of adult EFL learners in Saudi Arabia is devoid of assimilation in both contexts, at the citation phonological form of the lexis level as well as at the level of connected fast speech. We found in our preliminary analysis of learners' speech that the frequency of assimilation of nasal consonant sounds in the speech of Saudi undergraduate students was low.

The normal speech of selected students sounds entirely different, and thus non-standard, in comparison to the speech of native speakers. The selected students were found to be particularly prone to make errors in the assimilation of nasal sounds. The issue was noted to be quite persistent, and, therefore, needed further research. The existing research literature on the topic is insufficient to shed enough light on it.

1.2. Review of existing literature

1.2.1. Assimilation

Assimilation is a phonological process observed across languages where a sound segment affects the features of its adjacent sound—preceding or following segment—in sequence, within a word or between words (Autores, 2008; Celce-Murcia et al., 2010; Crowley and Bowern, 2010; Fromkin et al., 2014; Groll et al., 2021; Hall, 1997; Knight, 2012). Sound segments in words and phrases may affect each other in various ways. The affected sound segment may undergo a change in its place of articulation, manner of articulation, or in voicing. Thus, it may be place assimilation, manner assimilation, or voicing assimilation (Jansen, 2007). Depending on the location of the affected sound, assimilation may be progressive or regressive. If a sound affects the one next to it in sequence, i.e., the second sound takes the features of the first sound, it is called 'progressive assimilation'. For example, the plural marker phoneme /s/ and the past tense marker phoneme /d/ undergo progressive voicing assimilation. If a sound affects the one before it in sequence, i.e., the first sound takes the features of the sound next to it, it is called 'regressive assimilation'. For example, the phoneme /n/ in the prefix {-in} may undergo place assimilation where it takes the articulation features of the sound before it.

Assimilation may be 'partial' if one sound becomes similar to another sound only in one feature (place, manner, or voice). Assimilation is called 'full' if one sound becomes identical to the other sound. If a sound takes the voicing feature of its neighboring sound, it is called voicing assimilation. In addition, contextual assimilation is subject to the environment of sounds where historical assimilation has resulted from the development of the language. As a language develops, one sound may change into another if the two share the manner, place, and voicing features. For example, look at the following excerpt from Chaucer's *Canterbury Tales* (1387–1400):

And in his tyme swich a conquerour;
That gretter was ther noon under the sonne.
Ful many a riche contree hadde he wonne;
What with his wysdom and his chivalrie. (The Knight's Tale)

The vowel sounds in the Middle English words, such as swich, gretter, sonne, contree, wonne, and wysdom (Modern English equivalents—such, greater, sun, country, won, and wisdom, respectively) changed to different vowels in Modern English.

1.2.2. Research on assimilation in general

There has been a continuous stream of research in this prolific area in phonetics and phonology, in native as well as non-native English settings. Researchers have focused attention on the assimilation of place, manner, and voicing in English in general and on the assimilation behaviour of nasals in particular (Colantoni et al., 2023; Koffi and Schwintek, 2023; Singh and Cheng, 2023). For instance, grammarians and linguists have pointed out that in connected speech sounds at word boundaries assimilate very frequently and produce a different sound, such as in phrases like 'good night' the final alveolar plosive /d/ in 'good' assimilates to alveolar nasal /n/ of 'night' (ALIC, 2023). Articulation place assimilation is very common in English; so, very frequently, /d/ and /j/ assimilate to produce /dʒ/, /z/ and /j/ make /ʒ/, /t/ and /j/ make /tʃ/, /s/ and /j/ make /ʃ/, and so on. Researchers also note that word-final alveolar consonants in English, i.e., /t/, /d/, /n/, /s/, and /z/ change their place of articulation to that of the following word-initial consonant. For instance, as noted by Renwick et al. (2013, p. 1), 'ran quickly' is pronounced as 'raŋ quickly'.

Voicing assimilation is also a very common occurrence in English where a voiceless sound close to a voiced sound becomes voiced, or a voiced sound close to a voiceless sound becomes voiceless. For instance, 'walked' is pronounced as 'wakt' since the voiceless phoneme /k/ triggers assimilation of the voiced phoneme /d/. On voicing assimilation, researchers maintain that phonologically voiceless obstruents, for example, /t/ and /s/ as well as phonologically voiced sounds, such as /z/, and to a large extent /d/, trigger certain forms of voicing assimilation in preceding obstruents (Jansen, 2007; Weismer, 1980).

In English, assimilation of place in the case of three nasal phonemes (/m/, /n/ and /ŋ/) is very common, e.g., pancake 'pæŋkek', incredible 'ɪn'kredəbəl' (ALIC, 2023). A general rule for the assimilation of nasals is that in English, the place of articulation of nasals assimilates to that of a following stop. For instance, in rapid speech, handkerchief is pronounced 'hæŋkə-tʃif', and handbag is pronounced 'hæmbæg'. Research studies show that word-final coronal nasal consonants undergo place assimilation toward following word-initial labial consonants (Hon, 2005). Ellis and Hardcastle (2002) state that English [nasal] place assimilation is a phonetically gradient process, that is, 'a process that involves fine phonetic differences, but involves no change in phonological categories' (Strycharczuk, 2019). Alveolar nasal /n/ assimilates more frequently than the velar and bilabial nasals /ŋ/ and /m/, respectively (Renwick et al., 2013).

1.2.3. Research on assimilation in Saudi EFL contexts

Research on pronunciation errors in English speech of non-native speakers includes loss of assimilation in their speech as one aspect of the comprehensive analysis. EFL speakers of Arabic origin, for instance, commonly pronounce the English words close to their written form, that is, in accordance with their spellings, as they may assume that written words perfectly symbolize the pronunciation and; hence, whatever is written is to be pronounced as it is. This phenomenon makes them miss assimilation. Research works on this phenomenon highlight the loss of assimilation in their English speech as one of the areas of their focus (e.g., Al-Jarf, 2022). Al-Jarf (2022) studied the inaccuracies in proper noun pronunciation in English by educated Arabic speakers and finds that the informants had problems in articulating English vowels, replaced consonants absent in L1 with their equivalents, geminated consonants, inserted vowels in consonant clusters, broke certain words into sub-words, and pronounced words as they are written. Previous research shows that it is generally hard for Arab speakers of English to differentiate between certain English sounds, such as the pairs of voiced and voiceless plosives (p/b,

t/d, k/g) and between nasal consonants (m/n, n/ŋ) in certain words (Abdelaal, 2017), although some speakers effectively manage to produce, and differentiate between, the intended sound segments. Saudi ESL learners, in general, have greater difficulty in categorizing /m/, /n/, and /ŋ/ as separate phonemes, dealing with these sounds as a single category. They have especially more difficulty in distinguishing /n/ from /m/ and /m/ from/ŋ/ (Alharbi, 2014).

It appears that Arab speakers of English fail to observe the contextual predictability rule for assimilation of nasals in English, as suggested by Turnbull et al. (2018).

Researchers, such as Ali (2012), have attempted to pinpoint the sources of this error through contrastive studies, such as finding the differences between English and Arabic sound systems, and state that mother-tongue influence plays a big role in the phenomenon. To EFL learners, place assimilation, especially pertaining to certain consonant sounds, particularly consonant clusters, appears to be the hardest to grasp and acquire. Researchers show that the difficulty arises owing to learners' lack of understanding of the articulatory and perceptual mechanisms involved in place assimilation as well as because of some crosslinguistic factors, such as mother-tongue influence (Jun, 1995).

EFL/ESL speakers' Voice Onset Time (VOT) analysis helps researchers to throw some light on issues, such as loss of assimilation in their English. Research studies on VOT analysis exist in a number of languages, though there is a scarcity of research on the acquisition of VOT of English sounds. In Saudi EFL contexts, this scarcity of research is especially a limiting factor. There are only a few studies to fill these gaps (e.g., Alanazi, 2018; Flege, 1980; Mitleb, 2009). Results from previous studies show that Arab learners of English choose not to apply consonantal assimilation, neither at a slow speech rate, as predicted by some researchers, nor at a fast speech rate. Moreover, it has been reported that neither leaner fluency nor speech rate influences the assimilation processes (Zuraiq and Abu-Abbas, 2009).

Research on bilingual subjects shows that they commonly develop socio-linguistically appropriate production patterns for each of their languages (Khattab, 2002). This set speech production pattern also affects learners' acquisition of assimilation in the target language, affecting their fluency and naturalness.

1.3. Research gap

A review of existing research literature on EFL/ESL learner issues in assimilation of nasals in English shows that there is a scarcity of research in this area, particularly pertaining to assimilation issues concerning Saudi adult learners. Previous studies on learners' problems in nasals' assimilation are either limited in their approach, as is the study by Al-Jarf (2022), or their focus is on learners categorizing the nasals as separate phonemes (Abdelaal, 2017; Alanazi, 2018; Alharbi, 2014), or they do not address issues pertaining solely to Saudi learners. Thus, there exists a research gap that the present study is an attempt to fill.

1.4. Research hypothesis

Assimilation of sound segments within words, and between words and phrases in fast speech, is a common phenomenon in the speech of native speakers of English. However, the practice is observed to be less prominent a feature in the speech of non-native speakers of English, marking it recognizably distinct from native speech. Identification of the issue as a research problem and a preliminary review of existing research literature on the topic led the researchers to formulate the hypothesis that adult Saudi EFL learners do not observe the contextual predictability rule for assimilation of nasals in English. Additionally, it was hypothesized that a time analysis of nasal sounds in words and at word boundaries spoken by adult Saudi learners may reveal some factors affecting their pronunciation.

1.5. Research questions

Keeping the formulated hypothesis in mind, the following research questions were framed to carry out empirical research on the issue:

RQ 1: Do Arab speakers of English fail to observe the contextual predictability rule for assimilation of nasals in English?

RQ 2: What specific factors affect the acquisition of assimilation of nasals for Arab speakers of English?

1.6. Research objectives

The primary objective of the present study has been to investigate whether adult Saudi EFL learners follow assimilation rules in their conversational English, specifically pertaining to nasal sounds. Also, if adult Saudi EFL learners are observed not to follow rules of assimilation of nasal sounds in English, the secondary objective of the study would be to identify the potential factors that affect the speech process of Saudi learners to bypass the phonological phenomenon in their English speech. In specific terms, the researchers designed this study to investigate the loss of assimilation of nasals in the normal as well as fast English speech of Saudi EFL undergraduate learners, and the factors affecting it, if any.

1.7. Scope of the present study

The present study has been limited in its scope as the researchers specifically investigated the loss of assimilation concerning the nasal sounds in English in the normal as well as fast English speech of Saudi EFL undergraduate learners. For the present study, only nasals in the citation phonological form (/m, n, ŋ/) are taken as the target of assimilation, while labial or velar plosives (/p, b, t, d, k, g/) are the triggers of assimilation.

2. Methodology

Mixed-Methods methodology was employed to conduct the present research. Thus, a blend of quantitative and qualitative methods was used for data collection and analysis. The quantitative method involved data collection by recording the words and phrases spoken by research participants. Participants were asked to read out the given words/phrases aloud, and their responses were recorded using the Praat software (Boersma and Weenink, 2012). Further, Praat was used to draw spectrograms of the recorded sounds and to measure the time the participants took in articulating the assimilated sounds.

Buckeye Corpus of spontaneous speech (Pitt et al., 2007) was used as a reference manual for comparison of sounds uttered by Saudi speakers with the sounds uttered by native speakers of English recorded in the Corpus to investigate the variation, if any, in the two utterances. The Buckeye Corpus includes audio recordings of one-on-one interviews with forty native English speakers from Columbus, Ohio, and it is comprised of roughly 300,000 words. For comparative purposes, sixteen different utterances in which assimilated nasal sounds were used, were selected from the Corpus and their time durations were measured using Praat.

The qualitative analysis involved interpretation of the results obtained through data analysis. The numerical data obtained through statistical analysis was interpreted qualitatively to be presented in narrative format for ease of understanding.

2.1. Research design

In order to achieve the research objectives and find answers to the research questions, the following design was set to conduct empirical research:

To find an answer to the first research question, time taken by adult Saudi EFL learners to pronounce assimilated nasal sounds, and a few other sounds, used in words and phrases was calculated and compared with the time native speakers of English take in articulating the same sounds in similar environments in their recorded utterances. The logic behind the idea to find exact time duration has been that if Saudi learners take more time in articulating the assimilated sounds in comparison to native speakers of English, Saudi learners are not following the rules of assimilation of the sounds in question.

To find an answer to the second research question, certain assumptions will be tested against the data obtained by previous researchers and the data generated through the present research.

2.2. Theoretical framework

The present research was carried out employing the contextual predictability rule for the assimilation of nasals in English as suggested by Turnbull et al. (2018): '*More assimilation should be observed for more contextually predictable target words, while "less" assimilation should be observed for contextually more predictable "trigger" words*.' Contextual predictability of words lies in the talker's assumption that phonetic reduction in words would not hamper communication. For example, in the phrase 'hang glider', 'hang' is the target word for which 'glider' is the trigger word. Triggered by the initial sound /g/, in 'glider' the final sounds /ng/ in 'hang' are phonetically reduced to be /ŋ/ as the speaker can safely assume that the reduction wouldn't hamper communication. We assume in the present research that adult Saudi EFL learners do not follow this phonetic rule because they assume that phonetic reduction would hamper communication.

Perceptual Assimilation Model, PAM and PAM-L2 (Best and Tyler, 2007) was employed to make sense of the factors affecting the pronunciation of Saudi EFL learners. In accordance with the model, adults perceive the speech sounds of a non-native language based on similarities and differences with their native language. For instance, Saudi speakers of English may not follow assimilation of nasal sounds in English since in Arabic language assimilation, whether of nasal sounds or other sounds, takes an altogether different form, and in general, ends up in overlapping of sounds (Saleem, 2022).

2.3. Participants

A total of 16 students took part in the present study as participants. Random selection process was followed to select participants for the study since almost all the students in the class out of which participants were selected stood at the same level as regarded assimilation of sounds in their English speech. All the participants were undergraduate students enrolled in the second semester of the Preparatory-Year program at the university. The participants had undergone six years of training in English as a foreign language and were able to speak the language with moderate fluency. To avoid psychological bias in data collection, the purpose of the speaking test was not revealed to the participants, as it would have made them over-conscious.

2.4. Data collection and measurement

2.4.1. Data collection

The collected data consisted of sound files recorded through Praat. 16 sound files were created as each participant was given the speaking opportunity individually. Similarly, utterances of 16 different interviewees were culled from the Buckeye Corpus and analyzed using Praat.

2.4.2. Procedure

Utterances of participants were recorded in the quiet language lab. Panasonic noise-cancelling headsets with built-in microphone connected to a laptop were used for the recordings. The recordings were directly saved on the Praat software at 5000 Hz sampling rate. The reason behind this choice of frequency range is that the human auditory system is 'most sensitive to frequencies between 2000 Hz and 5000 Hz' (Gelfand, 2011, p. 87). The sound sections consisting of assimilation were zoomed out, played repeatedly for exactitude, and their visible parts were checked for the exact time duration of the articulated sounds. The sound-play and time-check process was repeated with the sound segments collected from the Buckeye Corpus.

2.4.3. Measurement: Instrument

The Praat software program is a very convenient tool for the measurement and analysis of sound segments. The software provides on a single platform the facility for recording, drawing spectrogram, measuring the time duration, and presenting the formant pictures. Mean values, standard deviation and variance in time duration for each assimilated sound segment uttered by participants and interviewees were calculated using an online platform (EasyCalculation, 2023). The mean values of time for the selected assimilated sound segments obtained from the utterances of the Arab participants and English interviewees were compared. The significance of difference, if any, in the mean values of time duration for each sound segment was calculated using *t*-score analysis.

2.4.4. Trigger data

The selected trigger words/phrases largely used assimilated nasal sounds, though to make the data collection and analysis more reliable, a few single words and phrases containing other assimilated sounds were also used as triggers. Assimilated English nasal sound segments (m, n, ŋ) were specifically selected for the analysis since assimilation is particularly common in the case of nasal sounds (McMahon, 2002). Following is a list of trigger words/phrases and the sentences they were used in.

1) Incorrect

Sentence: *Circle the incorrect answers.*

2) Incredible

S: *The athlete has achieved an incredible feat*.

3) Ten million

S: *The community event was attended by ten million people*.

4) Handkerchief

S: *Do you always carry a handkerchief?*

5) Handbag

S: *What's there in your handbag?*

6) Income

S: *His income is always less than his expenditures.*

7) Pancake

S: *Do you use eggs to make a pancake?*

8) Humpback

S: *A humpback whale jumped out of the sea*.

9) Green boat

S: *While they were swimming, a green boat passed by*.

10) Hang gliding

S: *Hang gliding is an interesting sport*.

11) It's

S: *It's no more there*.

12) That's

S: *That's what I wanted to say*.

13) Don't you.

S: *Don't you take a bath every day?*

14) Did you

S: *Did you say yes or no*?

15) Good night

S: *Good night, dear. Sweet dreams!*

16) Encouraging

S: *The progress in this project is encouraging*.

17) Don't beat

S: *Don't beat the child, please.*

Table 1 below presents a summary description of the trigger data used in the research.

Table 1. Summary description of trigger data.

	Word/phrase	IPA transcription	Assimilated sound	Assimilation type
1	Incorrect	/ínkərékt/	/n/ > /ŋ/	Regressive, place
2	Incredible	/ɪnkrédəbəl/	/n/ > /ŋ/	Regressive, place
3	Ten million	[tɛn mɪljən]	/n/ > /ŋ/	Regressive, place
4	Handkerchief	/háŋkəʧɪf/	/n/ > /ŋ/	Regressive, place
5	Handbag	[ˈhæn(d)ˌbæg]	/n/ > /m/	Regressive, place
6	Income	/ˈɪnkʌm/	/n/ > /ŋ/	Regressive, place
7	Pancake	[pæŋkek]	/n/ > /ŋ/	Regressive, place
8	Humpback	/ˈhʌmpbæk/	/p/ > /b/	Regressive, place
9	Green boat	[griːn'bəʊt]	/n/ > /m/	Regressive, place
10	Hang gliding	[háŋ glɑjdɪŋ]	/g/ > /ŋ/	Regressive, place
11	It's	[ɪts]	/z/ > /s/	Progressive, voicing
12	That's	[ðæts]	/z/ > /s/	Progressive, voicing
13	Don't you	[dontʃ ju]	[tʃju] > [tʃu]	Regressive, place, manner
14	Did you	[dɪdʒ ju]	[dj] > [dʒ]	Regressive, place, manner
15	Good night	[gʊn naɪt]	/d/ > /n/	Regressive, place, manner
16	Encouraging	/ɪnkárɪdʒɪŋ/	/n/ > /ŋ/	Regressive, place
17	Don't beat	[dəʊn 'biːt]	/n/ > /m/	Regressive, place

3. Analysis and discussion

3.1. Data analysis

The collected sound files were processed using Praat to obtain the precise time duration of the sound signals and to draw spectrograms for each assimilated sound segment. **Figure 1**, given below, represents a sample spectrogram displaying the precise time duration for the trigger word "incorrect", pronounced by one of the participants. The time clocked here is 0.170975 s, i.e., roughly 17 cs (See Appendix for spectrograms of all the trigger sound segments).

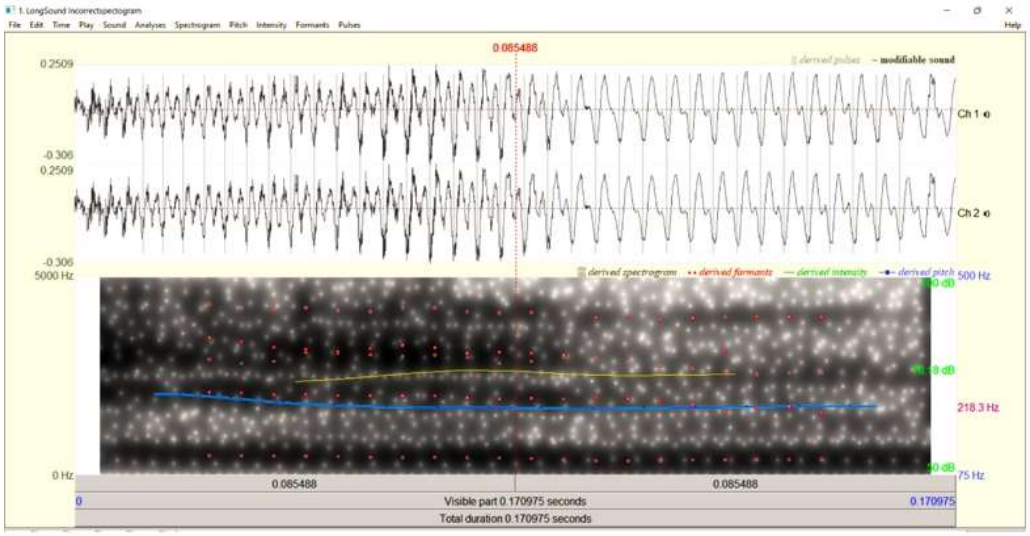

Figure 1. Spectrogram and formant picture for the trigger word 'incorrect'.

Thus, we obtained the duration of time of articulation of the selected assimilated sound segments (i.e., just one sound in question, e.g., /m/, /n/ or /ŋ/) for each participant and the selected few native speakers of English. The obtained raw scores are presented in **Tables 2** and **3**, respectively.

Table 2. Time taken by participants to pronounce assimilated sounds in words and phrases.

1	Assimilated sound	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	Mean	SD	Variance
2	/n/ > /m/	28	22	23	24	25	26	27	22	22	24	21	22	23	24	25	26	24	2.033	3.87
3	/n/ > /ŋ/	14	14	21	20	17	16	18	20	20	17	14	16	16	17	18	18	17	2.236	4.68
4	/p/ > /b/	23	22	23	22	22	22	23	20	22	22	22	22	23	22	23	20	22	0.0928	0.808
5	/d/ > /n/	12	14	11	10	12	10	12	12	10	9	8	10	10	12	9	11	10	1.527	2.18
6	/g/ > /ŋ/	18	18	19	20	22	19	17	14	21	13	22	22	19	19	23	19	19	2.768	7.18
7	/z/ > /s/	13	14	12	13	14	13	16	12	13	14	13	12	14	14	13	12	13	1.064	1.06
8	[tʃju] > [tʃu]	28	27	29	26	25	28	26	27	27	29	30	30	30	25	24	25	27	1.983	3.68
9	[dj] > /dʒ/	22	21	20	20	22	21	22	19	19	18	22	23	24	18	18	17	20	2.061	3.98

N = 16, time unit: centisecond.

Table 3. Time taken by native speakers to pronounce assimilated sounds in words and phrases.

1	Assimilated sound	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	Mean	SD	Variance
2	/n/ > /m/	10	10	12	12	12	12	10	10	12	12	10	10	12	12	10	10	11	1.03	1
3	/n/ > /ŋ/	12	12	10	10	12	12	10	10	10	12	12	12	12	10	10	10	11	1.03	1
4	/p/ > /b/	14	12	13	12	12	14	14	14	13	12	12	14	13	14	12	13	13	0.89	0.75
5	/d/ > /n/	10	12	12	10	10	8	12	10	8	12	10	10	8	12	12	8	10	1.61	2.43
6	/g/ > /ŋ/	13	13	12	12	12	14	12	12	12	12	10	14	10	12	10	12	12	1.21	1.37
7	/z/ > /s/	4	8	4	6	4	6	8	4	6	8	4	6	8	8	6	6	6	1.63	2.5
8	[tʃju] > [tʃu]	12	12	14	10	12	12	13	10	10	12	13	12	12	12	12	14	12	1.21	1.37
9	[dj] > /dʒ/	10	12	12	10	10	8	12	10	8	12	10	10	8	12	12	8	10	1.61	2.43

N = 16, time unit: centisecond.

3.2. Results

Our primary focus in data analysis had been to find out if the research participants clocked more time in articulation of each sound segment compared to the native speakers. So, we calculated the mean time-duration for each sound segment pronounced by the sixteen participants as well as by native speakers. Further, to investigate if the difference in the mean scores bears any statistical significance, we obtained the variance in the means and calculated their *t* scores. The results obtained thus are summarized in **Table 4**.

Table 4. Summary of obtained results.

		Participants			Native speakers				
	Assimilated sounds	**Mean**	**SD**	**Variance**	**Mean**	**SD**	**Variance**	***t* score**	**Results**
1	/n/ > /m/	24	2.033	3.87	11	1.03	1	4.061	Significant at $p < 0.05$
2	/n/ > /ŋ/	17	2.236	4.68	11	1.03	1	–1.461	Not significant at $p < 0.05$
3	/p/ > /b/	22	0.0928	0.808	13	0.89	0.75	–19.66	Not significant at $p < 0.05$
4	/d/ > /n/	10	1.527	2.18	10	1.61	2.43	–8.629	Not significant at $p < 0.05$
5	/g/ > /ŋ/	19	2.768	7.18	12	1.21	1.37	2.58	Significant at $p < 0.05$
6	/z/ > /s/	13	1.064	1.06	6	1.63	2.5	0.280	Not significant at $p < 0.05$
7	[tʃju] > [tʃu]	27	1.983	3.68	12	1.21	1.37	5.64	Significant at $p < 0.05$
8	[dj] > /dʒ/	20	2.061	3.98	10	1.61	2.43	4.20	Significant at $p < 0.05$

For clarity and ease of understanding, the obtained results are represented graphically in **Figure 2**.

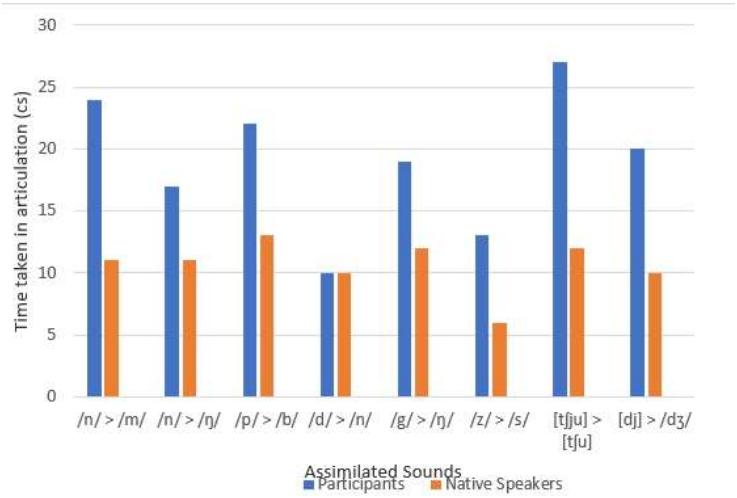

Figure 2. Graphical representation of time taken by participants and native speakers of English to articulate the assimilated sounds.

3.3. Discussion

A quick glance at **Table 4** reveals that in comparison to native speakers of English, the research participants have clocked higher time durations to pronounce each sound segment, except one, i.e., /d/ > /n/, which exactly shows the same time duration in both cases. For example, the average time taken by the research participants to pronounce the assimilation of /n/ > /m/ (/n/ assimilated to /m/) is 24 cs, whereas the average time duration for the same sound pronounced by a native speaker is 11 cs. Our interpretation of Saudi EFL learners taking longer duration of time to pronounce assimilated sounds is that though they are trying to be as close in assimilation of sounds as native speakers, certain factors may be hindering them to be exactly close to native speech. However, there are two significant points worth considering at this juncture. First, the difference in the duration of time, that is, the Saudi learners taking more time, is minimal, almost negligible since the difference in many cases is not even perceptible in

speech. Second, the difference in the case of four out of eight sound segments (/n/ > /ŋ/, /p/ > /b/, /d/ > /n/, and /z/ > /s/) is statistically NOT significant, which means Saudi learners use as accurate assimilation in those sound segments as possible. The difference is significant only in the case of four out of eight sounds (/n/ > /m/, /g/ > /ŋ/, [tʃju] > [tʃu], and [dj] > /dʒ/).

Although the findings from the present study show an encouraging trend among Saudi EFL learners towards categorization and articulation of English sounds, it is probable that Saudi learners lean more towards categorical sounds since they are stable sounds and represent clear distinct phonological categories (Ernestus, 2011), and do not realize gradient sounds as they may change during their realization and may simultaneously represent different phonological categories (Ernestus, 2011) as happens in the assimilation of sounds. What we assume is that the adult Saudi EFL learners who participated in the present study were not accustomed to the articulation of sounds in native English speech that, in the process of assimilation, are begun with one feature (such as voicing) but end up with other feature (e.g., devoicing), or the total loss of a sound, such as the loss of /p/ in humpback /hʌmpbæk/. Therefore, they end up avoiding assimilation altogether. The phenomenon may result from mother tongue influence as well, and that takes us to our second research question. The native tongue of the participants, i.e., Arabic language, differs considerably from English regarding certain phonological features. Saleem (2022) indicates that assimilation in the Arabic language is categorically different from assimilation in English.

A comparative look would have contextualized our findings in a better way; however, given the general lack of previous studies on the subject in the Saudi Arabian academic context, we can't have a comparative perspective on the obtained results. Nevertheless, it is obvious that our results show a progressive trend as regards pronunciation of English sounds by Saudi EFL learners. For example, results from the present study present a better picture of Saudi EFL learners than the results obtained by Al-Jarf (2022) in her study on the inaccuracies in proper noun pronunciation in English by educated Arab speakers of English. Similarly, our results show an improvement over Abdelaal's (2017) findings as well who observed that it is generally hard for Arab speakers of English to differentiate between certain English sounds, such as nasal consonants (m/n, n/ŋ) in certain words. We realize that our participants did find it difficult to differentiate /n/ from /ŋ/, and that's why they had assimilation issues where the distinction played a significant role, but they had no difficulty in differentiating /m/ from /n/. Findings from the present study also indicate that Saudi ESL learners are making improvements in categorizing /m/, /n/, and /ŋ/ as separate phonemes, not dealing with these sounds as a single category as observed by Alharbi (2014). They have no specific difficulty in distinguishing /n/ from /m/ or /m/ from /ŋ/. Our findings also present a contrast with the findings of Zuraiq and Abu-Abbas (2009) who report that Arab learners of English choose not to apply consonantal assimilation, neither at a slow speech rate, nor at a fast speech rate.

4. Conclusion

Findings from the present study can be used to safely conclude that adult Saudi EFL learners, particularly the students in the classes from which the participants for the present study were selected, are making fast progress in articulating, and distinguishing between, typical English consonant sounds properly, which are reported in previous studies to pose difficulties for them (such as, the pairs of voiced and voiceless plosives p/b, t/d, k/g, and the nasal consonants m/n, n/ŋ in certain words). Although Saudi learners take slightly longer time to pronounce the assimilated sound segments compared to native speakers, they seem to observe the contextual predictability rule for assimilation of nasals in English. The

slight difference in the time duration may result from certain specific factors, such as mother tongue influence, socio-cultural environment, lack of exposure to native speech, lack of good English teachers, lack of awareness of assimilation of sounds in certain words/phrases, and so on. The phenomenon may also be affected by two factors—articulatory overlap and rate of speaking. Articulatory overlap is common for native English speakers, while a slow rate of speaking is more common for Arab speakers.

4.1. Significance of the present research

The present study is significant in two ways. First, the study throws some light on a neglected issue, and therefore, may draw the attention of educationists and practitioners towards teaching assimilation of sounds in English to adult Saudi EFL learners. Second, this study is an attempt to fill the gap in research on issues concerning adult Saudi EFL learners' difficulties in learning assimilation of certain English sounds, and thus, may prove to be a valuable part of the existing literature on the topic.

4.2. Further recommendations

Nasal sounds in English are commonly affected by regressive assimilation. English nasal assimilation is also time constrained. Arab speakers, in general, fail to grasp this phonological phenomenon, which needs to be studied extensively. Therefore, there is an urgent need for further research in this area.

For constraints of time and resources, the researchers could not include certain aspects of the use of assimilation in the speech of adult Saudi EFL learners. For instance, further researchers can investigate whether there are any differences between the performances of male and female learners or between the performances of English medium school students and Arabic medium students as regards assimilation of nasal sounds.

Author contributions

Conceptualization, PS and MK; methodology, PS; Software, PS; validation, PS, MK and MS; formal analysis, MS; investigation, PS; resources, MS; data curation, MS; writing—original draft preparation, PS; writing—review & editing, MS. All authors have read and agreed to the published version of the manuscript.

Conflict of interest

The authors declare no conflict of interest.

References

Abdelaal NM (2017). Instrumental analysis of the English stops produced by Arabic speakers of English. *International Journal of Education and Literary Studies* 5(3): 8–15. doi: 10.7575/aiac.ijels.v.5n.3p.8

Alanazi S (2018). *The Acquisition of English Stops by Saudi L2 Learners* [PhD thesis]. University of Essex.

Alharbi TNL (2014). *The Perception of English Syllable-final Nasals by Saudi ESL Learners* [Master's thesis]. Southern Illinois University Carbondale.

Ali ZA (2012). A phonological study of English and Arabic assimilation: A contrastive study. *Journal of the College of Languages (JCL)* (25): 156–211.

Al-Jarf R (2022). Proper noun pronunciation inaccuracies in English by educated Arabic speakers. *British Journal of Applied Linguistics* 2(1): 14–21. doi: 10.32996/bjal.2022.2.1.3

Analyzing Language in Context (ALIC, 2023). Available online: https://alic.sites.unlv.edu/chapter-11-8-assimilation (accessed on 1 September 2023).

Autores V (2008). *Concise Oxford English Dictionary*, 11th ed. Oxford University Press.

Best CT, Tyler MD (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In: Munro MJ, Bohn O (editors). *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*. John Benjamins Publishing.

Boersma P, Weenink D (2012). Praat: Doing phonetics by computer (Version praat6309_win64 software). Available online: http://www.fon.hum.uva.nl/praat/ (accessed on 1 September 2023).

Celce-Murcia M, Brinton DM, Goodwin JM (2010). *Teaching Pronunciation: A Course Book and Reference Guide*, 2nd ed. Cambridge University Press.

Chaucer G (1387–1400). The canterbury tales. Available online: https://chaucer.fas.harvard.edu/pages/text-and-translations (accessed on 1 September 2023).

Colantoni L, Kochetov A, Steele J (2023). L1 influence on the L2 acquisition of English word-final nasal place contrasts: An electropalatographic study of L1 Japanese and Spanish learners. *Laboratory Phonology* 14(1). doi: 10.16995/labphon.6434

Crowley T, Bowern C (2010). *An Introduction to Historical Linguistics*, 4th ed. Oxford University Press.

Cruttenden A (2014). *Gimson's Pronunciation of English*, 8th ed. Routledge.

Dawood HAS, Atawneh A (2015). Assimilation of consonants in English and assimilation of the definite article in Arabic. *American Research Journal of English and Literature* 1(4): 9–15. doi: 10.21694/2378-9026.15011

EasyCalculation (2023). Available online: https://www.easycalculation.com (accessed on 1 September 2023).

Ellis L, Hardcastle WJ (2002). Categorical and gradient properties of assimilation in alveolar to velar sequences: Evidence from EPG and EMA data. *Journal of Phonetics* 30: 373–396. doi: 10.1006/jpho.2001.0162

Ernestus M (2011). Gradience and categoricality in phonological theory. In: Oostendorp M, Ewen CJ, Hume E, et al. (editors). *The Blackwell Companion to Phonology*. Wiley-Blackwell. pp. 2115–2136. doi: 10.1002/9781444335262.wbctp0089

Flege JE (1980). Phonetic approximation in second language acquisition. *Language Learning* 30(1): 117–134. doi: 10.1111/j.1467-1770.1980.tb00154.x

Fromkin V, Rodman R, Hyams N (2014). *An Introduction to Language*, 10th ed. Wadsworth.

Gelfand S (2011). *Essentials of Audiology*. Thieme.

Groll MD, Hablani S, Stepp CE (2021). The relationship between voice onset time and increase in vocal effort and fundamental frequency. *Journal of Speech, Language, and Hearing Research* 64(4): 1197–1209. doi: 10.1044/2021_JSLHR-20-00505

Hall TA (1997). *The Phonology of Coronals*. John Benjamins Publishing.

Heselwood B, Watson J (2013). The Arabic definite article does not assimilate. *Leeds Working Papers in Linguistics and Phonetics* 18: 34–53.

Hon EA (2005). *An Acoustic Analysis of Labialization of Coronal Nasal Consonants in American English* [Master's thesis]. Massachusetts Institute of Technology.

Jansen W (2007). Phonological 'voicing', phonetic voicing, and assimilation in English. *Language Sciences* 29(2–3): 270–293. doi: 10.1016/j.langsci.2006.12.021

Jones D (1976). *An Outline of English Phonetics*, 9th ed. Cambridge University Press.

Jun J (1995). *Perpetual and Articulatory Factors in Place Assimilation: An Optimality Theoretic Approach* [PhD thesis]. University of California.

Khattab G (2002). *Sociolinguistic Competence and the Bilingual's Adoption of Phonetic Variants: Auditory and Instrumental Data from English-Arabic Bilinguals* [PhD thesis]. University of Leeds.

Knight RA (2012). *Phonetics: A Coursebook*. Cambridge University Press.

Koffi E, Schwintek E (2023). An idiolectal acoustic phonetic analysis of English nasal consonants. *Linguistic Portfolios* 12: 11.

Ladefoged P (2003). *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques*. Wiley-Blackwell.

Ladefoged P, Johnson K (2011). *A Course in Phonetics*, 6th ed. Wadsworth.

Laver J (1994). *Principles of Phonetics*. Cambridge University Press. doi: 10.1017/CBO9781139166621

Lecumberri M, Maidment JA (2000). *English Transcription Course*. Routledge.

McMahon A (2002). *An Introduction to English Phonology*. Edinburgh University Press.

Mitleb FM (2009). Voice onset time of Jordanian Arabic stops. *The Journal of the Acoustical Society of America* 109(5): 2474. doi: 10.1121/1.4744787

Pitt MA, Dilley L, Johnson K, et al. (2007). Buckeye corpus. Available online: www.buckeyecorpus.osu.edu (accessed on 1 September 2023).

Renwick ME, Baghai-Ravary L, Temple R, Coleman JS (2013). Assimilation of word-final nasals to following word-initial place of articulation in United Kingdom English. *Proceedings of Meetings on Acoustics* 19(1): 060257. doi: 10.1121/1.4800279

Roach P (2012). *English Phonetics and Phonology: A Practical Course*, 4th ed. Cambridge University Press.

Saleem MA (2022). Nasal assimilation in standard Arabic with reference to English: A phonological analysis.

Journal of Tikrit University for Humanities 29(1,1): 57–73. doi: 10.25130/jtuh.29.1.1.2022.25
Singh L, Cheng Q (2023). Words in context: Compensation for phonological assimilation in monolingual and bilingual toddlers. *First Language* 43(4): 407–430. doi: 10.1177/01427237231161135
Skandera P, Burleigh P (2011). *A Manual of English Phonetics and Phonology*. Gunter Narr Verlag.
Strycharczuk P (2019). Phonetic detail and phonetic gradience in morphological processes. *Oxford Research Encyclopedia of Linguistics.* doi: 10.1093/acrefore/9780199384655.013.616
Turnbull R, Seyfarth S, Hume E, Jaeger TF (2018). Nasal place assimilation trades off inferrability of both target and trigger words. *Laboratory Phonology* 9(1): 15. doi: 10.5334/labphon.119
Weismer G, Elbert M (1980). The duration of normally articulated and functionally misarticulated [s] in preschool children. *The Journal of the Acoustical Society of America* 67(S1): S39. doi: 10.1121/1.2018198
Zuraiq WMS, Abu-Abbas KHM (2009). Assimilation processes by Arab learners of English1. *International Journal of Arabic-English Studies (IJAES)* 10(1): 85–92. doi: 10.33806/ijaes2000.10.1.7

Appendix

Spectrograms of the selected sound segments

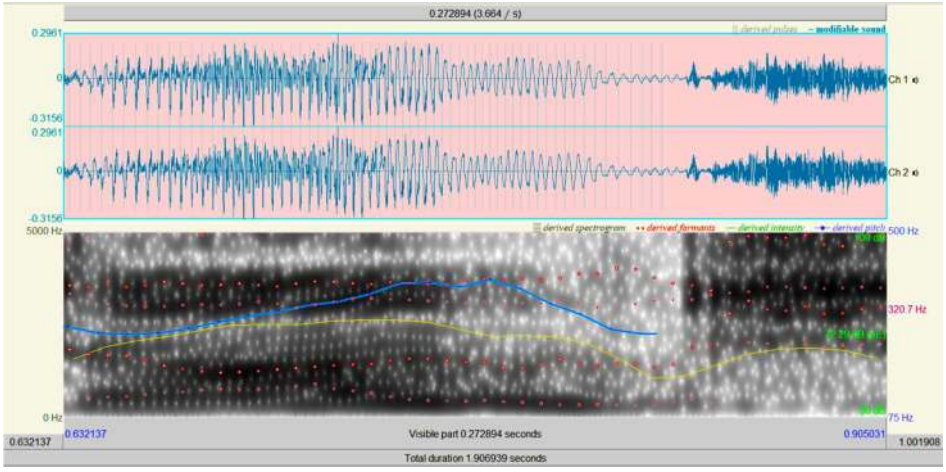

Figure A1. Don't you.

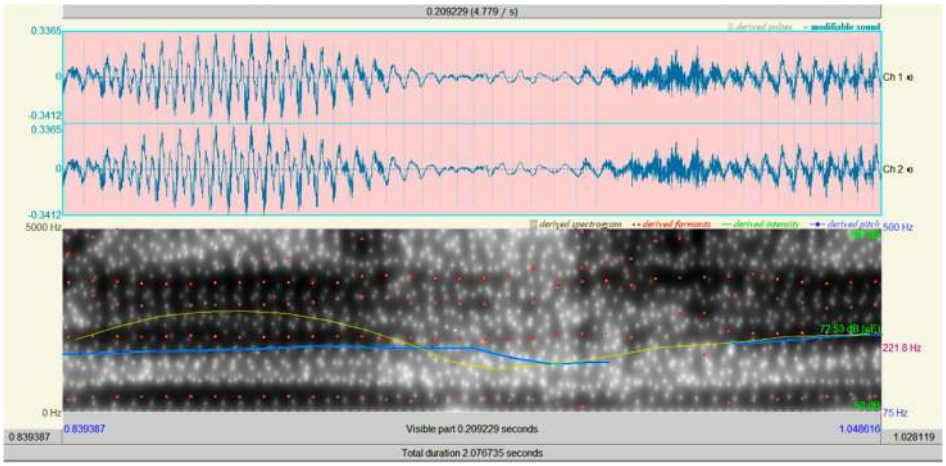

Figure A2. Did you.

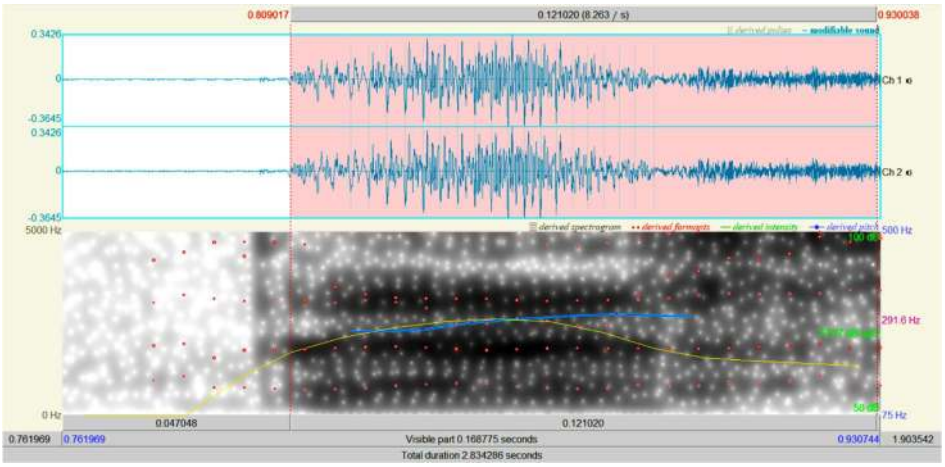

Figure A3. That's.

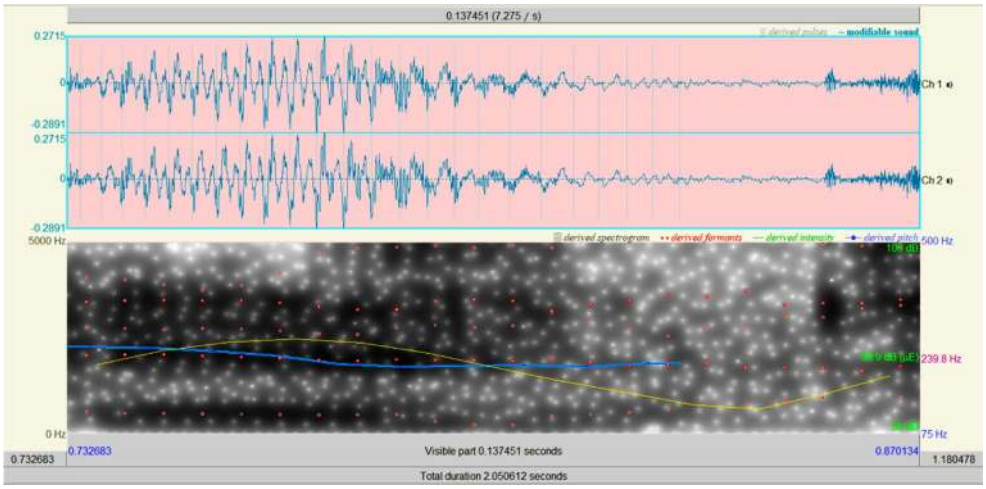

Figure A4. It's.

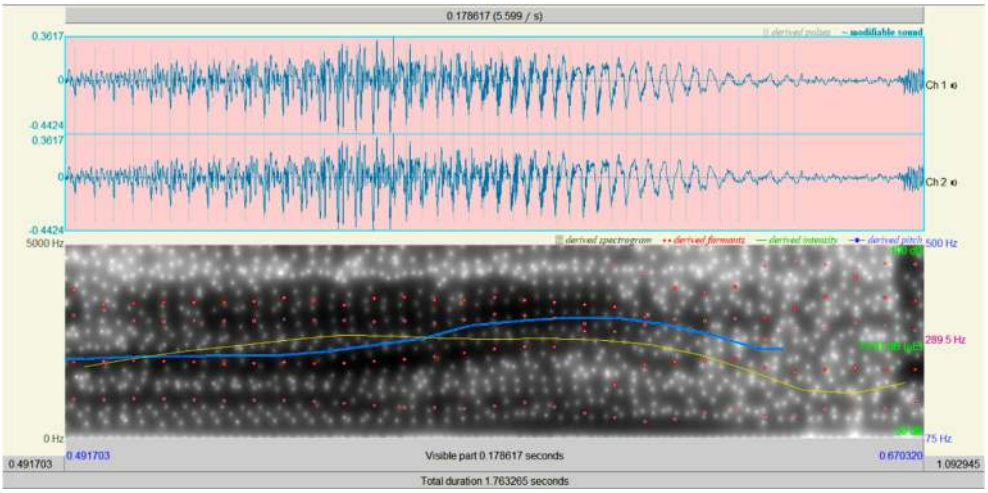

Figure A5. Encouraging.

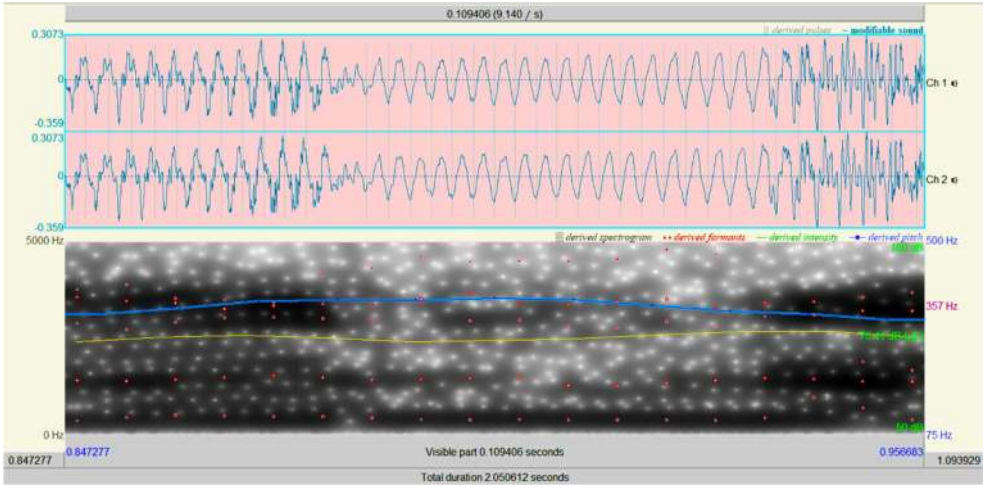

Figure A6. Good night.

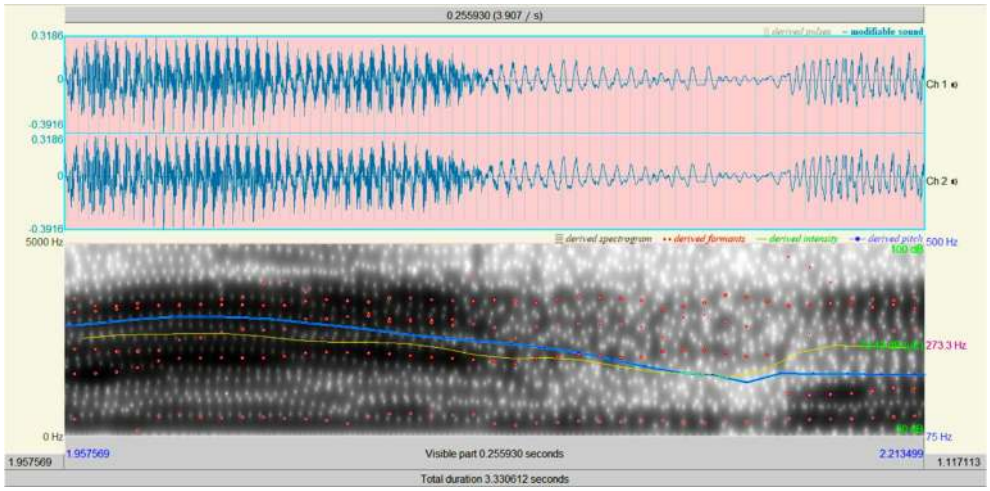

Figure A7. Green boat.

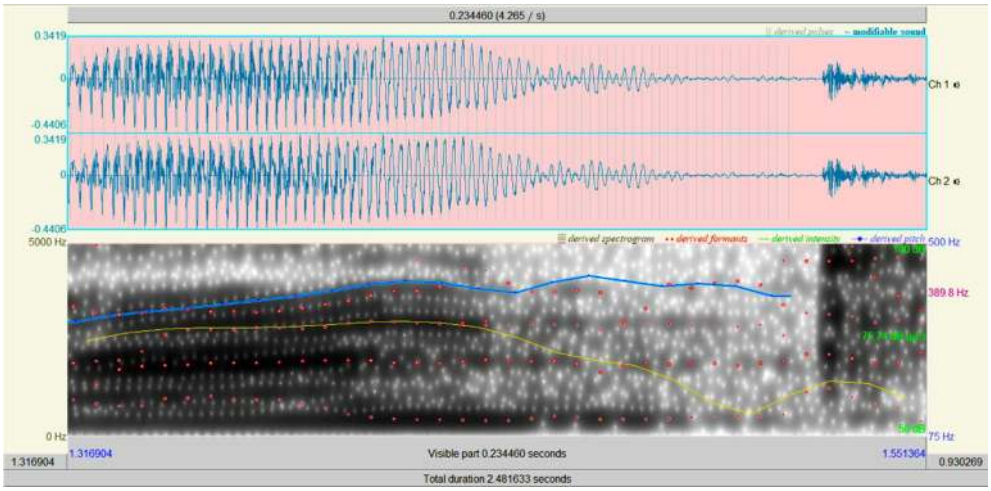

Figure A8. Pancake.

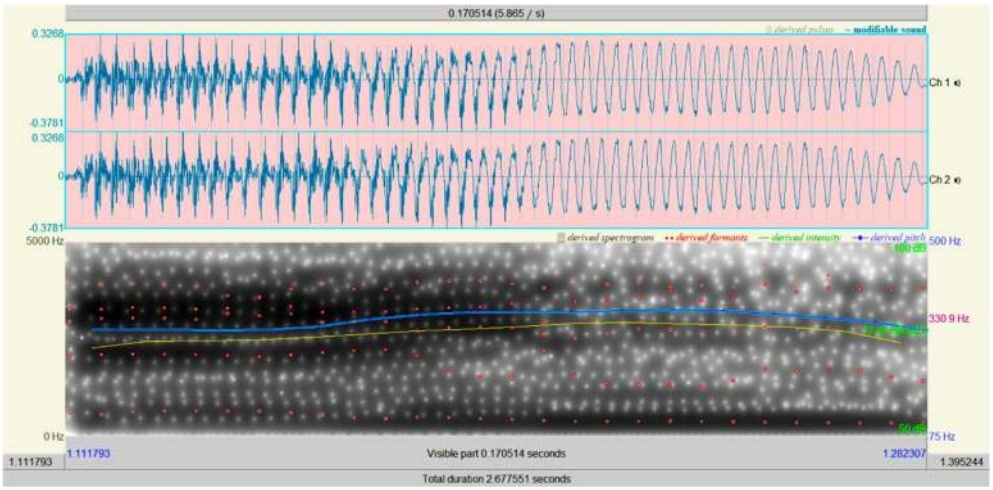

Figure A9. Income.

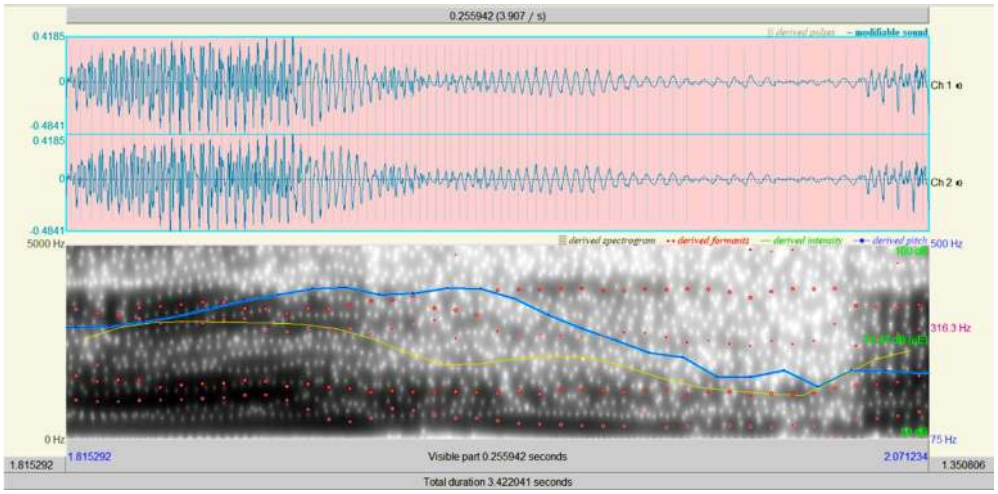

Figure A10. Humpbacked.

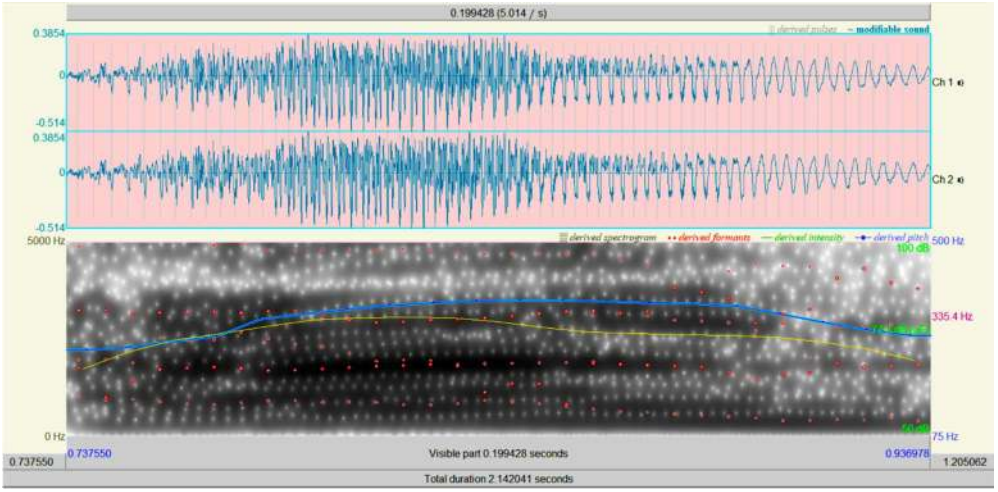

Figure A11. Hang glider.

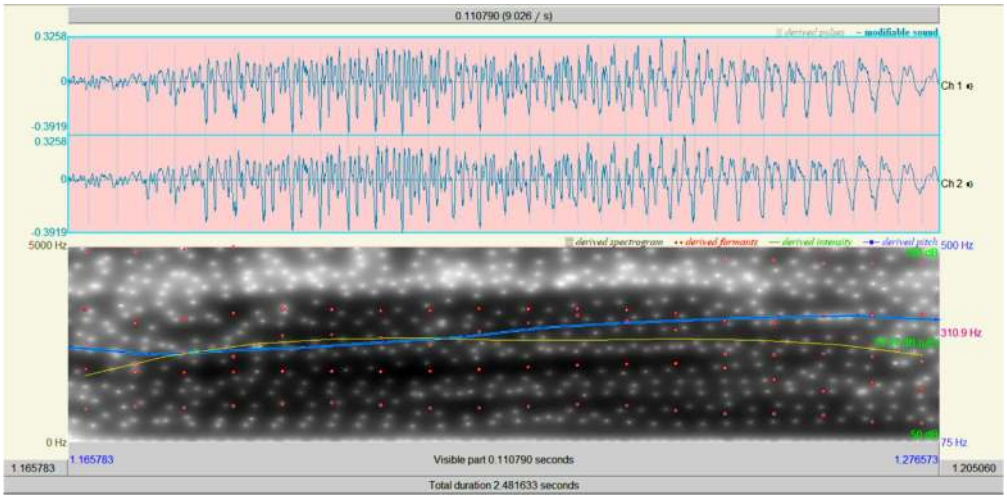

Figure A12. Handkerchief.

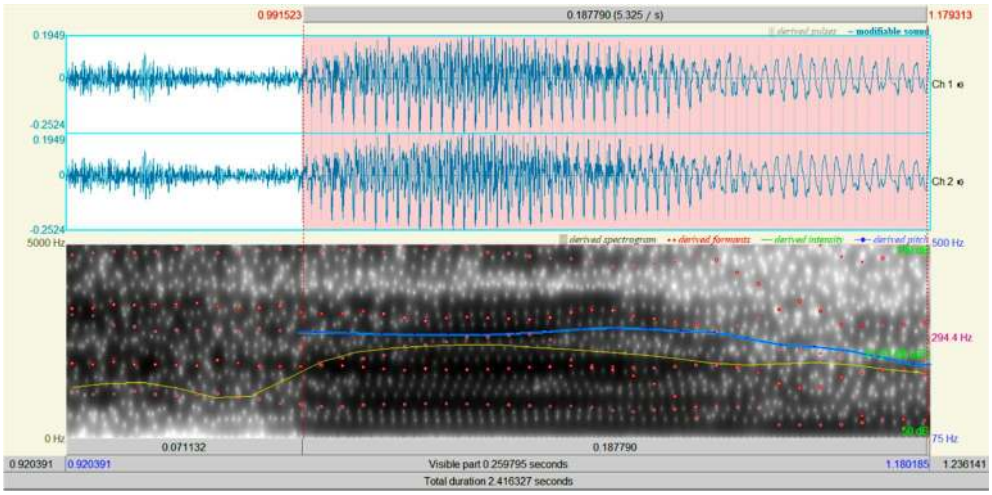

Figure A13. Handbag.

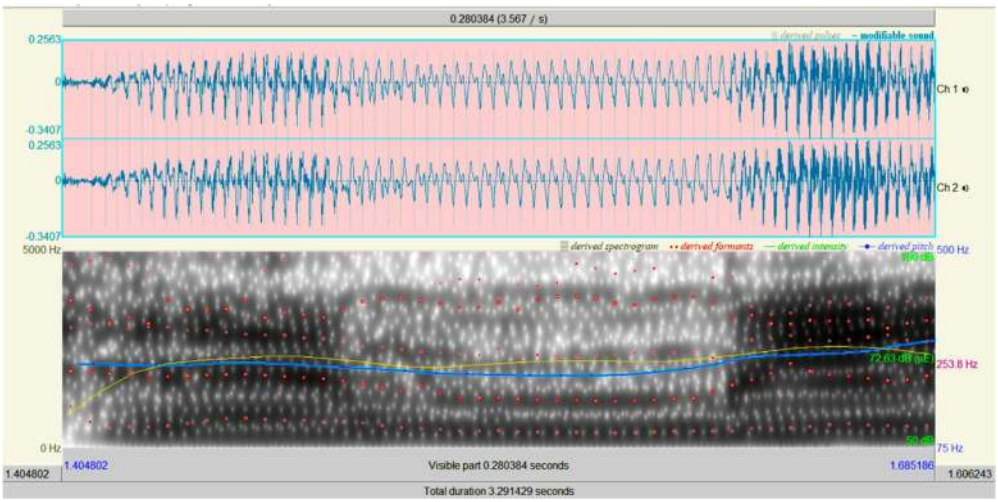

Figure A14. Ten million.

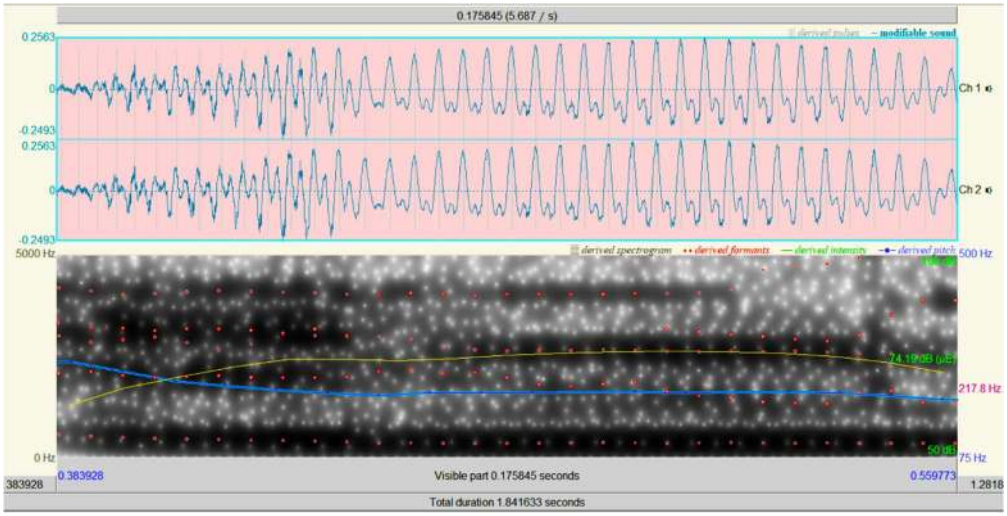

Figure A15. Incredible.

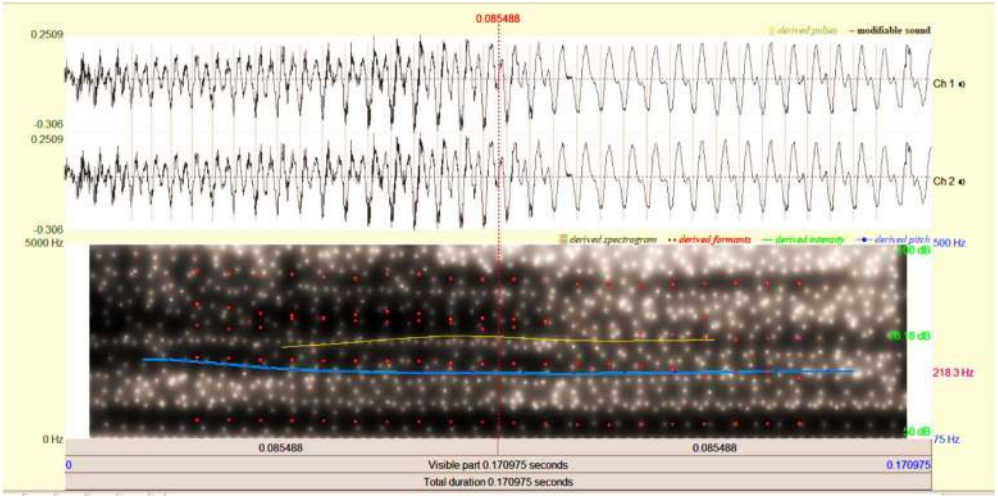

Figure A16. Incorrect.

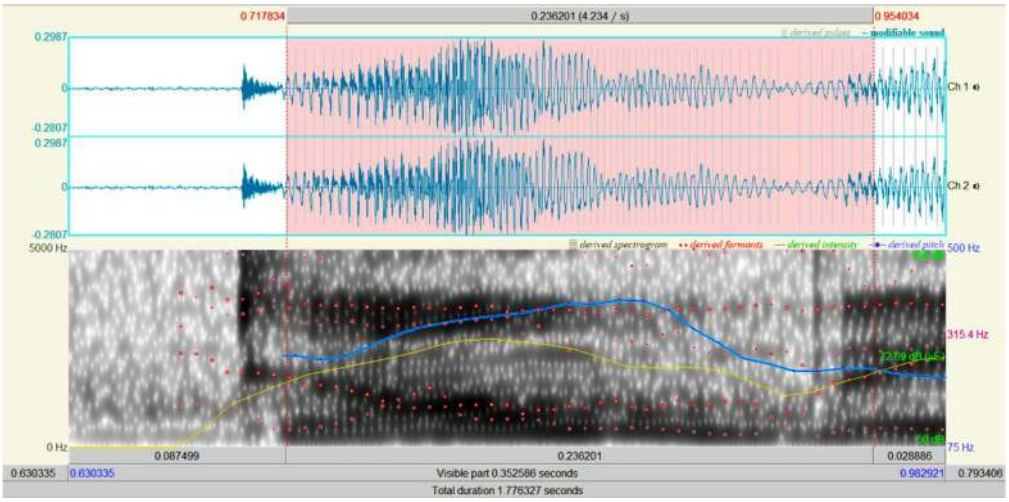

Figure A17. Don't beat.