CHAPTER II

REVIEW OF RELATED LITERATURE

This chapter discusses the framework that is related of the study. This chapter will discuss previous studies, description acoustic phonetics, elements of acoustic phonetics, and acoustic cues for consonants.

2.1 Previous Studies

The study of pronunciation had been discussed also by several students of Department of English. The writer found three works which discussed about pronunciation. First, the work from Alfriani Ndandara entitled *A Study English Consonant Pronunciation by Sumbanese Students Live in Kupang and Salatiga* (2009). This work only examined Sumbanese students who live in Kupang and Salatiga as informants to pronounce English consonant and this work was based on Dobrovosky and Katamba (1997). Its scope records of English consonant pronunciation by Sumbanese students. The result of study was there were 27 allophone variations from 24 consonant sounds based on the subject's pronunciation. Second, the work from Linda Susanti entitled *A Study of Pronunciation of English Sound by Sundanese Students of English* (2003). This work discusses about how Sundanese students from Department of English pronounce the English sound and the study based on the theory from Dobrovosky and Katamba (1997). Its scope records of English sound pronunciation by Sundanese students. The result of the study was there were 5 similarities between English and Sundanese consonant sounds based on the manner of articulation, 5 similarities based on the place of articulation and 14 similarities based on the way of sound are produced by speech organs. Thus, there were 5 differences between English and Sundanese consonant sound based on the manner of articulation, 3 differences based on the place of articulation, and 11 differences based on the way production of sounds. Also there were 10 similarities between English and Sundanese vowel sounds and 11 similarities between English and Sundanese vowel sounds. The work from Ima Halimatus Sa'diyah entitled *English Pronunciation by Madurese Students in Surabaya* (2015). This work discussed about how Madurese students in Surabaya pronounce English and Madurese students as informants. This work wa based on the theory from Michael Dobrovosky and Katamba (1996) and Soendjono

Dardjowidjojo (2009). Its scope records English sound pronunciation by Madurese students. The result of the study was that there were 10 consonant sounds that have been mispronounced in the initial position, 10 sounds in the middle position, and 12 sounds in the back position. And regarding the vowel sounds, the researcher also finds there are 8 sounds that have been mispronounced in the initial position, 11 sounds in the middle position, and 6 sounds in the final position. The researcher also finds 8 English consonant sounds from 24 English consonant sounds that are pronounce correctly by the subject.

2.2 Acoustic Phonetics

According to Francis Katamba (1989:2) acoustic phonetics is the study of the physical properties of speech sounds using laboratory instruments. In phonetic acoustics we will study the physical aspects of sound waves, from the way produce sound through the speech organs, from the lungs, then the throats, and passing out from the mouth, after being released the sound will turn into a wave through the air and then to the hearer.

2.3 Elements of Acoustic Phonetics

Acoustic phonetics has several elements, there are: Speech production, wave, energy, and stop. Each element will be discussed in this session.

2.3.1 Speech Production

Sound is produced when air is set in motion, when the air come out of the lungs goes up the windpipe and into the larynx, then produce through the vocal cords (Peter Ladefoged, 2009:4). According to Katamba (1989:4), the sound will come out when the human speaks and all the sounds they produce is the result of contracting muscles. The muscles are also called vocal cords. Vocal cords can be pulled apart when produce voiceless sounds [p], [t], [k], [s], [f], [h], [θ], [ʧ], and [ʃ] or drawn closer together when produce voiced sound [b], [d], [g], [z], [m], [n], [v], [l], [r], [w], [j], [ŋ], [ʒ], [ð], and [ʤ] especially at their back or posterior ends. As air passes through the space between the vocal tracts, which is called glottis, different glottal states are produce, based on the position of vocal cords. According Soenjono Djarjowdjojo (2003:36), when human produce voiceless sounds, the vocal cords are slightly open but not vibrating, if sound is produced with vibration then it is called voiced

2.3.2 Wave

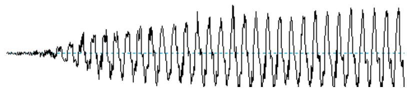

Figure 1: Showing the waveform

Above picture is the visual of waveform that shows the wave of the sound and if the sound is sufficiently short (Dr. Will Styler, 2017:13). According to Johnson (2003:4), articulatory sound that produced by human speech form into sound waves, than received by the eardrum. In the figure 2 is a sound waves model that produce by the vocal tract, it can be a vowel or a consonant sound.

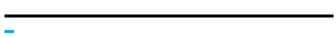

Figure 2: Waveform straight line waves

Figure 3: Waveform small waves

The waveform has several models. In figure 3 shown the waveform into straight line when the vocal cord do not produce the sound or there is no sound activity on the waveform. While in figure 4 shown the waveform into small waves, it means there is a sound activity that start produce by the vocal cords, it can be when vocal cords produces vibration or hissing sound.

2.3.2.1. Periodic and Aperiodic Waves

In phonetic acoustic there are two types of sounds: periodic and aperiodic waves. Periodic waves has two types: simple periodic waves and complex periodic waves. Periodic waves has regular wave and aperiodic waves has irregular wave.

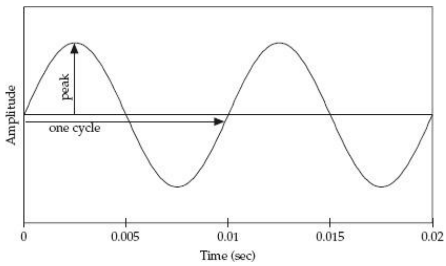

Figure 4: Simple periodic waves by Keith Johnson p. 7

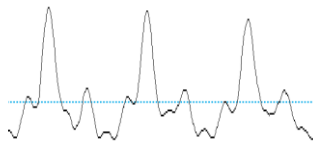

Figure 5: Simple periodic waves by the data

From the pictures above is the visually of periodic waves. Simple periodic waves has regular wave, this wave has regular period repetition. Each repetition of the pattern is called a cycle and the duration of a cycle is its period (Keith Johnson, 7:2003)

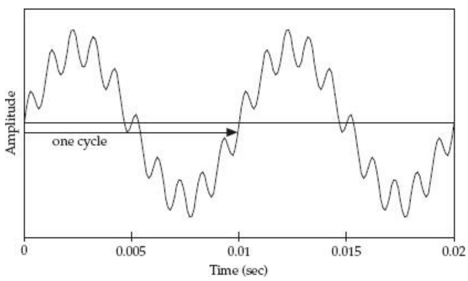

Figure 6: Simple periodic waves by Keith Johnson p. 8

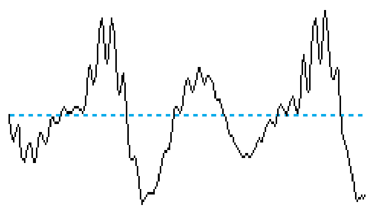

Figure 7: Simple periodic waves by the data

Complex periodic waves has the same characteristic like simple periodic waves. Complex waves has regular wave but forming into dense little waves. It can be found when producing sentences.

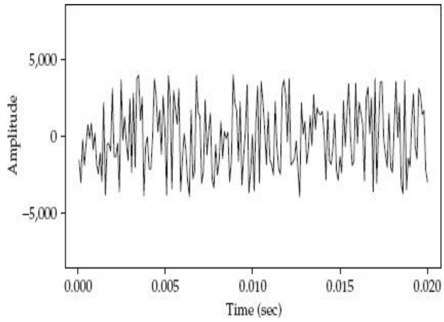

Figure 8: Aperiodic waves by Keith Johnson p. 13

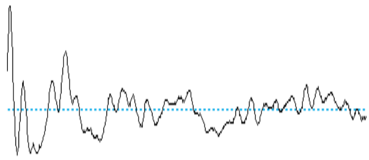

Figure 9: Aperiodic waves by the data

Aperiodic waves is not like periodic waves that have regular waves. This wave has an irregular shape. Can be seen in figure 16. They have either a random waveform or a pattern that doesn't repeat. In the aperiodic waves it may indicate the existence of hissing sound or white noise or these are various types of clanks and busrt which produce a sudden pressure fluctuation that is not sustained or repeated over time. For example door slames, baloon pops, and electrical clicks, or it can be consonants sound. (Keith Johnson, 12:2003).

2.3.3 Energy

The spectrogram bar position is under the waveform. Spectrogram is showing the spectral energy of the sound over time while (Dr. Will Styler, 2017:13).

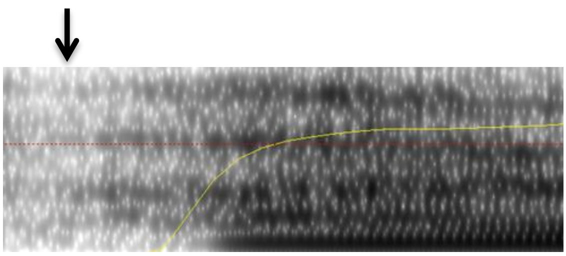

Figure 10: The spectrogram

The waveform indicates sound waves due to the sound activity produced by the vocal tract, so the spectrogram indicated the energy of the waveform. If there is no sound activity produced by the vocal tract, the spectrogram shows no dark colour as shown by the arrow. While spectrogram has a dark colour it means there is a sound activity that start produce by the vocal cords. The dark colour caused by the sound energy, it can be indicate whether the sound is bilabial, alveolar or velar.

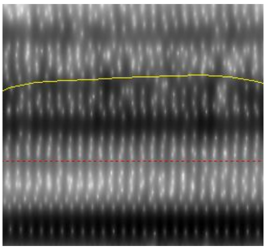

Figure 11: Spectrogram into shaded form

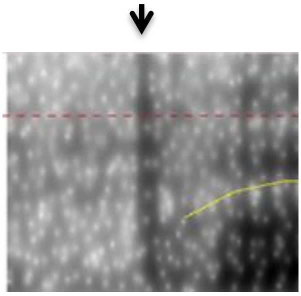

Figure 12: Spectrogram into vertical tight line

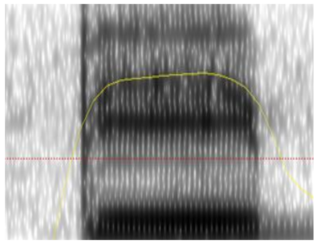

Figure 13: Spectrogram form a pattern

The spectrogram has several characteristics. Figure 6 shown the spectrograms into the shaded form when the image is zoomed. In figure 7 spectrogram into vertical tight line, it can be release energy. In figure 8, spectrogram can form a pattern, it can be bilabial, alveolar or velar sound.

2.3.3.1 White Noise and Hissing Sound

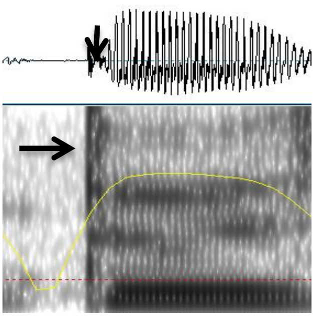

Figure 14: A piece of "white noise" picture on the spectrogram.

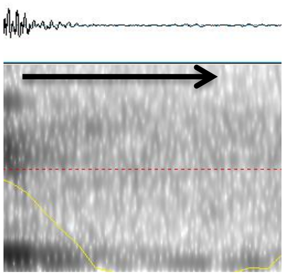

Figure 15: The spectrogram that has been zoomed in the "white noise" section

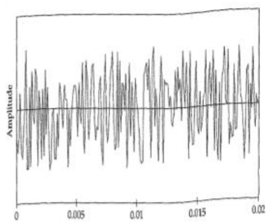

Figure 16: Waveform white noise that has been zoomed.

White noise is the condition of spectrogram that affected by surrounding environment. White noise also can produce when vocal tract produce fricative sound. So white noise is "natural sound" which is can affected from human production or surrounding environment. For example it sounds like radio static or wind blowing through trees. White noise has irregular waves that shown in figure 17 and has a faint black colour on spectrogram that shown in figure 16 (Keith Johnson, 2003:12).

2.3.4 Stop

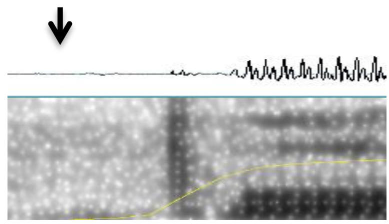

Figure 17: Showed the pattern on the spectrogram.

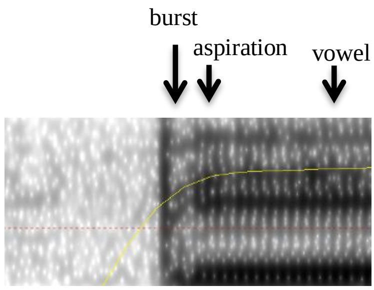

Figure 18: Showing the burst, aspiration and vowel.

The figure 17 and 18 are the spectrogram that shows the patterns sound on the spectrogram. On the spectrogram, the stops sound showed the aspiration. In the figure 22, after the spectrogram zoomed there is visible aspiration, the aspiration has short distance between burst and vowel. Stops are produced because of the obstacle that causes the airflow closure through to the vocal tract in a moment then prevents the airflow out of the mouth. It is affected by the source from which they came out, when the air comes out of the lungs and goes to the larynx and then when the air flow gets obstacle there. Also the waveform showing the small waves indicates a burst in figure 18. Stops or plosives are found as bilabial, dental, alveolar, palatal, velar, uvular and glottal points or articulation. To identify the stop there are three stages in time: shutting – closure – release.

2.4 Acoustic Cues for Consonants

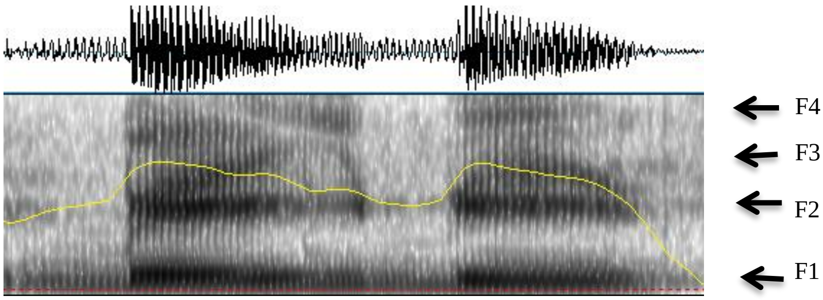

Figure 19: Spectrogram, waveform and formant.

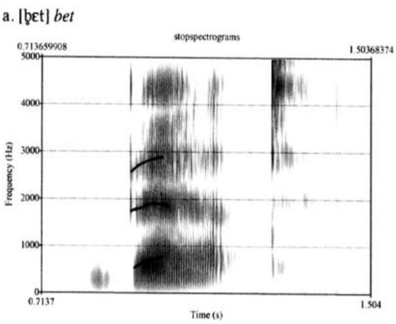

Figure 20: Bilabial *bet* constriction

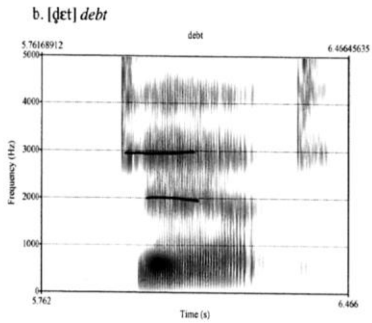

Figure 21: Alveolar *debt* constriction

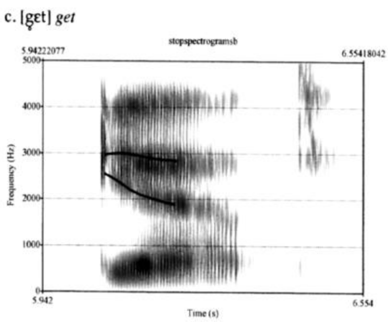

Figure 22: Velar *get* construction

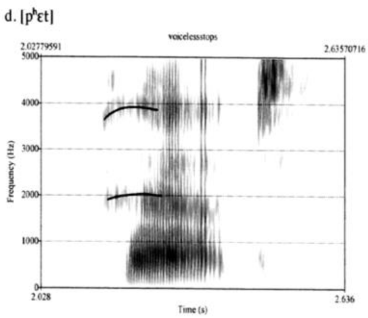

Figure 23: bilabial p^h*et* construction

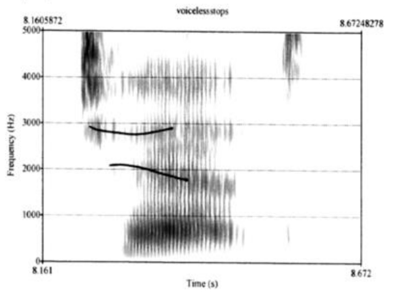

Figure 24: Alveolar t^h*et* construction

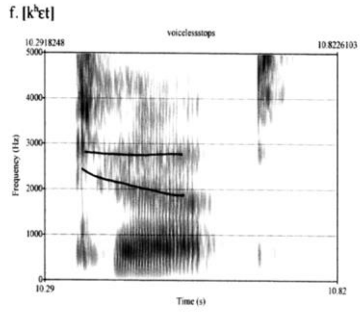

Figure 25: Velar k^h*et* construction

Waveform and spectrogram has related to each other. Waveform shows the indicate of waves, it can be periodic or aperiodic waves and can be vowels or consonants. While the spectrogram can show the energy and pattern of waveform, it can be bilabial, alveolar or velar sound. According to Ladefoged (2003:131), the acoustic dimensions represented by F1, F2, F3 and F4 do not correspond directly to the auditory/articulatory dimensions vowel height, vowel backness, and lip rounding. In F1 characterizes represent vowel height, F2 characterizes both backness and lip rounding, F3 helps distinguish the high of vowels, and F4 shows the quality of speaker's voice.

Figures 20, 21, 22, 23, 24, and 25 above shows voicing during the stop closure at the beginning of each of these and the stop closure at any place results in a low F1 locus. Figure 20 bilabial constriction lowers the second and third formants, so the F2 and F3 loci are relatively low. Alveolar constriction typically produces an F2 locus around 1700-1800 Hz as can be seen in Figure 21. Figure 22 shows constriction raises the F2 locus, and may lower the F3 locus, resulting in a common origin point for the two, this is known as *velar* pinch.

Figures 23, 24 and 25 are the formant transitions that take place in the aspiration noise. So they are harder to see clearly. In Figure 24 the burst noise of [t] has the highest frequency, with highest amplitude noise is above 3.500 Hz. Figure 25 [k] is the next highest in frequency – notice the high amplitude noise or lower frequencies than [t]. [p] has the lowest frequency burst noise. The burst noise of bilabial stop is also the lowest in amplitude, the alveolar and velar bursts are louder.

2.5 Duration

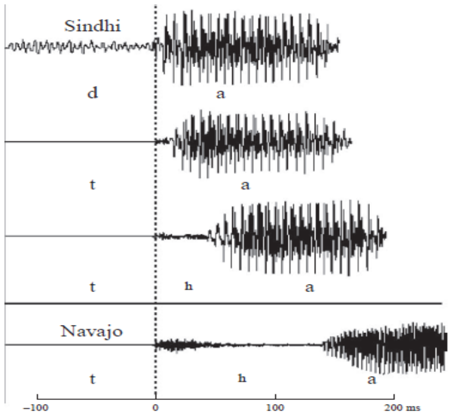

Figure 26: Waveform showing stops with different length duration

In general, duration is the time range of an phenomenon. In Acoustic phonetics duration is called voice onset time (usually abbreviated VOT). According to Ladefoged the interval between the release of a closure and the start of the voicing is called voice onset time. The VOT can also measure the length of duration in milliseconds (m) in Figure 26.

2.6 English Consonant Sounds

Katamba (2003:4) consonants are produced by obstructing in some way the flow of air through the vocal tract. Ladefoged (2011:36), English has 24 conconants ([p], [b], [t], [d], [k], [g], [s], [z], [m], [n], [f], [v], [l], [r], [h], [w], [j], [ŋ], [ʒ], [θ], [ð], [ʧ], [ʃ], and [ʤ].

2.7 Organ of Speech

The sound will come out when the human speaks and all the sounds they produce is the result of contracting muscles. According to Francis Katamba (1989:4) the organs of speech are:

1. Lips

2. Teeth
3. Alveolar ridge
4. (hard) palate
5. Velum
6. Uvula
7. Tongue tip
8. Tongue blade
9. Front of the tongue
10. Back of the tongue
11. Mouth cavity
12. Nose cavity
13. Pharynx
14. Epiglottis
15. Oesophagus
16. Glottis
17. Larynx

The term vocal tract is used to refer to the air passages which the air enters on leaing the larynx. The vocal tract has two parts: the oral cavity, which is the air passage offered by the mouth and the nasal tract, which is the air passages provided by the nose.

2.8 Place of Articulation

Based on the place of articulation the English consonants can divided into 9 groups (Daniel Jones, 1922:14) those are:

a. **Bilabial**: articulated by the two lips. Example: [p], [b], [w], and [m].
b. **Labio-dental**: articulated by the lower lip against the upper teeth. Example: [f] and [v].
c. **Dental**: articulated by the tip of the tongue against the upper teeth. Example: [θ] and [ð].
d. **Alveolar**: articulated by the tip or blade of the tongue against teeth ridge. Example: [t], [d], [l], [s], [z], and [n].
e. **Retroflex**: those in the formation of which the tip of the tongue is curled upwards towards the hard palate. Example: [r]

f. **Alveo palatal**: articulated by the blade of the tongue against the teeth ridge raising of the main body of the tongue towards the palate. Example: [ʧ], [ʤ], [ʃ], and [ʒ].
g. **Palatal**: articulated by the front of the tongue against hard palate. Example: [j].
h. **Velar**: articulated by the back of the tongue against the central and toward part of the soft palate. Example: [k] and [g].
i. **Glottal**: articulated in the glottis. Example: [h] and glottal stop [?].

2.9 Manner of Articulation

According to Jones (1922:14), manners of articulation are various configuration of the lips, tongue, velar (velum), and glottis positioned in different ways to produce different sound types. Based on the manners the articulation, the English consonants can be divided into 6 groups:

a. Plosive/stop: are made with a complete and momentary closure of airflow through the vocal track, thus preventing the escape of air via the mouth. Example: [p], [b], [d], [t], [k], [g], [m], [n], [ŋ], and [?]
b. Fricative: are consonants produces with continuous airflow through the mouth: example: [f], [v], [θ], [ð] [s], [z], [h], [ʃ], and [ʒ].
c. Affricative: when a stop articulation is released, the tongue moves rapidly away from the place of articulation. Some non-continuant consonant show a slow release of the closure. Example: [ʧ] and [ʤ].
d. Liquid: partial closure of speech organs and airflows over sides of tongue. Example: [l]. near closure of speech organs and the air flows down middle of tongue. Example: [r]
e. Nasal: when sounds (other than plosive and nasal consonants) are pronounced with simultaneous lowering of the soft palate, so that the air passes through the nose as well through the mouth. Example: [m], [n], and [ŋ]
f. Semi vowel: independent vowel-glides in which the speech organs start by forming a weakly articulated close or fairly close vowel and immediately move to another sound of equal or greater prominence: the initial vowel-position is not held on for any appreciable time. Example: [w] and [j].