

AN ACOUSTIC ANALYSIS OF VOICELESS ALVEOLAR PLOSIVE /t/ IN SUNDAHESE, INDONESIA, AND ENGLISH BY SUNDAHESE SPEAKERS

Jeri Sukmawijaya

Linguistics Program, Faculty of Cultural Sciences
Universitas Padjadjaran
jeri17001@mail.unpad.ac.id

Sutiono Mahdi

Linguistics Program, Faculty of Cultural Sciences
Universitas Padjadjaran
sutionomahdi@gmail.com

Susi Yuliawati

Linguistics Program, Faculty of Cultural Sciences
Universitas Padjadjaran
susi.yuliawati@unpad.ac.id

Abstract

A voiceless alveolar plosive /t/ is pronounced with many variations. In English, it is perceived to have an aspiration when it is in the initial segment of a word in a stressed syllable while Sundanese and Indonesian do not aspirate it. The study focuses on finding out Voice Onset Time (VOT) or duration of /t/ in Sundanese, Indonesian, and English produced by Sundanese speakers. The method used is quantitative and qualitative (mix method). The data are obtained from the voice recording of eighteen undergraduate students. They are all native Sundanese speakers who are learning English at IKIP Siliwangi Cimahi. The voice recordings are analyzed and measured by acoustic analysis by using PRAAT software. The result shows the average VOT or duration of /t/ in the initial segment of the word in Sundanese and Indonesian is identical, namely 20 ms. It indicates that they do not aspirate /t/ in Sundanese and Indonesian. Additionally, they transfer the way they produce /t/ in their two languages into English as their foreign language. It can be seen from the average VOT or duration of /t/ in the initial segment of the word in English that lasts 29 ms. This finding can be a reference to other researchers who want to know the necessary duration of /t/ in Sundanese and Indonesian produced by Sundanese speakers.

Keywords: VOT, /t/ in Sundanese, /t/ in Indonesian, Acoustic Analysis, PRAAT

Abstrak

Sebuah bunyi alveolar plosif /t/ diucapkan dengan banyak variasi. Dalam bahasa Inggris, bunyi /t/ dianggap memiliki aspirasi ketika ia berada di suku kata yang ditekan di awal kata, sementara bahasa Sunda dan Indonesia tidak mengaspirasikannya. Penelitian ini berfokus untuk menemukan Voice Onset Time (VOT) atau durasi dari bunyi /t/ dalam bahasa Sunda, Indonesia, dan Inggris yang dihasilkan oleh penutur bahasa Sunda. Metode yang digunakan adalah metode kuantitatif dan kualitatif (metode campuran). Datanya diperoleh dari rekaman suara delapan belas mahasiswa sarjana. Mereka semua adalah penutur asli bahasa Sunda yang belajar bahasa Inggris di IKIP Siliwangi Cimahi. Rekaman suara itu dianalisis dan diukur

dengan analisis akustik menggunakan perangkat lunak PRAAT. Hasilnya menunjukkan VOT atau durasi rata-rata dari bunyi /t/ di segmen awal kata dalam bahasa Sunda dan Indonesia itu sama, yaitu 20 ms. Itu mengindikasikan bahwa mereka tidak mengaspirasikan bunyi /t/ dalam bahasa Sunda dan Indonesia. Selain itu, mereka juga mentransfer cara mereka melafalkan bunyi /t/ dari dua bahasa mereka ke dalam bahasa Inggris sebagai bahasa asing mereka. Hal ini dapat dilihat dari VOT atau durasi rata-rata di bagian awal kata dalam bahasa Inggris hanya 29 ms. Temuan ini bisa menjadi referensi bagi peneliti lain yang ingin mengetahui durasi dari bunyi /t/ dalam bahasa Sunda dan Indonesia yang dihasilkan oleh penutur bahasa Sunda.

Kata kunci: VOT, /t/ dalam bahasa Sunda, /t/ dalam bahasa Indonesia, Analisis Akustik, PRAAT

I. INTRODUCTION

English has many sound variations of a phoneme, which can be affected by the surrounding sounds. It is commonly called allophones. Fromkin et al (2000) define that a phonetic variant of a phoneme is an allophone of that phoneme. The realizations of a phoneme that are entirely predictable from context are called allophone (Carr, 1999: 39). The existence of the allophones can be predicted by observing the adjacent sounds. The emergence of allophones is a uniqueness in some languages, where each language has a different allophone pattern. For instance, a voiceless alveolar plosive (/t/) is known to have many realizations in General American English (GAE). King and Koffi (2012: 1) reveal that the phoneme /t/ has anywhere from four to eight allophones in GAE. Meanwhile, Sundanese and Indonesian are considered having only two allophones (released and unreleased allophone) of /t/.

Therefore, it is not surprising if English learners, especially Sundanese speakers who have Sundanese as their mother tongue or first language and Indonesian as their second language, get difficulty pronouncing the allophones in English. It may be because their Sundanese and Indonesian have a strong effect on their English pronunciation. People continued using their L1 (first language) processing strategies of linguistic information to communicate in their second language (Culter & Norris, 1998; Koda, 1997) as cited in (Zaikovskii & Koffi, 2019: 74). The theory says the second language is a target language or a foreign language while a foreign language in the present study means English.

As a matter of fact, English is still viewed as a foreign language in Indonesia; it has not been used as an official language. English learners are called EFL learners. Learners that learn English in non-English speaking country are EFL learners (E.g. Japanese people who learn English in their country are EFL learners) while If learners learn English in the country that uses it as a tool for communication officially called as ESL learners (E.g. Hispanic people who are learning English) (Iwai, 2011: 150). In Indonesia, English is commonly learned and used by various groups, such as English students, employees, travelers, interpreters or translators, English teachers, etc. Therefore, Sundanese speakers here are EFL learners.

Nevertheless, the assumption that says the source language pronunciation can influence the target language pronunciation may be incorrect. It is proven by Zaikovskii and Koffi (2019) in Russian speakers, they thought that Russian speakers of English would transfer the short-lag (non-aspiration) Voice Onset Time (VOT) from

their native language into English, but it was not what they saw. They found the duration or VOT of [p] and [k] are as long as the speakers of GAE, even they also gained the VOT of the [t] was longer than the [t] in English. It indicates that we may not assume the way or strategy in producing the phoneme /t/ in Sundanese and Indonesian by Sundanese speakers must be transferred into English. The result may be the opposite like Zaikovskii and Koffi found in their study.

The VOT is a measurement of duration between a voiceless plosive and a vowel sound. According to Lisker and Abramson (1964) define VOT as the time span between the burst of a stop consonant and onset of voicing (as cited in Zaikovskii and Koffi, 2019: 75). The VOT is classified into two types such as positive and negative VOT. Positive VOT is also known as voicing lag that happens when the vocal cord activity begins after the release of the closure while negative VOT is also known as voicing lead that takes place prior to the closure release (Lisker & Abramson, 1964 as cited in Misnadin, 2016: 64). In other words, when the voicing lag appears after the release of the closure or after producing the plosive sound, it is called positive VOT. Meanwhile, if the voicing or the vibration happens before the release of the closure, it is called negative VOT.

According to Kong et al (2012: 726-727) in Zaikovskii and Koffi (2019: 75), there are three types of VOT in world languages, as depicted in the following figure.

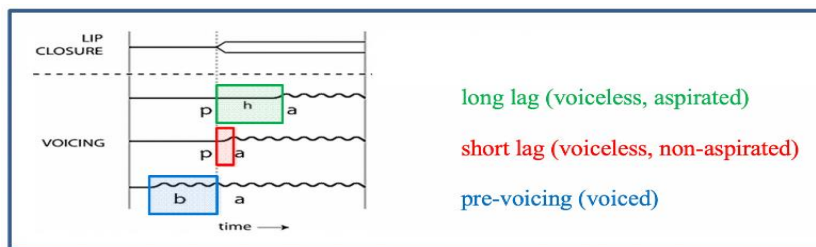


Figure 1. Three Types of VOT

The figure shows three types of VOT: long lag, short lag, and pre-voicing. Long lag or aspiration is producing voiceless plosives by a strong puff of air, short lag or non-aspiration is the opposite (there is not a strong puff of air), and pre-voicing is known as vibration or voicing before the release of the closure. Pre-voicing is also called as negative VOT because it occurs before the release of the closure. Dittmers et al (2017: 4) illustrate these categories of VOT in figure 2. They also list some languages that belong to these categories (as cited in Zaikovskii & Koffi, 2019: 76).

	pre-voicing	short lag	long lag
French, Russian	[b d g] <b d g> voiced	[p t k] <p t k> voiceless	
Turkish	[b d g] <b d g> voiced	[p t k] <p t k> voiceless	
German, English		[b d g] <b d g> voiced	[p ^h t ^h k ^h] <p t k> voiceless

Figure 2. Voicing in Different Languages

The figure above shows that French, Russian, and Turkish have pre-voicing or negative VOT but English and German do not. Yet, they have long lag or aspiration for voiceless stops, while French and Russian do not.

Meanwhile, the pronunciation of voiceless plosives in Sundanese and Indonesian are so simple. The languages have the same pattern where voiceless plosives especially voiceless alveolar plosive (/t/) is released in the initial and medial segment of the word. It is unreleased at the end of the word. For instance, the Sundanese words like *teurab*, *nangtung*, and *reungit* will be pronounced like [tərab], [nanʈun], [rəŋit̚] (Allitadram, 2011). Then, in Indonesian, the words *tinggi*, *satu*, and *sangat* will be pronounced like [tiŋgi], [satu], [saŋat̚] (Aminoedin et al, 1984: 128). If compared with the languages in figure 2, Sundanese and Indonesian only have short-lag types for voiceless plosives.

However, the allophones of /t/ can help Sundanese speakers to master English pronunciation. It can facilitate the learners to speak smoother, more fluent, and more relaxed like a native speaker (Mojsin, 2009: 35 & Cook, 2000: 78-79). As a consequence, Sundanese speakers have to realize and start learning the phonological rules like the allophones of /t/ in English to achieve good and fluent English communication. Even though, their mother tongue or first language (Sundanese) and second language (Indonesian) do not have all variations of /t/ that English has. They still have a chance to be skillful in English pronunciation as long as they learn and implement the allophones of /t/ in English in their English speaking activity. If the Sundanese speakers can find out their best way of learning pronunciation, they can master English pronunciation well. They can make use of sophisticated technology in this era. Even though they do not go abroad, they can still communicate, watch and listen to native English speakers through Youtube, English applications on the smartphone, etc.

The analysis of Sundanese speakers' speech sounds was conducted using acoustic analysis. The tool utilized to analyze was PRAAT software (<http://www.praat.org>) (Boersma & Weenink, 2018). It can show the physics of the speech sounds by showing a figure which consists of the spectrogram and waveform. In the spectrogram, we can see the characteristics of the speech sounds shown by striations or dark bands while in the waveform, we can see the waves of the speech sounds. Acoustic measurements like duration, intensity, and frequency can be found too on the software. The authors here only focus on finding out the duration or Voice Onset Time (VOT) of /t/ in Sundanese, Indonesian, and English produced by Sundanese speakers. By looking at the duration, we can see whether Sundanese speakers transfer the non-aspiration of /t/ from Sundanese and Indonesian into English or their languages are influenced by English instead. Therefore, the present authors intend to conduct the study.

The word environments may have an important role in allophones in English. The different word contexts produce eight realizations of a phoneme /t/ in GAE, such as the unaspirated allophone ([t]), the aspirated allophone ([t^h]), the unreleased allophone ([t̚]), the two flapped allophones (flap [ɾ] and nasalized flap [ɾ̃]), the glottalized allophone ([t̚ʔ]), the fricative allophone ([ʃ]), the affricated allophone ([tʃ]), and the deleted allophone ([∅]) (King & Koffi, 2012: 3-11). Let us see the word environments of the eight allophones of /t/ in GAE. The first allophone is an

unaspirated allophone, it is generally shown by the symbol [t]. Voiceless plosives like /p/, /t/, and /k/ are produced without strong aspiration if it is preceded by a voiceless alveolar fricative (/s/). A /t/ is unaspirated when preceded by a voiceless alveolar fricative, as in *spurt* (Carr, 1999: 37). It indicates that if voiceless plosives especially voiceless alveolar plosive is preceded by /s/, it will not be aspirated. For example, in words *stay*, *stone*, *still*, *star*, and *stomach*, they will be pronounced like [steɪ], [stoun], [stil], [stɑɪ], and [stʌmək].

The second allophone of /t/ is aspirated allophone, it is symbolized by adding superscript like [t^h]. It occurs in initial segment of the single word frequently. The phonological rule of this pattern is like /t/ → [t^h] / __ [+vocalic, +stress] (King & Koffi, 2012: 3). It shows that /t/ will be aspirated or produced with a strong puff of air if it is in the initial segment of the single word in the stressed syllable. The examples are *take* [t^heɪk], *tea* [t^hi], *tell* [t^hɛl], *time* [t^haɪm], *team* [t^him], and *term* [t^hɜːr̩m]. The third allophone is unreleased allophone, it is transcribed with adding the symbol [t̚] (Ogden, 2009: 106), or it is called as no audible release (IPA, 1999). This variation is common in American English, where the voiceless stops or plosives are held in the final segment of the word. It is called unreleased where the tip of the tongue is held in the alveolar ridge until producing the other sound; the phonological formula may be /t/ → [t̚] / __ # (King & Koffi, 2012: 4). That is why it is called unreleased allophone because the tip of the tongue keeps staying in the alveolar ridge until the speaker wants to produce the next sound. The examples are in the words *cat* [kæt̚], *cut* [kʌt̚], *put* [pʊt̚], *bite* [baɪt̚], *can't* [kænt̚], and *get* [gɛt̚].

The fourth allophone is two flapped allophones ([ɾ] and [ɾ̃]). Two flapped allophones are a flap and an alveolar nasal. Those are also the realizations of /t/ in certain environments. A flapped allophone occurs when the /t/ is in the medial segment of two vowels and the first vowel is in the stressed syllable, such as *city* ['sɪɹi] and *duty* ['dʊɹi]. The rule to predict these pronunciations can be stated formally like /t/ → [ɾ] / [+voc, +stress] __ [+syllabic] (King & Koffi, 2012: 5). The /t/ changes to flap if it is between two vowels (*better*), before /l/ (*little*), after /ɹ/ and before a vowel (*party*, *forty*), and it can appear in between two vowels in two words (*it is*, *get up*, *try it on*, *wait a minute*) (Mojsin, 2009: 37-38). The [ɾ] is the sound variation of a phoneme /t/ when it is within two vowels that have stressed syllable in the first vowel. Besides, it can also occur before /l/, after /ɹ/, and between two vowels in two words. Meanwhile, the nasalized flap occurs when there is an /n/ before a /t/ sound, such as in *winter* ['wɪ̃n̩ɹ], *interview* ['ɪ̃n̩ɹvju], and *twenty* ['twɛ̃ɹti]. The symbol is [ɾ̃]. The formula to predict this allophone is /t/ → [ɾ̃] / [+nasal, +alveolar] __ [+voc] (King & Koffi, 2012: 7). This variation is also called a silent /t/ when it is preceded by /n/ such as in words *disappointing*, *international*, *center*, *plenty*, *accountable*, *quantity* (Mojsin, 2009: 36-37). These allophones are the type of American accent where they are not used in another accent as British accent (King & Koffi, 2012: 7).

The fifth allophone is glottalized allophone ([ʔ]). This variation is produced by opening and closing suddenly the vocal folds like when we say [o o] there will be a little gap between the first [o] and the second [o], it occurs when /t/ is between two vowels, the first one is in the stressed syllable and the second vowel is unstressed and followed by /n/ such as *kitten* [kʰɪʔn] (King & Koffi, 2012: 7-8). Therefore, the words that have identical environments, such as in *button*, *mountain*, *eaten*, *written*, and

important will be pronounced like [ˈbʌʔn], [ˈmaʊnʔn], [ˈiʔn], [ˈɪʔn], [ɪmˈpɔɪʔnt]. The sixth allophone is a fricative allophone ([ʃ]). This sound is usually called voiceless postalveolar fricative. This sound can realize the letters *sh* such as in the word *bush and shore*, but in this context, the /ʃ/ is an allophone of a phoneme /t/. This sound frequently appears in suffixes *-tial* and *-tion*, such as in the words *partial* [ˈpɑːʃəl], *presidential* [ˌprezəˈdɛnʃəl], *initial* [ɪˈnɪʃəl], *association* [əˌsɔʊsiˈeɪʃən], *qualification* [ˌkwɒləfəˈkeɪʃən], and *election* [ɪˈleɪʃən].

The seventh allophone is affricated allophone [tʃ]. The affricate is the combination of a voiceless alveolar plosive and a voiceless postalveolar fricative sound. It is one of the allophones of /t/ too, when /t/ is in the middle of two vowels, the first vowel is in a stressed syllable and the second vowel has less stress. It commonly appears in the word that is ended by suffix *-ture*. In other words, the affricate occurs when it is preceded by a vowel in a stressed syllable and before a schwa or middle central vowel (/ə/) and rhotic (/ɹ/), the schwa can be mixed with rhotic or approximant in GAE becomes [ə̃]. The examples of this pattern are in *picture* [ˈpɪktʃə], *nature* [ˈneɪʃə], *adventure* [ædˈvɛnʃə], *feature* [ˈfiʃə], *future* [ˈfjuːʃə], and *agriculture* [ˈægrɪˌkʌltʃə]. The eighth allophone is deleted allophone ([∅]). This pattern is an omission of /t/ when it is in certain words, such as *often*, *listen*, *fasten*, and *soften* (King & Koffi, 2012: 10). This pattern is called deleted allophone because /t/ is totally removed from certain English words. Therefore, those words will be pronounced to [ˈɒfən], [ˈlɪsən], [ˈfæsən], and [ˈsɒfən].

II. METHODS

This study used quantitative and qualitative methods (mix method). The quantitative was used to measure the average of all participants' duration while the qualitative was used to present the data analysis descriptively. The participants were eighteen undergraduate students (nine males and nine females) of the English education study program of A1 class 2017, Faculty of Language Education at IKIP Siliwangi Cimahi in the fifth semester. They originated from Bandung, Tasikmalaya, Sukabumi, Cimahi, Purwakarta, Subang, Majalengka, and Cianjur. As we know, the cities or regencies have an authentic Sundanese Environment because it is still in west java province. Their mother tongue or first language is Sundanese, while their second language is Indonesian. Additionally, their foreign language is English, in other words, they are called Sundanese EFL learners (Sundanese speakers who are learning English as a Foreign Language in the non-English speaking country).

The authors collected the data by recording the participants' voices in the language laboratory. The recorder was Sony PX-470, while the microphone was Boya BY-M1. The step of the recording was the participants were asked to read as natural as possible the Sundanese word contained the phoneme /t/ in the initial segment of the word like *teurab*, then the Indonesian word like *terbit*, and the English word like *Tim* that showed the name of someone in the United States of America which was adopted from King and Koffi (2012). It is important to note, this study focuses on the duration or VOT where it was obtained in the voiceless plosive like /t/ in the initial segment of the word. Therefore, words like *teurab* in Sundanese, *terbit* in Indonesian, and *Tim* in English were selected because they had /t/ in the initial position of a word and the next sound was a vowel sound.

After obtaining the data, the authors converted the data using Freemake Audio Converter software. It was because the data produced by the recorder were mp3 files while PRAAT software could not read the mp3 files. Hence, the data needed to be converted into Wav files. Converting steps were clicking *Audio* → select the file → *to WAV* → *Optimal quality* → *Convert*. As mentioned in the preceding description, the present authors used PRAAT software to analyze the VOT or duration of the data. The techniques of data analysis using the software were opening files, view and edit, then measuring. In opening files, the authors clicked *Object* → *Open* → *Read from File...* (Styler, 2017). The display of the opening files will be like the following figure.

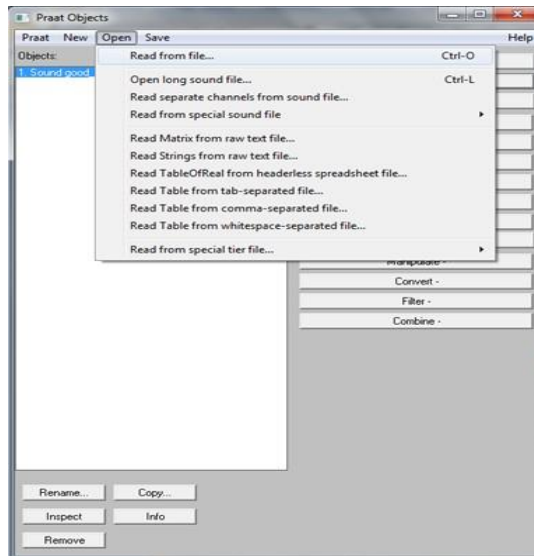


Figure 3. Object Window

View and edit were the process to show the image of the voice recording. The steps were clicking *Object* → *View & Edit* (Styler, 2017). Object means the selected data, view and edit were the options button in the right position. After clicking view and edit button, the display will be like the figure below.

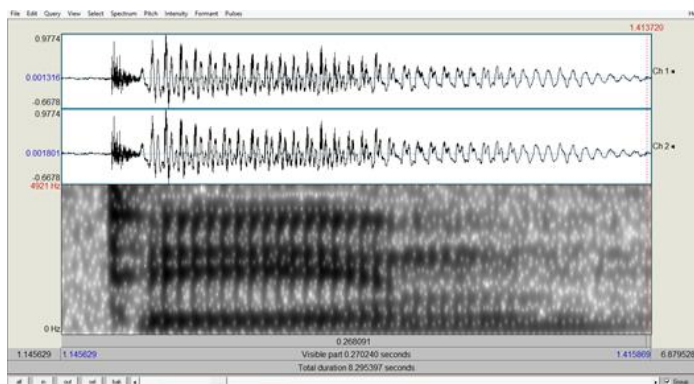


Figure 4. PRAAT Editor Window

The waveforms of the data showed two waves. It was because the data were gained from the external recorder had stereo signal (two signals).

The last step was measuring. It was the process to get the duration of the data. The authors adjusted the PRAAT editor window by choosing one of the five buttons in the bottom left corner of the window. There are *all* to see all of the files, *in* to zoom in, *out* to zoom out, *sel* to make current selection fill the window, and *bak* to zoom back to the previous zoom level (Styler, 2017). The authors then highlighted the desired segment such figure as follows.

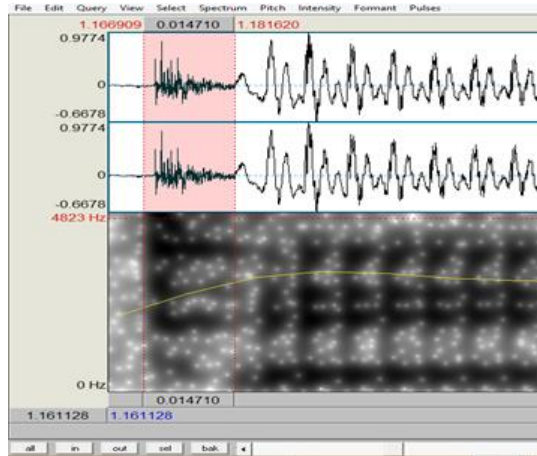


Figure 5. Highlighted Segment

The authors then selected *Query* in ten menus above the waveforms. The steps were *Highlight* → *Query* → *Query Time Domain* → *Get Total Duration* (Styler, 2017). As a consequence, the info window showed the duration of the data in millisecond (ms).

III. RESULT AND DISCUSSION

This section will present the result of acoustic measurement particularly the VOT or duration of /t/ in Sundanese, in Indonesian, and in English. It was focused on the initial segment of the word. It is important to know that the basic duration or non-aspiration of /t/ in GAE is around 32 ms, the symbol was [t] while the aspiration of /t/ is twice longer or more, the duration of aspiration of /t/ produced by the native American English speaker was 103 ms, the symbol was [t^h] (King & Koffi, 2012). As a consequence, if the duration or VOT of Sundanese speakers are twice longer or more than the basic duration, they aspirate the /t/ sound in Sundanese, Indonesian, and English. The opposite result occurs if the average duration is around the basic duration or shorter than it. The data of all participants' acoustic measurements will be shown in the following table.

Table 1. All participants' Duration of /t/

Participants' Code	Duration		
	<i>teurab</i>	<i>terbit</i>	<i>Tim</i>
1M	10 ms	21 ms	14 ms
2M	14 ms	18 ms	29 ms

3M	32 ms	30 ms	24 ms
4M	41 ms	23 ms	17 ms
5M	28 ms	29 ms	64 ms
6M	20 ms	8 ms	17 ms
7M	22 ms	29 ms	94 ms
8M	20 ms	30 ms	16 ms
9M	21 ms	17 ms	43 ms
10F	17 ms	19 ms	28 ms
11F	19 ms	12 ms	22 ms
12F	19 ms	13 ms	25 ms
13F	13 ms	17 ms	8 ms
14F	11 ms	19 ms	66 ms
15F	11 ms	21 ms	23 ms
16F	16 ms	20 ms	13 ms
17F	27 ms	20 ms	11 ms
18F	17 ms	12 ms	12 ms
Average	20 ms	20 ms	29 ms

Note: M is Male, F is Female, and ms is a millisecond

The table showed the duration of /t/ in initial segment of the word *teurab*, *terbit*, and *Tim* pronounced by Sundanese speakers. The average duration of /t/ in *teurab* and *terbit* were identical namely 20 ms. Meanwhile, in the word *Tim*, it had a longer duration namely 29 ms. Based on the parameter in the previous description, the duration of /t/ in those three languages (Sundanese, Indonesian, English) was characterized as non-aspiration or unaspirated allophone ([t]).

Now let us see the spectrogram and waveform of *teurab*, *terbit*, and *Tim* produced by Sundanese speaker. The figures were taken from one of the participants' voice. It was shown as the representation of the unaspirated allophone form in Sundanese, Indonesian, and English.

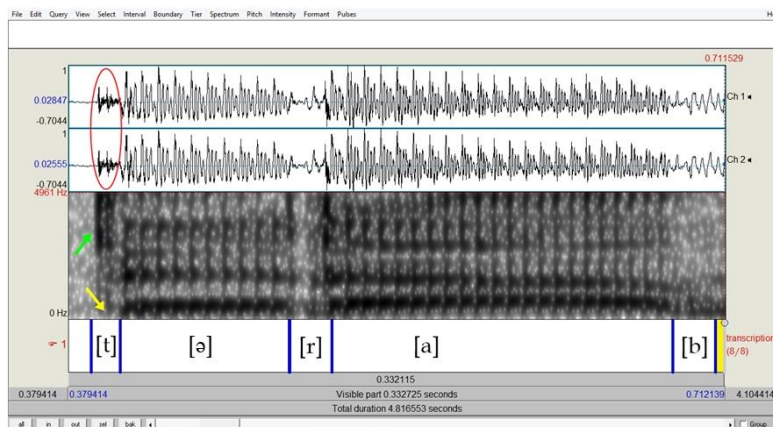


Figure 6. The Spectrogram and Waveform of [t] in *teurab*

The figure above is the [t] or unaspirated allophone of the word *teurab* in Sundanese word which means *burp* in English. First of all, the authors analyzed the spectrogram then the waveform because the spectrogram could be called as core of acoustic analysis that had more cues than the waveform. The authors only focused on the preceding segment of the word. As we can see, there are not dark bands in the voicing bar, it is shown by a yellow arrow. The position was called a voicing bar because we could see the sounds were voiced or voiceless sounds by looking at the bar. If there were horizontal dark bands, it was voiced sound while if there were not horizontal dark bands, the result was the opposite. The yellow arrow showed the absence of the dark bands, so it was definitely voiceless sound.

Additionally, the green arrow pointed to the dark vertical line. Actually, the dark vertical line was a strong hint related to the voiceless plosive sound, especially voiceless alveolar plosive. The dark vertical line was called a release burst (the process of releasing the tip of the tongue from the alveolar ridge). The striations after the release burst were called aspiration if the horizontal striations were long and strong or non-aspiration if the striations were so short and faint. Therefore, by looking at the spectrogram, we had realized the preceding segment of the spectrogram was a voiceless alveolar plosive which was unaspirated ([t]). Meanwhile, the waveform had two waves. But all of them were the same pattern. The characteristic of the unaspirated allophone of the phoneme /t/ could be found too by observing the abrupt pulse in the preceding segment of the waveforms (shown by red oval). The next segments after [t] in figure 6 were the remaining sounds like shown in the phonetic transcription under the spectrogram.

The next figure was the [t] in the Indonesian word taken from one of the participants' voice.

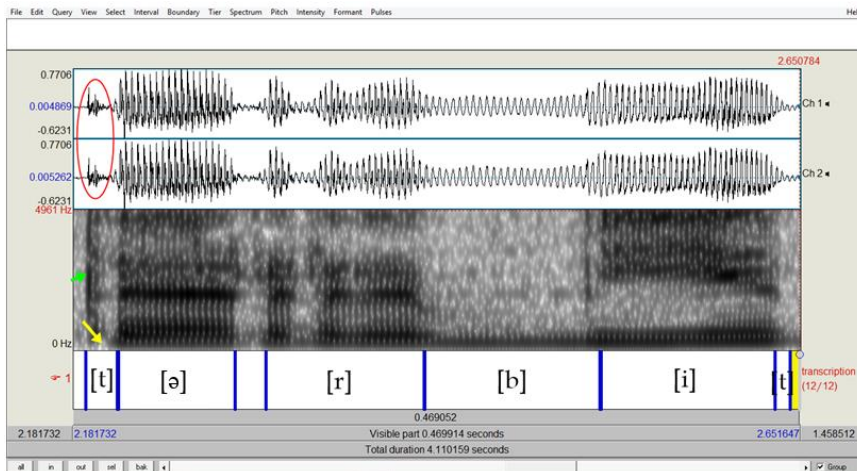


Figure 7. The Spectrogram and Waveform of [t] in *terbit*

Figure 7 shows the spectrogram and waveform of the [t] in the word *terbit* that means *rise* in English. The characteristic of the [t] in this figure was almost identical with the [t] in figure 6. There were not the horizontal dark bands in the voicing bar as shown by a yellow arrow. The dark vertical line of [t] in *terbit* that

represented a release burst was taller than [t] in *teurab*. Both waveforms of [t] in this figure were similar to the waveforms in figure 6. There was abrupt pulse or prominent vertical spike in the preceding segment of the waves which represented a release burst. A stronger release burst will be a higher abrupt pulse.

The last figure is the spectrogram and waveform of [t] in the English word. It is shown to see the similarities and the differences of [t] in Sundanese, Indonesian word and English word.

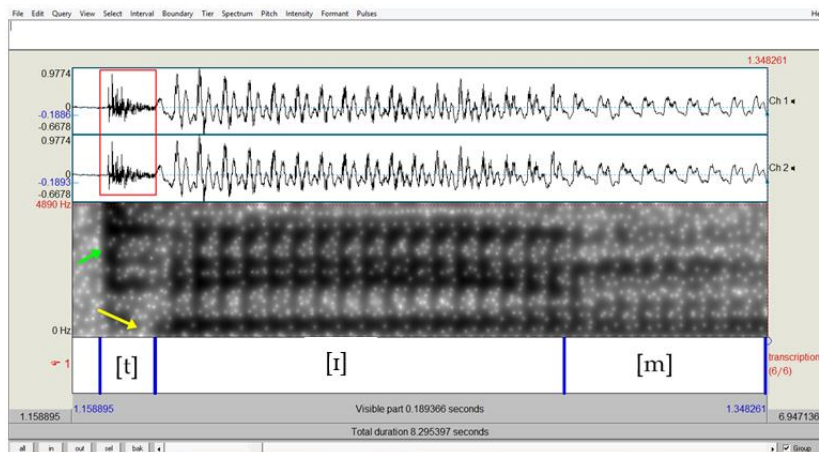


Figure 8. The Spectrogram and Waveform of [t] in *Tim*

The figure shows the spectrogram and waveform of *Tim*. Based on the result of duration analysis, Sundanese speakers pronounced the phoneme /t/ in *Tim* as unaspirated allophone ([t]). It was investigated not only in the duration analysis but also in the spectrogram and waveform analysis. Native speakers of American English pronounced the /t/ in the word *Tim* in English as aspirated allophone ([t^h]) with a length 103 ms (King & Koffi, 2012). Meanwhile, Sundanese Speakers produced it as unaspirated allophone ([t]) with an average duration of only 29 ms.

The form of [t] in spectrogram and waveform in *Tim* was only differentiated by longer striations after a release burst and the abrupt pulses were a bit higher and longer than [t] in *teurab* and *terbit*. It indicates that the similarity of [t] in the three languages was not only in the average duration but also in the form of the spectrogram and waveform. The following graph is the result of the average duration of [t] in the three languages.

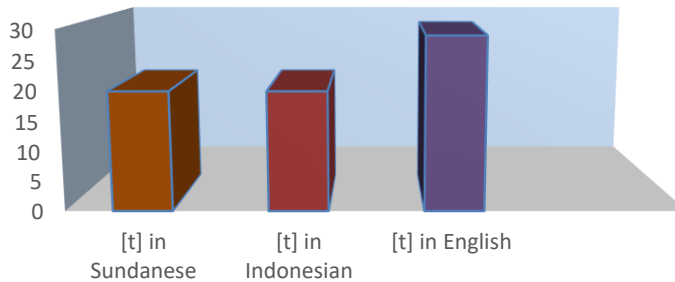


Figure 9. The Duration of [t] in Sundanese, Indonesian, and English

The result of the average VOT or duration of [t] produced by Sundanese speakers in Sundanese and Indonesian showed a surprising result where the durations were equal. The VOT of [t] in English produced by them was 29 ms, it was not too far from the two languages. It meant that the participants produced the phoneme /t/ in those two languages in the same way as an unaspirated allophone. Actually, the participants used Sundanese in their family and society environment while Indonesian was used in the formal situation or in certain conditions like when they talked with other people who did not understand Sundanese. Sometimes, they mixed those languages in their daily communication because the interference of Indonesian and even English into Sundanese was not unavoidable in this era.

IV. CONCLUSIONS

The result of this study showed the participants produced the phoneme /t/ as unaspirated allophone ([t]) in Sundanese and Indonesian in the same way. It can be seen from the average VOT of both languages that lasts 20 ms. They also transfer the way in producing the /t/ in the initial segment of the word in Sundanese and Indonesian into English. Therefore, they produce the /t/ in English as unaspirated allophone too, the average VOT is not too long, namely 29 ms. Based on the result, we can know the basic duration of /t/ in Sundanese and Indonesian in the environment like *teurab* and *terbit* or /t/ → [t] before /ə/ (**middle and central vowel or schwa**) in the initial segment of the word.

This finding can be valuable information related to VOT or duration of voiceless alveolar plosive /t/ in Sundanese, Indonesian, and English. Even though it does not generalize and represent to all Sundanese Speakers but at least it can be a reference to other researchers who have similar research.

REFERENCES

- Allitadram. 2011. "Fonetik Bahasa Sunda" Retrieved from:
<https://allitadram.wordpress.com/2011/12/18/fonetik-bahasa-sunda>.
- Aminoedin, N. A., Soedjatno, Rasaq, H. S., Marsoedi, I. L., & Taryono. 1984.
Fonologi Bahasa Indonesia: Sebuah Studi Deskriptif. Jakarta: Perpustakaan Pusat Jakarta.

- Boersma, P., & Weenink, D. 2018. "PRAAT: Doing Phonetics by computer [Computer Program]." Version 6.0.39. Retrieved from: <http://www.praat.org/>
- Carr, Philip. 1999. *English Phonetics and Phonology*. Malden: Blackwell Publishing.
- Cook, A. (2000). *American Accent Training*. USA: BARRON'S.
- Culter, Anne & Norris, Dennis. 1988. "The role of strong syllables in segmentation for lexical access." *Journal of Experimental Psychology: Human perception and performance*, 14(1), 113-121. <http://dx.doi.org/10.1037/0096-1523.14.1.113>
- Dittmers, Tetyana., Gabriel, Christoph., Krause, Marion., & Topal, Sevda. 2017. "Positive transfer from the heritage language? The case of VOT in German/Turkish and German/Russian learners of L3 French, Russian, and English." Universität Hamburg. Retrieved from: http://wa.amu.edu.pl/L3_workshop/Dittmers_Gabriel_Krause_Topal_2017_L3Poznan.pdf
- Fromkin et al. 2000. *Linguistics: An Introduction to Linguistic Theory*. USA: Blackwell Publishers.
- IPA (International Phonetic Association). 1999. *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge, U.K: Cambridge University Press.
- Iwai, Yuko. 2011. "The Effects of Metacognitive Reading Strategies: Pedagogical Implications for EFL/ESL Teachers." *The Reading Matrix*. 11 (2), 150-159.
- King, Amber & Koffi, Ettien. 2012. "An Acoustic Account of the Allophonic Realization of /t/." *Linguistic Portfolios*: Vol. 1, No 12.
- Koda, Keiko. 1996. Orthographic knowledge in L2 lexical processing: A cross-linguistic perspective. In J. Coady & T. Huckin (Eds.), "Second Language Vocabulary Acquisition: A Rationale for Pedagogy (Cambridge Applied Linguistics, pp. 35-52)." Cambridge: Cambridge University Press. doi:10.1017/CBO9781139524643.005
- Kong, E. J., Beckman, M. E., & Edwards, J. (2012). "Voice onset time is necessary but not always sufficient to describe acquisition of voiced stops: The cases of Greek and Japanese." *Journal of phonetics*, 40 (6), 725-744.
- Lisker, Leigh & Abramson, Arthur S. 1964. "A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements." *WORD*, 20:3, 384-422, doi: 10.1080/00437956.1964.11659830
- Misnadin. 2016. "The Phonetics and Phonology of the Three-Way Laryngeal Contrast in Madurese." Thesis. Edinburgh: The University of Edinburgh.
- Mojsin, Lisa. 2009. *Mastering the American Accent*. United States of America: Barron's Educational Series.
- Ogden, R. (2009). *An Introduction to English Phonetics*. Edinburgh: Edinburgh University Press Ltd.
- Zaikovskii, Mikhail & Koffi, Ettien. 2019. "An Acoustic Phonetic Account of VOT in Russian- Accented English." *Linguistic portofolios*: Vol. 8, No 7.