CHAPTER 2
# LITERATURE REVIEW

This chapter presents a brief explanation of some theories that support the study. The theories relate to the nature of pronunciation, techniques and activities in pronunciation learning, factors affecting pronunciation learning, types of pronunciation learning strategies, and difficulties and strategies in pronunciation learning.

2.1 The Nature of Pronunciation

Pronunciation refers to the way on how the individual determines the utterances. As Pennington and Richards (1986) stated that pronunciation is accepted as the basis for individual expression of meaning in the form of phonological segments. Similarly, pronunciation refers to the production of phonological or linguistic sounds which include stress, intonation, and how it affects the listener so that they can understand the content (Antaris & Omolu, 2019; Burns & Claire, 2003). Thus, this reflects the expression that pronunciation leads to the production of individual words through sounds which become the language for communication.

As the initial recognition by Pennington and Richards (1986) that pronunciation contains segments, the following are segments classified by Burns and Claire (2003) which are a way of distinguishing the main characteristics of pronunciation.
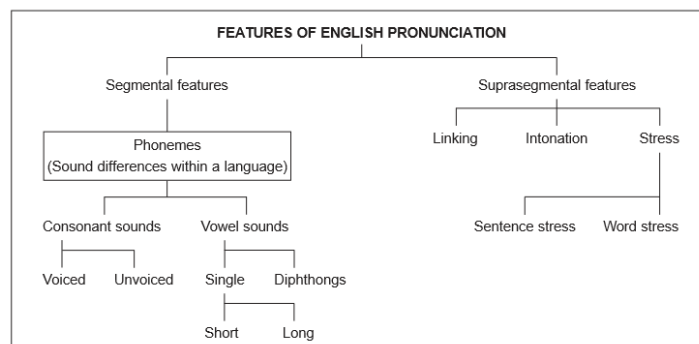
Figure 2. 1 Diagram of Pronunciation Features

The figure 2.1 shows classification features of English pronunciation namely segmental features and supra-segmental features which can be differentiated as below.

2.2.1   Segmental Features (Micro Level)

Segmental features also known as phonemes which are sound differences within a language. Phonemes can change the meaning of the word when an individual pronounces them incorrectly but can still be interpreted reasonably when an individual accurately considers the meaning of how each sound is produced. This feature is a minimal unit of sound defined in phonetic terms (Pennington & Richards, 1986) as the following figure.

Figure 2. 2 The Phonemic Chart
(source: https://albaenglish.co.uk/blog/improve-english-pronunciation-phonemic-chart/)

As depicted in figure 2.2, the set of segmental features consist of two categories, namely consonant sounds and vowel sounds (Burns & Claire, 2003).

a)   Consonant Sounds

Consonant sounds can be voiced where a part of the mouth is closed and the air behind it is released suddenly. They also can be unvoiced when the air is pushed through a narrow part of the mouth.

b)   Vowel Sounds

Vowel sounds are all voiced, and probably single or a combination. It called as diphthongs as well as there is a movement from one vowel sound to

another (like /eɪ/ in tr<u>ai</u>n, say, or plane), or if it describes three vowel sounds, it called as triphthongs like /aʊə/ in our or power).

2.2.2   Supra-segmental Features (Macro Level)

This feature is related to speech extending across whole stretches of language (prosody) or it is a kind of group of segmental is generally applied. They are classified as word stress, intonation, and linking words.

a)   Word stress

Word stress is usually when one syllable in a word sounds more prominent than the others which focuses on making it long and loud to convey the meaningful part. Thus, in the meaningful part, in order to foreground which meaning is important some words are given more prominence than others. For example, compare these utterances:

Table 2. 1 Word stress

Can YOU take the phone?
Can you take the PHONE?

In catch the comparison, it can be seen by the uppercase word of each utterance. First, the word "YOU" is stressed to describe that the subject is not someone else. While, second, the word "PHONE" is used to emphasise that the object is the phone, for example not the laptop.

b)   Intonation

Intonation related to the melody of the way sounds are rising or falling pitch of an utterance. For example, notice the differences between the following utterances:

Table 2. 2 Intonation

Can you take the phone?
Can you take the phone?

Along with the differences, when the arrow goes up, it noted that the sound of the utterance is the rising pitch which is used for a request. Besides, when the arrow goes down, it describes the utterance as a falling pitch which is used for a command.

c) Linking Words

Linking or connection word refers to the way the last sound of one word is joined to the first sound of the next word. There are some ways to produce the linking word, it runs together to link consonant to vowel, consonant to consonant, vowel to vowel, and shorten some sounds and leave others out altogether, as in the following table.

Table 2. 3 Linking words

Consonant to vowel	An Indonesian animal
Consonant to consonant	next week
	seven months
Vowel to vowel	Where (r) are you?
	You (w) ought to
	Saturday (y) evening
Shortened sounds	When do they arrive?
	/ə/ /ə/
	Five o'clock
	/ə/

Additionally, in the part of vowel to vowel, it should be noticed that some sounds such as /r/, /w/ and /j/ (y) are inserted to link adjacent words ending and beginning with a vowel. Besides, for the shortened sounds, when words begin with an unstressed sound they are often pronounced as a short schwa /ə/ sound.

By the whole of nature pronunciation, it indicated pronunciation as a foundation of making meaning that articulated in sound towards communicative effectiveness.

2.2 Techniques and Activities in Pronunciation Learning

The wealth of techniques and activities in pronunciation learning provides considerable opportunities for students to integrate into pronunciation class. In particular, in EFL contexts, students need to learn to hear the differences between phonemes, especially if these differences are not present in their L1. In order to achieve these differences, Department of English Education students must recognise a number of techniques and activities in learning pronunciation, for example as suggested by Kelly (2000) as follows.

2.2.1 Drilling Technique

The most common technique practised in the classroom might be drilling and it helps students achieve better pronunciation and helps them remember new language items. Practically, this technique is similar to a spelling activity that encourages students to bring up a previous phrase or structure through prompts, pictures, mime, etc. Students usually listen to a native speaker's audio and highlight the way they pronounce words and then repeat it. As students implement this technique into their class, they will simultaneously build their confidence to pronounce some words or structures. Besides, they are needed to ensure that certain words or phrases are intelligible. Thus, students must spend time putting new words and phrases into their vocabulary (Kelly, 2000).

2.2.2 Taping Technique

This technique helps students who experience pronunciation difficulties that are difficult to overcome. In this case, students can compare their sounds, stress, and intonation with, for example, native speakers, other students or fellow L1 speakers who do not have the same difficulties. Students usually listen to speakers speak through the audio, then record their voices and compare the results repeatedly until they feel the pronunciation is better (Kelly, 2000).

2.2.3 Minimum Pairs Activity

The activity concerns words or utterances that only differ by one phoneme. Practically, students can create words that have certain sounds to understand individual phonemic symbols relevant to English. Students pronounce at least two or three words that have certain sounds until they understand the differences in pronunciation or sound production of each word. Apart from that, this activity can also implement students' vocabulary in order to practise sounds. Therefore, it will be more successful if students choose their current words in the previous vocabulary learning to indicate sounds in context (Kelly, 2000).

2.2.4 Homograph and Homophone Activities

Similar to the previous activity, homophones are words that have the same pronunciation but different spellings. Practically, students listen to native speaker audio and differentiate each word. This activity allows students listening practice as they have to listen to a word or sentence and have to choose which one from a list of correctly spelled words they heard (Kelly, 2000).

2.2.5 Listening Activity

Along with having the goal of language learning within the classroom, students also pursue language instruction that they can better comprehend and implement outside the class. Alternatively, listening activities involve a natural speaker voice at a normal speed, aiming to sound as realistic as possible. In this regard, students need to listen to how language is used and pronounced in the narration or conversation context. For instance, before engaging in a listening task, students can pay attention to the meaning and pronunciation of specific language aspects and practise them in a highly controlled way. Usually, they listen to native speaker audio, or sometimes they watch a video without subtitles in the form of narrative or conversation. Therefore, this can play a key role in helping students pay attention to the existence of pronunciation features (Kelly, 2000).

2.2.6 Reading Activity

Even though the medium is the written word, reading activities can be successfully integrated with pronunciation. Here, students are allowed to get the gist of the text they read or by establishing the type of text used, followed by some more detailed assignments to focus on specific details. In this activity, they are also aware of pronunciation features in terms of spelling, stress and intonation, and linking sounds between words in connected speech (Kelly, 2000).

2.3 Factors Affecting Pronunciation Learning

English is an international communication language that can influence many countries in daily life. For example, in Singapore, English is spoken typically following the Singaporean English dialect. This difference in English pronunciation may cause discomfort and misunderstandings for other listeners (Zhang, 2009). Therefore, pronunciation is highlighted for EFL students, especially in countries where English is not officially the state language. In line, Fadillah (2020) mentioned that someone who lives in a non-English speaking country faces many difficulties in pronouncing English and it leads to the factors that affect their pronunciation. The researchers have conducted research to determine factors that influence students difficulties in English pronunciation, such as students are low motivation, lack of confidence, unsupported environment, lack of knowledge, less attention, influenced by their mother tongue, and foreign language instructions (Fadillah, 2020; Antaris & Omolu, 2019).

Meanwhile, Kenworthy (1987) classified six factors that affect pronunciation learning, namely the native language, the age factor, amount of exposure, phonetic ability, attitude and identity, and motivation and concern for good pronunciation. The following is the explanation of each factor.

2.3.1 The Native Language

The factor is a noticeable impact on students when they learn a new language. Specifically, some phonemes that are present in English may not be present in a student's mother tongue. For example, Motair and Mahmed (2022) mentioned that Arabic EFL students had problems with consonant sounds,

because the Arabic and English sound systems differ in many aspects; English contains 24 consonants, while Yemeni Arabic contains 28 consonants. Therefore, some English consonants are absent in the Yemeni Arabic system, such as /v/, /p/, /ŋ/. Meanwhile, in Tasikmalaya, students speak Sundanese where the characteristic is that they tend to replace consonant such as /f/ and /v/ sounds into /p/ sound. Thus, students may find it challenging to pronounce words correctly when they come across novel phonemes in words in a new language. Students may find it difficult to pronounce some words correctly since they do not have equivalent sounds in their original tongue, which might hinder their language learning process as a whole.

2.3.2 The Age Factor

This factor underlines the significant impact of age on pronunciation. When individuals learn a second language, they tend to show different pronunciation patterns based on age as they begin their language acquisition journey. If a person begins learning a second language in childhood, their ability to imitate the sounds of the language is often more pronounced, allowing them to develop an accent in the second language that is similar to the native speaker. In contrast, if a person begins the journey of learning a second language in adulthood, they may face greater challenges in acquiring native-like pronunciation as their brain's plasticity, or ability to adapt and learn new things, tends to decrease with age. As a result, adults may find it more difficult to imitate a native speaker's accent, and they may retain certain nuances or elements of pronunciation that are characteristic of their native language. In essence, the age at which a person begins learning a second language plays an important role in determining the extent to which they can achieve native-like pronunciation (Kenworthy, 1987).

2.3.3 Amount of Exposure

This factor is related to the intensity of exposure to English received by students. Some students are live in English-speaking countries but are often in non-English-speaking environments, such as their homes where another language is spoken. In this case, students may not have consistent exposure to English in

their daily lives, and this limited linguistic environment may impact their ability to develop strong pronunciation skills in English. On the other hand, there are students who live in non-English speaking countries but still use English in various aspects of their lives, such as in educational settings. This way can make students actively engaged with the language that makes a significant contribution to improving their pronunciation (Kenworthy, 1987).

2.3.4   Phonetic Ability

Phonetic ability, also known as having an ear for a language, is crucial in language acquisition. People who have strong phonetic skills often find it easier to learn a new language as they have been exposed to it since childhood. Additionally, some individuals may have a phonetic code or understanding for interpreting and reproducing the sounds of a language that differentiates them from others. In contrast, individuals who do not have high phonetic abilities may find it more difficult to understand the intricacies of pronunciation. However, it is important to note that although some people may have a natural inclination to learn phonetics, language acquisition is a skill that can be developed and improved through practice and exposure over time (Kenworthy, 1987).

2.3.5   Attitude and Identity

In this factor, attitude includes the learner's character, approach and emotional attitude towards the target language and its speakers. A positive attitude can foster a more receptive and engaged learning experience, potentially contributing to improved pronunciation skills. In addition, a person's sense of identity and feelings of group affiliation significantly influence the acquisition of accurate pronunciation in a foreign language. When individuals strongly identify with a particular cultural or linguistic group, they may be more motivated to adopt and imitate that group's pronunciation patterns. The feelings of group affiliation refer to the acceptance and recognition of individuals in the community to encourage students to perfect their pronunciation skills (Kenworthy, 1987).

2.3.6   Motivation and Concern for Good Pronunciation

Motivation functions as the driving force behind the learner's efforts to achieve and maintain good pronunciation. For those who recognise the importance of pronunciation, motivation emerges as an important factor in their efforts to achieve good pronunciation. This motivation may come from a variety of sources, such as the desire to communicate effectively or the desire to be perceived as a skilled speaker. In contrast, students who consider pronunciation to be less important may be influenced by different motivations or priorities in their language learning journey. They may prioritise other aspects of language acquisition, such as vocabulary acquisition or grammar proficiency, over pronunciation. In short, recognition of the importance of pronunciation in language learning is subjective and varies among students (Kenworthy, 1987).

Entrenched on these, it can conclude that the ability to learn a new language, particularly pronunciation is influenced by several factors that might be faced by the EFL students, where English is not officially the state language.

2.4 Types of Pronunciation Learning Strategies

As foreign learners, specifically EFL students, pronunciation has become challenging in language skills. For instance, based on this preliminary study, students admitted they have difficulties during pronunciation learning, especially when they were pronouncing consonant and vowel sounds. To overcome it, a number of ways should be taken by the students to make their pronunciation easier and more effective, and it refers to the pronunciation learning strategies. At this point, Oxford (1990) suggests several learning strategies that involve memory strategy, cognitive strategy, compensation strategy, metacognitive strategy, affective strategy, and social strategy. Furthermore, the strategies are also interpreted on the basis of Rokoszewska (2012) strategy taxonomy as follows.

2.4.1   Memory Strategy

The strategy helps students link the pronunciation knowledge that they store in memory into information that they can use for a long time. For example, the students can repeat a word aloud or silently many times, connect a word's

pronunciation with a situation where they heard or read it, visualise the word, and connect sounds with pictures. They can use phonetic symbols or their own codes, regularly review new pronunciations, highlight important information in their notes, and listen to recorded word lists (Rokoszewska, 2012; Oxford, 1990).

2.4.2   Cognitive Strategy

Cognitive strategy helps students understand and implement their pronunciation knowledge through cognitive processes to learn it or complete the tasks. In this strategy, the students can listen to recordings to identify the pronunciation of new words, imitate the target language speakers or phonetic drills, and use media (Rokoszewska, 2012; Oxford, 1990).

2.4.3   Compensation Strategy

This strategy is able to overcome the limitations and inability of individual knowledge in pronunciation. It is possible for students to guess intelligently and overcome limitations in pronunciation (Rokoszewska, 2012; Oxford, 1990).

2.4.4   Metacognitive Strategy

Metacognitive strategy maintains the student's pronunciation learning process as a whole by emphasising learning sequences. For example, centring learning, arranging and planning learning, and evaluating learning (Rokoszewska, 2012; Oxford, 1990).

2.4.5   Affective Strategy

This strategy helps students control their emotions, attitudes, motivation, and values in the process of learning pronunciation. For example, to reduce anxiety, the individuals can listen to music, laugh, and meditate after learning pronunciation. Besides, they can increase self-confidence by making positive statements. Then, to regulate emotional temperature, students can discuss with colleagues when they have problems, try listening to the body's voice when they are tired of learning pronunciation (Rokoszewska, 2012; Oxford, 1990).

2.4.6   Social Strategy

Social strategy helps students combine their pronunciation learning abilities and social skills by engaging directly in social activities. For example, asking questions, cooperating or empathising with others (Rokoszewska, 2012; Oxford, 1990).

Thus, by utilising these strategies, students can systematically address and overcome their pronunciation difficulties, leading to clearer and more accurate speech in the target language.

2.5   Difficulties and Strategies in Pronunciation Learning

English pronunciation can be one of the challenging aspects of learning a language, particularly for EFL students for whom English is not their first language. As Ababneh (2018) stated that most language students' main difficulty is learning how to pronounce, and they consider pronunciation as the main communication problems. Several factors formulated by Kenworthy (1987) may be the reason that students experience difficulties in pronouncing words or mispronunciation. This refers to using language in a way that is regarded as improper, inaccurate, or unusual (Silalahi, 2016) and it is often experienced by EFL students in consonants, vowels, and diphthongs.

Students typically struggle to pronounce consonants correctly (Jahara & Abdelrady, 2021; Almuslimi, 2020; Adeline, 2020; Falahuddin et al., 2019; Purba, 2019; Komariah, 2018). For example, the difficulty in pronouncing consonant /θ/, /ð/, and /ʧ/ that sounds like /s/, /d/, and /c/. In other words, they shift what should be pronounced to the sound that only they can produce. Apart from that, problems with vowel sounds are common for EFL students. Students tend to pronounce words as they are written. As in the word bat and but, students pronounced bat and but instead of /bæt/ and /bʌt/ (Antaris & Omolu, 2019). They admitted that they had difficulty pronouncing vowel sounds because English phonetic symbols were tough to recognise.

In addition, in the context of a Chinese learner, the difficulty arises in distinguishing long and short vowels Utami (2020) because Chinese has a

different vowel system from English. Chinese does not distinguish /i-ɪ/, /u-ʊ/, /ɛ/-/æ/, or /ʌ/-/ɒ/ and vowel duration (long and short) does not function as a meaning as in the English phonological system (Zhang et al., 2015). Lastly, related to the students' difficulties in pronunciation of diphthongs. Students tend to have difficulty pronouncing diphthongs and shifting them to monophthongs (Donal, 2016). For example, the diphthong /ɪə/ is pronounced as /e/ or /i:/, and the diphthong /eɪ/ is pronounced as /i/, instead of the original pronunciation.

Essentially, students are able to overcome these difficulties, for example by utilising technology. The use of applications such as WeChat allows students to get pronunciation tasks in a convenient way and repeat it as many times (Dai & Wu, 2021). Similarly, the ELSA Speak app leverages Artificial Intelligence (AI) and speech recognition to help improve and perfect students' English pronunciation. In the bag of this technology, ELSA can detect students' pronunciation errors with an accuracy of more than 95%. Students can also receive follow-up feedback to correct pronunciation errors (Darsih & Asikin, 2020). In addition, Xiao and Park (2021) utilised ASR technology as a pronunciation learning tool. Students admitted that the technology was useful in terms of expanding practice opportunities, diagnosing pronunciation weaknesses, and providing speech models to native speakers.

In spite of technology, YouTube videos can be used as a resource for students self-regulated pronunciation practice (Al-Jarf, 2022; Zitouni et al., 2021; Darsih & Asikin, 2020). Operationally, YouTube videos are able to increase students' motivation through exposure to authentic English material, including increasing independent learning. In addition, Berk (2009) stated that YouTube videos have a significant influence on students' minds and senses because they can equip students with various learning styles and techniques, such as visual, musical, verbal and emotional intelligence. Furthermore, Hasanah and Suparwoto (2022) complemented that YouTube videos help students to learn more English content and learn to native speakers directly so that they know how to pronounce the words correctly and how the speakers speak, such as their accent and style.

Finally, the number shows the pronunciation difficulties experienced by EFL students in pronunciation learning, such as difficulties in consonant sounds and vowel sounds. Besides, there are also ways that students can overcome these difficulties through the use of various mobile applications.

2.6 Study of Relevant Research

In recent years, there have been plenty of studies investigating English pronunciation difficulties. Besides, the studies suggest a number of strategies in covering difficulties in learning pronunciation. Therefore, the writer intends to classify several studies related to pronunciation difficulties and strategies in learning pronunciation briefly as follows.

First, the research related to English pronunciation difficulties conducted by Motair and Mahmed (2022) in Yemen, was reported that Arabic EFL students had problems with one of the segmental features, namely consonant sounds, because the Arabic and English sound systems differ in many aspects; English contains 24 consonants, while Yemeni Arabic contains 28 consonants. Therefore, some English consonants are absent in the Yemeni Arabic system, such as /v/, /p/, /ŋ/. In addition, in collecting the data, the writer used interview techniques and found that students faced several problems with /r/, /dʒ /, /ð/, and /p/ vs /b/ and /f/ vs /v/. This problem occurs due to the influence of the mother tongue, students' lack of listening practice, and some Yemeni EFL students not using various types of audio and video aids to improve their speaking skills.

Likewise, the research conducted by Lestari et al. (2020) using qualitative descriptive methods and interview and document as data collection, was explained there were pronunciation errors experienced by students of the English Language Education Department, which included segmental features such as vowels, consonants and diphthongs. The most common errors found in vowels are /ɪ/, where students intend to pronounce it as /e/ sound. The most common errors found in consonants are /θ/, where students intend to pronounce it as a /t/ sound. The errors found in diphthongs are /a/ and /eɪ/, where students intend to pronounce it as /ɪ/ or /e/. These errors are caused by several factors, such as

students' internal factors, which include lack of preparation and practice in giving a speech as well as feeling anxious when giving a speech. Apart from that, the external factor includes students' lack of time and opportunities to use English in everyday life, while the interlinguas factor is interference from the mother tongue.

Second, research related to strategies in learning pronunciation was completed by Utami and Morganna (2022) in Indonesia. They use action research design with junior high school students to examine the effectiveness of shadowing techniques to assist students in resolving their English pronunciation issues. As Casillas (2020) stated that shadowing is a listening activity in which students trace the spoken utterances they listen to and repeat them as accurately as possible. In addition, they target the minimum student test score to be 70, and the results show that the average student score is 76, which is higher than the minimum. As a result, it is possible to conclude that shadowing is an effective technique in improving the students' English pronunciation competence.

Another finding is located in Hong Kong that utilised Automatic Speech Recognition (ASR) to facilitate English pronunciation assessment and learning in EFL students. As study conducted by Xiao and Park (2021) aims to investigate the effectiveness of ASR technology in diagnosing English pronunciation errors and to explore teachers' and students' attitude towards using ASR technology as a pronunciation assessment tool and as a learning tool. The findings demonstrate that there were overlaps between human rating and machine rating at the segmental level. However, teachers and students gave positive responses to the use of ASR in the pronunciation learning process. The teacher agreed that the use of ASR technology helped develop more intelligent English pronunciation and he indicated his willingness to use this technology in teaching pronunciation in the future. In addition, most students responded positively to the use of ASR because it was useful in terms of expanding practice opportunities, diagnosing pronunciation weaknesses, and providing speech models to native speakers.

Finally, Utami (2020) conducted the research in helping Chinese students learn pronunciation by using rubber bands. The rubber band serves as a visual representation of syllable length. As Gilbert (2008) stated that it is a visual

instruction in lengthening stressed syllables, as Chinese students have a common problem that they cannot shorten syllables so that they can differentiate the length of each syllable pronounced. To apply it, students can pull the rubber band wide between their thumbs when pronouncing a word in English or for more advanced students they can use it to practise dialogue or pronounce longer sentences.

By the whole of the previous studies, it can be concluded that recently there are students who still have English pronunciation difficulties even though they are English language students. However, they can overcome these difficulties by applying several strategies.