CHAPTER II

REVIEW OF RELATED LITERATURES AND THEORETICAL FRAMEWORK

This chapter presents the review of literature that relates to the topic of the research. They are the basic concept of phonology and phonetics, segmental feature, vowel and consonant sounds, place of articulation, manner of articulation, articulatory of vowel, factors affect in pronunciation, previous research, and theoretical framework. The topic is presented below.

2.1 Phonology and Phonetics

People pronounce some words to convey the idea when they are communicating with others. When people speak and communicate with others, automatically they produce words and sounds (phonology) from their speech organs (phonetics). Phonology and phonetic are highly affects in pronouncing English words. In English, when a word pronounced incorrectly, it can change the meaning of the word. Study about sounds in English is called phonology. According to Yule (2010, p. 42) phonology is essentially the descriptions of the systems and patterns of speech sounds in a language. It is studies the phonemes (phonemic transcription). Each one of these meaning-distinguishing sounds in a language is described as a phoneme. For examples of phonemes, compare the changes of meaning in: 'snack – snake'; 'desert' – 'dessert'; 'bake – back'; 'stick' – 'steak'. In fact, many people may have difficulty with particular sounds,

sound combinations or with putting particular sounds in particular positions. Another expert mentioned phonology describes the way sounds function within a given language or across languages Roach (as cited in Andi-pallawa, 2013, p. 104). So, it can conclude that phonology is the study about sounds pattern which influence the meaning of the words in English pronunciation.

Phonetic is the study of the sounds made in the production of human languages. Finegan (2008, p. 80) stated that phonetic has two principal branches; articulatory phonetic and acoustic phonetic. Articulatory phonetics focuses on the human vocal apparatus and describes sounds in terms of their articulation in the vocal tract; it has been central to the discipline of linguistics. Acoustic phonetics uses the tools of physics to study the nature of sound waves produced in human language; it is increasingly important in linguistics with attempts to use machines for interpreting speech patterns in voice identification and voice-initiated mechanical operations.

Pronunciation is the manner where someone utters a word. Pronunciation in English (also known as phonology), refers to the production of sounds that we use to make meaning. Sounds have been pronounced by human, will produce different meaning. Burns and Claire (as cited in Akyuz, 2017, p. 390) emphasize pronunciation refers to the phonology of the language – or the meaningful perception and production of the sounds of that language and how they impact on the listener.

Tench (2011, p. 4) stated that phonology refers to pronunciation as a system in itself – how many vowels there are in the spoken form of the language (not the five vowel letters), and how many consonants there are, where the sounds can occur in words, what combination of sounds are allowed, etc. Phonetics refers to the pronunciation of the sounds themselves – how they are made, how they differ, how they sound in different positions of a word and how they sound in different combinations, etc. And for the study of the pronunciation of words in English, an extra set of symbols is needed to extend the use of the letters of the alphabet.

The 'square' brackets, [t], enclose letters of the alphabet. Whole words in ordinary spelling which are used as examples are in italics. 'Slant' brackets, /t/, enclose phonetic symbols in broad transcription, i.e. phonemes; whole words in broad transcription are also enclosed in 'slant' brackets, e.g. /rat/ rat. Mirza (2015, p. 487) states that the English pronunciation components are divided into two there are the segmental and suprasegmental features of the language. Segments are the sounds of vowels and consonants known as phonemes. The suprasegmental aspects focus on intonation, stress, and timing.

2.2 Segmental Feature

Pronunciation includes the segmental and suprasegmental features of the language (Yates as cited in Mirza, 2015, p. 487). The segmental aspects of pronunciation have been focused in traditional methods mainly because they are the easiest to notice and pronounced related to the written letters (sounds of vowels and consonants). Recently, however, some speakers see that the suprasegmental aspects of pronunciation, including intonation, rhythm, and stress might highly affect intelligibility and can convey better impression of the speaker.

2.2.1Consonant Sounds

Sound is produced when air is set in motion. The air supply is provided by the lungs. The sound source is in the larynx, where a set of muscles called the vocal folds (or vocal cords) is located. The filters are the organs above the larynx: the tube of the throat between the larynx and the oral cavity, which is called the pharynx, the oral cavity, and the nasal cavity. These passages are collectively known as the vocal tract (O'Grady et al., 1997, p. 19). Every speech has different sound from every other speech because of a unique combination of features in the shape of mouth and tongue and move parts of the vocal apparatus while speaking.

Knight (2012, p. 9) said consonants are sounds made with a lot of constriction in the mouth, so that the air coming up from the lungs gets squashed. Consonant in English pronunciation is included in segmental phonemes. According to Roach (as cited in Hadi, 2015, p. 49), there are 44 sounds in the British English language sound system, which are 24 consonants and 12 vowels.

Consonants						
Р	pin	f	Fan	h	hello	
В	bin	v	Van	m	more	
Т	to	θ	Think	n	no	
D	do	ð	The	ŋ	sing	
Κ	cow	S	Sun	1	live	
G	got	Z	Zoo	r	red	
t∫	church	ſ	She	j	yes	
dʒ	judge	3	Measure	W	wood	

 Table 2.1 Example of Consonants

1. Consonant Articulation

A consonant is a speech sound produced by a partial or complete closure of part of the vocal tract (Finegan, 2008, p. 85). Speech sound can be characterized in in terms of their articulatory properties – by where in the mouth and how they are produced and consonant can be describe in terms of three properties: voicing, place of articulation, and manner of articulation.

a. Voicing

It can be seen by distinguishing between /s/ and /z/. When you pronounce a long, continuous /z/ and a long continuous /s/, you'll notice that the position of your tongue within your mouth remains the same but, these sounds are noticeably different. You can feel if you touch your larynx. There is the vibration that you feel from your larynx if you say /z/ in along continues, it called voicing.

b. Place of Articulation

Each point at the airstream can be modified to produce a different sound is called a place of articulation. Places of articulation are found at the lips, within the oral cavity, nasal, in the pharynx and at the glottis. According to Yule (2010) there are seven kinds place of articulation; bilabials, labiodentals, dentals, alveolar, palatal, velar, and glottal which is the explanation can be seen below.

1) **Bilabials**

These are sounds formed using both (= bi) upper and lower lips (=labia). The initial sounds in the words pat, bat and mat are all bilabials. They are represented by the symbols [p], which is voiceless, and [b] and [m], which are voiced. We can also describe the [w] sound found at the beginning of way, walk and world as a bilabial.

2) Labiodentals

These are sounds formed with the upper teeth and the lower lip. The initial sounds of the words fat and vat and the final sounds in the words safe and save are labiodentals. They are represented by the symbols [f], which is voiceless, and [v], which is voiced. Notice that the final sound in the word cough, and the initial sound in photo, despite the spelling differences, are both pronounced as [f].

3) Dentals

These sounds are formed with the tongue tip behind the upper front teeth. The initial sound of thin and the final sound of bath are both voiceless dentals. The symbol used for this sound is $[\theta]$, usually referred to as "theta." It is the symbol you would use for the first and last sounds in the phrase three teeth.

The voiced dental is represented by the symbol [ð], usually called "eth." This sound is found in the pronunciation of the initial

sound of common words like the, there, then and thus. It is also the middle consonant sound in feather and the final sound of bathe. The term "interdentals" is sometimes used for these consonants when they are pronounced with the tongue tip between (= inter) the upper and lower teeth.

4) Alveolar

These are sounds formed with the front part of the tongue on the alveolar ridge, which is the rough, bony ridge immediately behind and above the upper teeth. The initial sounds in top, dip, sit, zoo and nut are all alveolar. The symbols for these sounds are easy to remember – [t], [d], [s], [z], [n]. Of these, [t] and [s] are voiceless whereas [d], [z] and [n] are voiced.

Other alveolar are the [l] sound found at the beginning of words such as lap and lit, and the [r] sound at the beginning of right and write.

5) Palatals

If you feel back behind the alveolar ridge, you should find a hard part in the roof of your mouth. This is called the hard palate or just the palate. Sounds produced with the tongue and the palates are called palatals (or alveo-palatals). Examples of palatals are the initial sounds in the words shout and child, which are both voiceless. The "sh" sound is represented as $[\int]$ and the "ch" sound is represented as [t]. So, the word shoe brush begins and ends with the voiceless palatal sound $[\int]$ and the word church begins and ends with the other voiceless palatal sound [t].

One of the voiced palatals, represented by the symbol [3], is not very common in English, but can be found as the middle consonant sound in words like treasure and pleasure, or the final sound in rouge. The other voiced palatal is [dʒ], which is the initial sound in words like joke and gem. The word judge and the name George both begin and end with the sound [dʒ] despite the obvious differences in spelling.

One other voiced palatal is the [j] sound used at the beginning of words like you and yet.

6) Velars

Even further back in the roof of the mouth, beyond the hard palate, you will find a soft area, which is called the soft palate, or the velum. Sounds produced with the back of the tongue against the velum are called velars. There is a voiceless velar sound, represented by the symbol [k], which occurs not only in kid and kill, but is also the initial sound in car and cold. Despite the variety in spelling, this [k] sound is both the initial and final sound in the words cook, kick and coke. The voiced velar sound heard at the beginning of words like go, gun and give is represented by [g]. This is also the final sound in words like bag, mug and, despite the spelling, plague.

The velum can be lowered to allow air to flow through the nasal cavity and thereby produce another voiced velar, represented by the symbol $[\eta]$ typically referred to as "angry." In written English, this sound is normally spelled as the two letters "ng." So, the $[\eta]$ sound is at the end of sing, sang and, despite the spelling, tongue. It occurs twice in the form ringing. Be careful not to be misled by the spelling of a word like bang – it ends with the $[\eta]$ sound only. There is no [g] sound in this word.

7) Glottals

There is one sound that is produced without the active use of the tongue and other parts of the mouth. It is the sound [h] which occurs at the beginning of have and house and, for most speakers, as the first sound in who and whose. This sound is usually described as a voiceless glottal. The "glottis" is the space between the vocal folds in the larynx. When the glottis is open, as in the production of other voiceless sounds, and there is no manipulation of the air passing out of the mouth, the sound produced is that represented by [h].

c. Manner of Articulation

Manner of articulation is the way how English consonants pronounced. According to Yule (2010, p. 31) manner of articulation in consonants are consists of stops, affricative, nasal, liquid, fricative and glides which stated below.

1) Stops

Phonemes [p], [b], [t], [d], [k], [g] are all produced by some form of "stopping" of the air stream (very briefly) then letting it go abruptly. This type of consonant sound, resulting from a blocking or stopping effect on the air stream, is called a stop (or a "plosive").

2) Affricative

If you combine a brief stopping of the air stream with an obstructed release which causes some friction, you will be able to produce the sounds $[\mathfrak{Y}]$ and $[d\mathfrak{z}]$. These are called affricates and occur at the beginning of the words cheap and jeep. In the first of these, there is a voiceless affricate $[\mathfrak{Y}]$, and in the second, a voiced affricate $[d\mathfrak{z}]$.

3) Nasal

Most sounds are produced orally, with the velum raised, preventing airflow from entering the nasal cavity. However, when the velum is lowered and the air stream is allowed to flow out through the nose to produce [m], [n] and [ŋ], the sounds are described as nasals. These three sounds are all voiced. The words morning, knitting and name begin and end with nasals.

4) Liquid

The initial sounds in led and red are described as liquids. They are both voiced. The [1] sound is called a lateral liquid and is formed by letting the air stream flow around the sides of the tongue as the tip of the tongue makes contact with the middle of the alveolar ridge.

5) Fricative

The manner of articulation used in producing the set of sounds $[f], [v], [\theta], [\delta], [s], [z], [f], [3]$ involves almost blocking the air stream and having the air push through the very narrow opening. As the air is pushed through, a type of friction is produced and the resulting sounds are called fricatives.

6) Glides

The sounds [w] and [j] are described as glides. They are both voiced and occur at the beginning of *we, wet, you* and *yes*. These sounds are typically produced with the tongue in motion (or "gliding") to or from the position of a vowel and are sometimes called semi-vowels.

The conclusion of the consonant articulation including the place of articulation and manner of articulation are shown in the table below.

	PLACE OF ARTICULATION										
			Bilabial	Labiodent al	Dental	Alveolar	Postalveol ar	Alveopala tal	Palatal	Velar	Glottal
	Plosive	Vl	р			t				k	
Z		Vd	b			d				g	
[OI]	Affricative	Vl						t∫			
LA		Vd						dʒ			
CU	Nasal	Vl									
RTI		Vd	m			n				ŋ	
ΕЧ	Lateral	Vl									
02		Vd				1					
NEI	Fricative	Vl		f	θ	S		ſ			Η
AN		Vd		V	ð	Z	r	3			
Σ	Semi vowel	Vl									
		Vd	W						j		

Table 2.2 Chart of Consonants

2.2.2 Vowel Sounds

Vowel sounds are produced by passing air through different shapes of the mouth, with different positions of the tongue and of the lips, and with the air stream relatively unobstructed by narrow passages except at the glottis (Finegan (2008, p. 89). Sounds in which are produced with relatively little obstruction in the vocal tract is called vowel (O'Grady et al., 1997, p. 22). Vowels are produced when the airstream is voiced through the vibration of the vocal cords in the

larynx, and then shaped using the tongue and the lips modify the overall shape of the mouth. The position of the tongue is a useful reference point for describing the differences between vowel sounds.

Tench (2011, p. 25) stated that, in English the movement of the tongue has three possible directions: either higher towards the front of the roof of the mouth, that is, in the general direction towards the /I/ or /i/ vowel; or higher towards the back, that is, in the general direction towards the / σ / or /u/ vowel; or towards a central area, that is, in the general direction of the / Λ / or /3:/ vowel. These three directions are called front closing, back closing and centering. In addition, Barman (2010, p. 25) mentioned that, vowel can be described with their placement in what is called 'cardinal vowel diagram'. A cardinal vowel diagram is drawn like a trapezium as seen in the following figure.



Figure 2.1 Cardinal Vowel Diagram (Barman, 2010)

The trapezium illustrates the shape of the tongue. The frontness/backness and openness/closeness of vowels are visually explicit with their relative positions (Barman, 2010, p. 25). Vertically, the trapezium has four lines, based on the height of the tongue: close, half-close, half-open and open; and horizontally, it has three lines, based on the location of the tongue: front, central and back. Vowel sounds are produced with a relatively free flow of air and they are all typically voiced.

English vowels are divided into two major types, simple vowels (also called pure vowels or monophthongs) and diphthongs. According to Tench (2011), if the tongue is relatively steady, they are called monophthongs (or 'pure' vowels); if there is a degree of movement by the tongue, they are called diphthongs.

	Vowels	Diphthongs			
i:	bead	ег	cake		
Ι	hit	IC	toy		
υ	book	аі	high		
u:	food	IƏ	beer		
e	left	ບຈ	fewer		
ə	about	39	where		
3:	shirt	ວບ	go		
э:	call	av	house		
æ	hat				
Λ	run				
a:	far				
D	dog				

 Table 2.3 Example of Vowels

1. Simple Vowel (Monophthong)

Simple vowels do not show a noticeable change in quality. The vowels of pit, set, cat, dog, but, put and the first vowel of suppose are all simple vowels. The symbols [ə] (called schwa) and [Λ] (called caret or wedge) represent similar sounds. Both occur in the word above /əb Λ v/.

We use [ə] to represent a mid-central vowel in unstressed syllables, such as the second syllable of buses /bʌsəz/ and the second and third syllables of capable /kepəbəl/.

2. Diphthong

Diphthongs are the combinations of two vowel sounds, one gliding towards the other and there are eight diphthongs in English. Finegan (2008, p. 91) stated that diphthong is a vowel sound for which the tongue starts in one place and glides to another. English diphthongs show changes in quality that are due to tongue movement away from the initial vowel articulation towards another vowel position (Barman 2010, p. 30), This change in vowel quality is clearly perceptible in words such as say, buy, cow, ice, loud, go and boy.



Figure 2.2 Chart of Diphthong (Roach, 1980)

3. Articulatory of Vowel

Finegan (2008, p. 91) stated that, to create differences among vowels, languages can exploit other possibilities besides tongue height and tongue backness. Vowels can have tenseness, rounding, lengthening, nasalization, and tone.

a. Tenseness

Languages can make a distinction between vowels that is characterized as tense versus lax. These labels represent a set of characteristics that distinguish one set of vowels from another. The contrast between [i] of peat and [1] of pit is in part a tense/lax contrast; likewise for the vowels in bait/bet and in cooed/could.

b. Rounding

In English high front vowels tend automatically to be unrounded (and high back vowels to be rounded), some languages have rounded and unrounded front vowels.

c. Length

English has two of each vowel type—one long, the other short. The pronunciation of long vowels is held longer than that of short vowels. Long vowels are commonly represented with a special colon after them in phonetic transcriptions. The contrast of letter [i] and [i:] to sense differences in the duration of vowels, pronounce the English words beat, bead, bit.

d. Nasalization

All vowel types can be nasalized by pronouncing the vowel while passing air through the nose (as for nasal stops) and through the mouth. Nasal vowels are indicated by a tilde (~) placed above the vowel symbol.

e. Tone

Vowel may be pronounced on several pitches and be perceived by the native speakers of these languages as different sounds. Typically, a vowel pronounced on a low pitch contrasts with the same vowel pronounced on a higher pitch. In English there are symbol of rising tone (´); a falling-rising tone (`), in which the pitch begins to fall and then rises sharply; and a falling tone (`), in which the pitch falls sharply.

2.3 Factors Affect in Pronunciation

Kenworthy (as cited in Dhillon, 2016, p. 125) mentioned there are factors that affected pronunciation, they are:

- The native language. The more differences of sounds characteristics of the native language and English, the more difficulties someone will have in pronouncing English.
- 2. The age factor. This factor is contradictory. Some researchers found that age determines the accuracy of a learner's pronunciation; other

researchers argued that age confers no immediate advantage in pronouncing foreign sounds.

- 3. Amount of exposure. It is tempting to view this simply as a matter of whether the speaker is living in an English-speaking country or not. If this is the case, then the speaker is 'surrounded' by English and this constant exposure should affect pronunciation skills.
- 4. Phonetic ability. This skill has been variously termed 'aptitude for oral mimicry', 'phonetic coding ability' or 'auditory discrimination ability'. Researchers have designed tests which measure this ability and have demonstrated that some people are able to discriminate between two sounds better than others, and/or are able to mimic sounds more accurately.
- 5. Attitude and identity. It has been claimed that factors such as a person's 'sense of identity' and feelings of 'group affiliation' are strong determiners of the acquisition of accurate pronunciation of a foreign language.

2.4 Previous Research

In the previous study, some researchers have already done with the same research about pronunciation. First, according to Demircioglu (2013), he has done the same research about pronunciation, but Demircioglu concern the diphthongs, the voiced and the unvoiced "th" $/\theta$ / and $/\delta$ / sounds in state primary, secondary and high school in Turkish. In his research, He concluded that, Turkish language is hardly to say vowel in juxtaposition or in double vowel. They only say with a single pure vowel sound. In the other hand, the diphthongs are difficult for Turkish to articulate when they speak English. So, the Turkish learners are used to articulating the word forming monophthongs sounds because of the structure of Turkish language.

Second research, according to Wewalaarachchi et al (2017), in their research, they concern in vowel, consonant and lexical tones in Monolingual and bilingual English-Mandarin by toddlers. In the result, there are three primary findings emerged from this study; 1) There were no differences between bilingual and monolingual toddler responses to correct pronunciations, nor were there differences in sensitivity to vowel, tone, and consonant mispronunciations, 2) There were differences in the efficiency with bilingual and monolingual children processed correct pronunciations and mispronunciations. 3) Bilingual and monolingual children exhibited different patterns of relative sensitivity to vowel, consonant, and tone mispronunciations; Bilingual children were most sensitive to vowel variation, followed by consonant variation and then tone variation, whereas monolingual children were equally sensitive to vowel and tone variation and less so to consonant variation.

2.5 Theoretical Framework

This session presents the theoretical framework of the research. The main theory is taken from Finegan's theory in term of vowel sounds; single vowel (monophthong) and diphthong, and consonant sounds which is following by Yule's theory. The theoretical framework has shown as below.



Figure 2.3 Theoretical Framework