Chapter II
Literature Review

This chapter discusses some previous studies and theories which were relevant to the study. These theories are used as the basis in analyzing the problem of the research. They are including the theories of linguistics, phonology, pronunciation, formant of vowels, and PRAAT application. In addition, this chapter explains the object of the research.

2.1 Previous Studies

Regarding to this research, there are some previous studies that have done discussed about phonological errors, specifically in vowel pronunciation. The first study is about An Analysis of Students' Errors in Pronouncing English Vowel at the Seventh Year of SMPN 8 Palopo by Lalu Rudy Siswandi, a student of the English Study Program, 2014. This study has the purpose of analyzing the kinds of pronunciation English vowel errors and the causes of students' errors in pronunciation English vowels at SMPN 8 Palopo. This type of research is classified as descriptive research. The result of the analysis shows that students still need theory about pronunciation correctly because they make errors with pronunciation especially English vowel. There are five types of dominant errors. There are vowel /æ/:, vowel /ɒ:/, vowel /eə/, vowel /ʊə/ : student's errors, and vowel /i:/.

The second research is in an Indraprasta PGRI University journal by Widya and Erika Agustiana, students of the English Education Program, Faculty of Languages and Arts. This journal is entitled "English Vowels Pronunciation

Accuracy: An Acoustic Phonetics Study with PRAAT" and published in 2020. This research is aimed at measuring the accuracy of the pronunciation of English vowel sounds by third-semester students majoring in English education by comparing them to the standard pronunciation of English native speaker and identifying factors causing pronunciation problems. The researcher used the descriptive-qualitative method. The data were collected from native and non-native English speakers' pronunciation. The results showed that there are vowel sounds that are often mispronounced by participants. The most common mistake was when the participants had to distinguish long and short vowels ([i:]: [I] and [u:]: [ʊ]). The participants failed to distinguish long and short vowels, even though they had been informed that they were different.

The third research is a thesis by Firta Amalia, a student of English Education Program at IAIN Parepare and published in 2021. The title of the thesis is "Analyzing Student's Error Pronouncing Minimal Pair Words by Using PRAAT Application of English Program at IAIN Parepare. This research was aimed to described pronunciation error and correct pronunciation in English consonant and Vowel by the participants. The researcher chosen minimal pair words which are focused on fourth teen English consonant sounds: /0/ /ð/ /t/ /d/ /s/ /z/ /f/ /v/ /ʒ/ /k/ /g/ /?/ /tʃ/ /dʒ/, then twelve vowel sounds, there are: /əu//ei//eə//ə/ // /e/ /i://a:////æ/. The result showed that the most minimal pair word errors in consonant sound made by the students are /f/ and /v/ sounds, Meanwhile, the most minimal pair word errors in vowel sound made by the students are /ɔ:/ and /ə/ sounds. The causes of error in pronouncing minimal pair

words by students of English programs are an interlingual error or the error of students' native language reference and the intralingual transfer. Also, the lack of knowledge and absence of certain sounds in students' native language also found to be the cause of students' errors in pronouncing minimal pair words.

The fourth is a journal by Anisa Larassati, Nina Setyaningsih, Valentina Widya Suryaningtyas, and Setyo Prasiyanto Cahyono (2022), entitled “Using Praat for EFL English Pronunciation Class: Defining the Errors of Question Tags Intonation”. This journal is aimed to investigate students’ question tag pronunciation errors by using Praat software, and tries to get the students’ perspective on the use of Praat as a teaching and learning aid for English Pronunciation course. This research employed qualitative descriptive method. The results showed that errors commonly occur on the pronunciation of down tags, both in conversation and sentences (68% and 53%). And also, the results of the analysis clearly explicate that Praat can help the students evaluating their own pronunciation as it increases the accurate pronunciation of intonation for both the up tags and down tags. The students also gave positive review on the use of Praat in the English Pronunciation Class.

The last Research is a journal under the title “Using PRAAT for Analysing Segmental Features of Speech Produced by the Students of English Study Program of Universitas Brawijaya" Istiqomah Wulandari, Iis Nur Rodliyah, and Fatimah, the English Study Program of Universitas Brawijaya. This journal wants to investigate the similar problematic features of pronunciation of the second semester students of English Study Program of

Universitas Brawijaya and later will try to implement some teaching pronunciation techniques to modify their pronunciation production. For this particular research employs a qualitative research design. The data were analyzed into narrative description, interpretation and textual. The participants were asked to read aloud the given 14 minimal pair words list. The recorded data were saved as wave files (.wav) to analyze them by using PRAAT to visual display of the students' First Formant (F1) and Second Formant (F2). Then, these visual displays were compared to those of native speakers in PRAAT window. The result of this research reveals that due to the lack of how to disambiguate the correct pronunciation of vowel sounds within the given words the participant tends to produce inaccurate pronunciation. Which are reflected by the result of the spectrum score range: F1 and F2. Some proposed techniques of teaching segmental were drilling, teaching minimal pairs, chanting with jazz chant and rhyming with tongue twisters, to modify learner segmental features. As the suggestion all targeted parties (students, Lectures and Authority) should be willing to do what are necessary to maintain the success of the teaching and learning pronunciation in Study program of English.

2.2 Linguistic

According to Crane, Yeager, and Whitman (1981, p. 28), Linguistics is the study of language, most concerned with spoken language and its anthropological, psychological, and sociological ramifications. During some periods, linguists have been most interested in the changes in languages through history; during other periods, they have concentrated on the study of languages at just one time. Some

linguists have concentrated on describing how language is used; others, on how it should be used. Language has been seen by some linguists as a mirror of the mind and a key to the understanding of thought; to other linguists, mind is irrelevant to the study of language. Linguistics consists of two major fields, namely; (1) Micro-linguistics, namely the field of linguistics that studies language from within in other words studying the structure of the language itself; (2) Macro-linguistics, namely the field of linguistics that studies language in relation to factors outside of language, including interdisciplinary and applied fields (Benedet et al., 1998; Booij & Lieber, 1993) cited in Vocroix (2021, p. 3).

2.2.1 Micro-linguistic

Based on Skandera and Burleigh (2005, pp. 1-2), The system or structure of a language (langue or competence) can be described at four different levels, which form the core areas of linguistics, sometimes called micro-linguistics:

1) Phonetics and phonology deal with pronunciation, or, more precisely, with speech sounds and the sound system.
2) Morphology covers the structure of words.
3) Syntax explains sentence patterns. (Morphology and syntax, often combined into morphosyntax, have traditionally been referred to as grammar.)
4) Lexicology and semantics describe the vocabulary, or lexicon, and explore different aspects of meaning.

2.2.2 Macro-linguistic

According to Skandera and Burleigh (2005, pp. 2-3), utilizing the core areas are various other branches of linguistics, sometimes referred to as macro-linguistics. Most of these are interdisciplinary fields because they overlap with other sciences. The first four branches are concerned with language variation, and are therefore often subsumed under the label variational linguistics:

1) Dialectology is at the interface between linguistics and geography. It is the study of regional variation within a language.
2) Sociolinguistics connects linguistics with sociology. It is concerned with language variation according to age, sex, social class, etc.
3) Ethno-linguistics overlaps with anthropology and investigates language variation and the part language plays in ethnic groups.
4) Discourse analysis, text linguistics, and stylistics are related branches that also deal with language variation.
5) Contrastive linguistics describes the similarities and differences between two or more modern languages, especially in order to improve language teaching and translation.
6) Psycholinguistics overlaps with psychology and explores mental aspects of language, such as language learning.
7) Neurolinguistics overlaps with medical science and investigates the connection between language and the nervous system. It is especially

interested in the neurological processes necessary to produce speech sounds and in language disorders [*Sprachstörungen*].

8) Computational linguistics [*Computerlinguistik*] overlaps with artificial intelligence. Some of its concerns are machine translation, automatic speech recognition, and speech simulation.
9) Applied linguistics, is to the field of foreign language teaching. This term is contrasted with general or theoretical linguistics, which denotes a more theoretical orientation, but is not usually considered a separate branch.

2.3 Phonology

The sounds of language spoken in everyday life certainly have different systems, syllable arrangements, and meanings. According to Crane et al. (1981, p. 72), phonology is the study of how speech sounds are organized and how they function in language. It uses the classifications of sounds derived. from phonetics to describe and analyze how sounds occur in speech. Phonology, on the other hand, according to McMahon (2002, p. 3) involves a reduction to the essential information, to what speakers and hearers think they are saying and hearing. Put simply, phonology is the study of the sound patterns of language, how speech sounds are grouped by speakers to affect communication. Phonology is concerned with sound structure in individual languages, specifically how sound distinctions are used to differentiate linguistic items and how the sound structure of the 'same' element varies as a consequence of the other sounds in the context.

According to Odden (2005, p. 2) phonetics and phonology both deal with language sound, they address different aspects of sound. Phonetics deals with "actual" physical sounds as they are manifested in human speech, and concentrates on acoustic waveforms, formant values, measurements of duration measured in milliseconds, of amplitude and frequency. Phonetics also deals with the physical principles underlying the production of sounds, namely vocal tract resonances, and the muscles and other articulatory structures used to produce those resonances. Phonology, on the other hand, is an abstract cognitive system dealing with rules in a mental grammar: principles of subconscious "thought" as they relate to language sound. According to Skandera and Burleigh (2005, p. 5), phonology can be divided into two branches: (1) segmental phonology and (2) suprasegmental phonology. The first is segmental phonology that based on the segmentation of language into individual speech sounds provided by phonetics. Unlike phonetics, however, segmental phonology is not interested in the production, the physical properties, or the perception of these sounds, but in the function and possible combinations of sounds within the sound system. And the second is suprasegmental phonology, also called prosody, that concerned with those features of pronunciation that cannot be segmented because they extend over more than one segment, or sound. Such features include stress, rhythm, and intonation (also called pitch contour or pitch movement. Thus, it can be concluded that phonology is a branch of language science (linguistics) that studies the sounds of language and the processes of

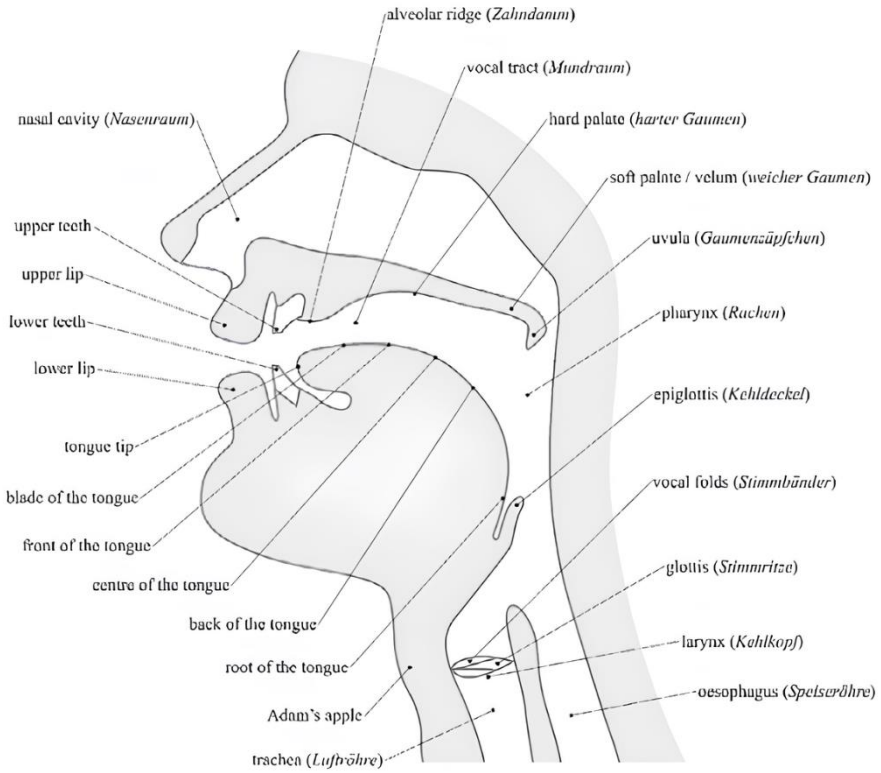
their formation and change. Phonology studies language sounds in general and functionally.

In general, the study of phonology is divided into two objects of study. The first object of study in phonology is language sounds (phonics), also known as the sound system (phonetics), and the second is phonemes, also known as the phoneme system (phonemics).

2.3.1 Phonetic

According to Brinton (2000, p. 10), Phonetics is the study of the speech sounds of human language in general, either from the perspective of their production (articulatory phonetics), their perception (auditory phonetics), or their physical properties (acoustic phonetics). On the other hand, based on Roach (2009, p. 44), phonetics is about how the sounds function of a language and its relationship to other sounds. It means the segment of utterances produced by the speaker has interrelationship to others, because a word appears due to the collection of sounds, and it forms a context or meaning from pronunciation. Moreover, its meaning becomes different when its sound is represented by a different speaker. Hence, it is all understood in the study of sounds. Phonetics is the study of how sounds are produced without paying attention to issues of meaning. In phonetics, we will discuss in detail the human speech organs, or human articulators, starting from the lips, tongue, teeth, oral cavity, nasal cavity, hard and soft palate, glottis, vocal cords, larynx, esophagus, windpipe, lungs, and diaphragm.

Figure 1. Speech Organ of Human



Source: Paul Skandera & Peter Burleigh (2005)

According to Skandera and Burleigh (2005, p. 3), phonetics first of all divides, or segments, concrete utterances into individual speech sounds. Phonetics can then be divided into three distinct phases: (a) articulatory phonetics, (b) acoustic phonetics, and (c) auditory phonetics.

a) Articulatory phonetics describes in detail how the speech organs, also known as vocal organs or articulators, in the vocal tract are used in order to produce, or articulate, speech sounds.

b) Acoustic phonetics is the study of the physical properties of speech sounds, or the way in which the air vibrates as sounds pass from speaker to listener. A spectrograph is a machine that measures the duration,

frequency, intensity, and quality of sound waves and displays them as visuals called spectrograms or sonograms.

c) Auditory phonetics studies the listener's perception of speech sounds, specifically how sounds are transmitted from the ear to the brain, and how they are processed.

According to Becker and Bieswanger (2017, p. 1), Phonetics provides objective ways of describing and analyzing the range of sounds humans use in their languages. More specifically, articulatory phonetics identifies precisely which speech organs and muscles are involved in producing the different sounds of the world's languages. Those sounds are then transmitted from the speaker to the hearer, and acoustic and auditory phonetics focus on the physics of speech as it travels through the air in the form of sound waves, and the effect those waves have on a hearer's ears and brain. It follows that phonetics has strong associations with anatomy, physiology, physics and neurology.

Thus, it can be concluded that phonetics is a branch of phonology that studies and describes language sounds from the point of view of speech, as well as how to form them so that they become air vibrations that can be received by hearing. Phonetics will be examined only from an articulatory perspective, i.e. the study of how the speech organs work to produce speech sounds and how these sounds are classified.

Figure 2. The International Phonetic Alphabet (IPA)

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Palatal	Velar	Glotal
(Oral) Stop	p b			t d			k g	ʔ
(Nasal) Stop	m			n			ŋ	
Tap or Flap				ɾ				
Affricate					tʃ dʒ			
Fricative		f v	θ ð	s z	ʃ ʒ			h
Approximant	w			ɹ		j	w	
Lateral approximant				l				

		Front	Central	Back
High	Tense	i		u
	Lax	ɪ		ʊ
Mid	Tense	e		o
	Lax	ɛ	ə ʌ	(ɔ)
Low	Lax	æ	a	

Source: Cambridge Dictionary Website

2.3.2 Phonemic

A phoneme is a distinctive or contrastive sound in a language which is the sound makes a difference in meaning and has communicative value Brinton (2000, p. 47). Phonemics is a branch of phonology that studies the sounds of language by paying attention to whether these sounds function as differentiators of meaning. Phonemic is the science that studies how sounds can differentiate meaning. A small number of regularly used sounds (vowels and consonants) that we call phonemes; for example, the vowels in the words ‘pin’ and ‘pen’ are different phonemes, and so are the consonants at the beginning of the words ‘pet’ and ‘bet’ Roach (2009, p. 2).

Every word or sentence spoken by humans is essentially a sequence of linguistic sounds. Changing a sound in the sequence can change the meaning. The change in meaning in question can change meaning or lose meaning. A phoneme defined as the smallest distinctive, or contrastive, unit in the sound system of a language. A phoneme, in other words, contrasts meaningfully with other speech sounds. Slash marks are conventionally used to indicate a phoneme, /p/, an abstract segment, as opposed to the square brackets, as in [p], used for each phonetic or physically produced segment.

Table 2. 1. Symbols for Phonemes

ɪ as in 'pit'	eɪ as in 'way'	p as in 'pea'	b as in 'bee'
e as in 'pet'	aɪ as in 'buy'	t as in 'toe'	d as in 'doe'
æ as in 'pat'	ɔɪ as in 'boy'	k as in 'cap'	g as in 'gap'
ʌ as in 'putt'	əʊ as in 'go'	f as in 'fat'	v as in 'vat'
ɒ as in 'pot'	aʊ as in 'cow'	θ as in 'thing'	ð as in 'this'
ʊ as in 'put'	ɪə as in 'peer'	s as in 'sip'	z as in 'zip'
ə as in 'about', upper'	eə as in 'pear'	ʃ as in 'ship'	l as in 'led'
i: as in 'key'	ʊə as in 'poor'	h as in 'hat'	r as in 'red'
ɑ: as in 'car'		m as in 'map'	j as in 'yet'
ɔ: as in 'core'		n as in 'nap'	w as in 'wet'
u: as in 'coo'		ŋ as in 'hang'	dʒ as in 'gin'
ɜ: as in 'cur'		ʧ as in 'chin'	

2.4 Pronunciation

Pronunciation is the understanding of how words, chunks, and blocks of English should sound. Pronunciation refers to the way we make the sound of a word. In other terms, pronunciation is the way a word or language is spoken or pronounced. Pronunciation (Cook, 1996) which was cited in Marulanda and Restrepo (2018, p. 10) defined as the practice and production of the variety of sounds in oral language. In other hand, Kelly (2000, p. 11) states that pronunciation is one of the important things in learning English in order to make a good communication. To make a good communication needs to pronounce the words correctly. The pronunciation is when we use all the same organs of speech to produce the sounds in particular away. Furthermore, as explained earlier, that the first thing that native speakers notice during a conversation is pronunciation. Grammar and vocabulary are important elements of language and they can be useless if the speakers cannot pronounce those elements or words accurately. Native speakers can understand people, despite their grammatical errors, if they use accurate pronunciation. Communicative efficiency can be guaranteed by correct pronunciation. Pronunciation is an essential part of communication and without correct pronunciation nobody can say that he/she knows the English language perfectly Harmer (2001, p. 249).

If learners want to change the way of pronouncing English words, they have to change the way they think about the sounds of those words. This is true both for individual sounds and the bigger parts of speech such as syllables, stress patterns, and rhythm Gilakjani (2012, p. 1). Richard cited in Heydari and

Bagheri (2012, p. 1584) that based on different types of errors, he distinguished three sources of errors: Interference errors, Intralingual errors, and Developmental errors. Interference errors are resulting from the use of elements from one language while speaking/writing another. Intralingual errors reflect general characteristics of the rule learning such as faulty generalization, incomplete application of rules and failure to learn conditions under which rules apply, and Developmental errors occur when learners attempt to build up hypotheses about the target language on the basis of limited experiences.

Apart from that, pronunciation includes articulation, emphasis, and intonation. Pronunciation relates to pronunciation techniques. Having good and correct pronunciation is the dream of many people. If you don't have good pronunciation, other people will have difficulty understanding what the speaker is saying. Good pronunciation is one of the aspects that will be assessed when speaking because pronunciation is the first impression that the interlocutor catches when speaking. If Indonesian only has two vowel sound systems: a, i, u, e, o, whereas in English you will pronounce many types of vowel letters such as æ, e, ə, ɔ, ɑ, ɒ, ɜ, i, ʊ, u, ʌ, ɪ. Of course, these vowel sounds are not easy to pronounce for those who are not used to them. Therefore, we must learn and practice pronunciation correctly. There are two elements of pronunciation, namely segmental features (included phoneme) and suprasegmental features.

1. Segmental Features of Pronunciation

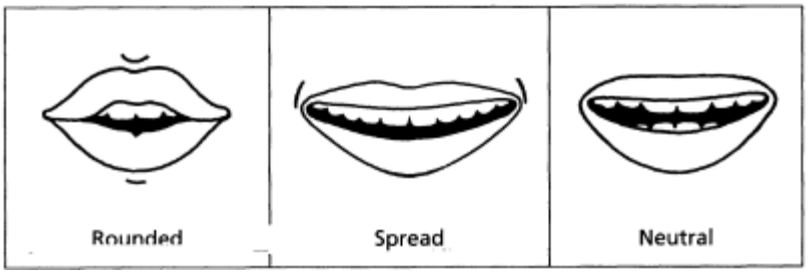
   Segmental features are also called phonemes. According to Becker and Bieswanger (2017, p. 59), segmental as individual sounds function in a

language which is called segments. In this segment, we can describe reasonably accurately how each sound is produced. Segments consist of two categories: vowels and consonants.

a. Vowel

According to Kelly (2000, pp. 29-46), Vowel sounds are all voiced, and may be single (like /e/, as in let), or a combination, involving a movement from one vowel sound to another (like /eɪ /, as in late); such combinations are known as diphthongs. An additional term used is triphthongs which describes the combination of three vowel sounds (like aʊ + ə = auə, as in hour or power). Single vowel sounds may be short (like /ɪ/, as in bit) or long (like /i:/, as in heat). The symbol /:/ denotes a long sound. The illustration below shows the basic lip positions which are used in describing the articulation of vowel sounds:

Figure 3. Basic Lip Positions



Source: Kelly (2000, p. 30)

- Rounded: the lips are pushed forward into the shape of a circle.
- Spread: the corners of the lips are moved away from each other, as when as smiling.
- Neutral: the lips are not noticeably or spread.

Vowels are articulated when a voiced airstream is shaped using the overall shape of the mouth with the tongue and the lips to modify. Vowels are classified based on whether the high or low the tongue is, if the tongue is in the front or back of the mouth, and whether or not the lips are rounded. All vowels can be divided into two main categories: diphthongs and monophthongs.

a) Diphthongs

Diphthong—two sounds. If the sound of a vowel continually changes within a single syllable, then we call it a diphthong. Complex vowels, which are characterized by movement, are called diphthongs. The total number of diphthongs is eight Crane (1981, p. 26).

Table 2. 2 Diphthongs

Diphthongs	Characteristics
eɪ as in 'way'	The glide begins in the positions for /e/, moving up and slightly back towards /ɪ/. The lips are spread.
aɪ as in 'buy'	The glide begins in an open position, between front and centre. The lips move from neutral, to loosely spread.

ɔɪ as in 'boy'	The glide begins in the position for /ɔ:/, moving up and forward towards / ɪ/. The lips start open and rounded, and change to neutral.
əʊ as in 'go'	The glide begins in the positions for /ə/, moving up and back towards /ʊ/. The lips are neutral, but change to loosely rounded.
aʊ as in 'cow'	The glide begins in a position a quite similar to / ɑ:/, moving up towards /ʊ/. The lips are neutral, but change to loosely rounded.
ɪə as in 'peer'	The glide begins in the position for /ɪ/, moving down and back towards /ə/. The lips are neutral, but with a small movement from spread to open.
eə as in 'pear'	The glide begins in the position for /e/, moving back towards /ə/. The lips remain neutrally open.
ʊə as in 'poor'	The glide begins in the position for /ʊ/, moving forwards and

	down towards /ə/. The lips are loosely rounded, becoming neutrally speed.

Source: Kelly (2000, pp. 35-36)

b) Monophthongs

Monophthong is one sound. If the sound of a vowel remains relatively unchanged during its production, then we call it monophthong. In simple vowels, or monophthongs, the tongue body has a relatively stable position throughout. But there are other vowels where the tongue body does not stay in one place, even in the most abstract diagrams with artificial slices Crane (1981, p. 24). A monophthong is a pure vowel sound that has a single auditory occurrence. There are twelve monophthongs in English. Simple vowels or monophthongs are divided into two kinds, there are short vowels and long vowels. Each vowel in short vowel is described in relation to the cardinal vowels. And there are five long vowels in English. The symbol consists of one vowel symbols plus a length mark made of two dots, [:].

Table 2. 3 Monophthongs (Short Vowels)

Short vowels	Characteristics
ɪ as in 'pit'.	The part of the tongue slightly nearer the center is raised to just above the half-close position. The

	lips are spread loosely, and the tongue is more relaxed
e as in 'pet'	The front of the tongue is between the half-open and half-close positions.
æ as in 'pat'	The front of the tongue is raised to just below the half-open position.
ʌ as in 'putt'	The center of the tongue is raised to just above the fully open position. Lips are neutrally open.
ɒ as in 'pot'	The back of the tongue is in the fully open position. Lips are lightly rounded.
ʊ as in 'put'	The part of the tongue just behind the center is raised, just above the half-loose position. The lips are rounded, but loosely so.
ə as in 'about', upper'	The center of the tongue is between the half-close and half-open positions.

Table 2. 4 Monophthongs (Long Vowels)

Long vowels	Characteristics
iː as in 'key'	The front of the tongue is slightly behind and below the close front position.
ɑː as in 'car'	The tongue, between the center and the back, is in the fully open position.
ɔː as in 'core'	The back of the tongue is raised to between the half-open and half-close positions.
uː as in 'coo'	The back of the tongue is raised just below the close positions.
ɜː as in 'cur'	The center of the tongue is between the half-close and half-open positions.

Source: Kelly (2000, pp. 31-33)

b. Consonant

Consonant sound may be voiced or unvoiced. It is possible to identify many pairs of consonants which are essentially the same except for the element of voicing (for example /f/, as in fan, and /v/, as in van). Consonants are speech sounds produced with a narrowing of the vocal

tract which is sufficient to prevent them from functioning as syllable nuclei (the nucleus is the 'heart' of the syllable, carrying stress, loudness, pitch information and usually consisting a vowel). In short, consonants are sound produced with a constriction or occlusion in the oral cavity Crane (1981, p. 57).

According to Crane (1981, pp. 59-62), in English, consonants can be divided into seven groups of the articulatory process based on where in the vocal tract obstruction of the air flow occurs, namely: bilabial, labiodental, dental/interdental, alveolar, palatal, velar and glottal/pharyngeal.

- Bilabial: sounds are made with both lips. There are five such sounds possible in English: [p] pat, [b] bat, [m] mat, [w] with, and [wh] where (present only in some dialects). We could say that the lower lip is the active articulator and the upper lip the passive articulator, though the upper lip usually moves too, at least a little.
- Labiodental: consonants are made with the lower lip against the upper front teeth. English has two labiodentals [f] as in fat and [v] as in vat. The lower lip is the active articulator and the upper teeth are the passive articulator.
- Dental/Interdental: are made with the tip of the tongue between front teeth. There are two interdental sounds in English: [θ] thigh and [ð] thy.

- Alveolar: behind your upper front teeth there is a small ridge called the alveolar ridge. English makes seven sounds at or near this ridge: [t] tab, [d] dab, [s] sip, [z] zip, [n] noose, [l] loose and [r] red.
- Palatal: sounds made near the hard part of the roof of the mouth are said to be palatal.
- Velar: the soft part of the roof of the mouth behind the hard palate is called the vellum. Sounds made near the velum are said to be velar. There are three velar sounds in English: [k] kill, [g] gill, and [ŋ] sing.
- Glottal/Pharyngeal: The space between the vocal folds is the glottis. English has two sounds made at the glottis. The first is easy to hear: [h] as in high and history. The second is called a glottal stop.

Table 2. 5 The Words of Consonants Pronunciation

p as in 'pea'	b as in 'bee'
t as in 'toe'	d as in 'doe'
k as in 'cap'	g as in 'gap'
f as in 'fat'	v as in 'vat'
θ as in 'thing'	ð as in 'this'
s as in 'sip'	z as in 'zip'
ʃ as in 'ship'	l as in 'led'

h as in 'hat'	r as in 'red'
m as in 'map'	j as in 'yet'
n as in 'nap'	w as in 'wet'
ŋ as in 'hang'	dʒ as in 'gin'
ʧ as in 'chin'	

2. Suprasegmental Features of Pronunciation

Suprasegmental is something that accompanies the phoneme which can be in the form of sound pressure (intonation), long-short (pitch), and sound vibrations that show certain emotions. Suprasegmental features are speech characteristics that follow consonants and vowels but are not limited to single sounds and often extend over syllables, words, or phrases. As Kelly (2000, p. 3) explained that suprasegmental features are features of speech which generally apply to groups of segments or phonemes. The features which are important in English are stress, intonation, and how sounds change in connected speech.

With regard to utterances, we can analyze and teach intonation as well as stress, although as features, they can at times be quite hard to consciously recognize and describe. Stress gives rhythm to speech. One or more words within each utterance are selected by the speaker as worthy of stress and thus made prominent to the listener. Each stressed syllable in a word in isolation also has a change in the pitch, or the level of the speaker's

voice, and the vowel sound in that syllable is lengthened. Stress can fall on the first, middle, or last syllables of words. Intonation, on the other hand, is the way in which the pitch of the voice goes up and down in the course of an utterance. (When discussing speech, the term utterance is used rather than 'sentence', as it refers to anything we say, including grammatically incomplete sentences and different ways of saying the same sentence). The term intonation refers to the way the voice goes up and down in pitch when we are speaking. It is a fundamental part of the way we express our own thoughts, and it enables us to understand those of others. It is an aspect of language that we are very sensitive to, but mostly at an unconscious level.

As well as helping to determine meaning, intonation gives us clues about the attitude of the speaker, or how he feels about what he is saying. When listening to people speak, we get clear messages about their attitude from the way things are said. We can get a good idea, for example, as to whether someone is interested, bored, kind, honest, or lying, and so on. Utterance stress and intonation patterns are often linked to the communication of meaning. For example, in the following utterance, the speaker is asking a question for the first time. In this particular instance, as you can hear on the CD, the pitch of her voice starts relatively high and falls at the end, finishing relatively low.

2.5 Formant Frequencies for English Vowels

A formant is a concentration of acoustic energy around a particular frequency in the speech wave. There are several formants, each at a different

frequency, roughly one in each 1,000Hz band. Each formant corresponds to a resonance in the vocal tract. Phonetic scientists like to describe vowels in terms of numbers. It is possible to analyze sounds so that we can measure the actual frequencies of the formant. According to Ladefoged and Johnson (2006, p. 307), each vowel has three formants, i.e. three overtone pitches. The first formant (F1) is inversely related to vowel height. The second formant is related to the degree of backness of a vowel. Formants can be seen in a wideband spectrogram as dark bands. The frequencies of the formants are determined by the shape and size of the mouth and pharynx (and in some cases also the nasal) cavities Hayward (2000, p. 124). Quality of vowel phonemes is determined by frequency components called formants. To identify the quality of a vowel, two lowest formants are used. According to Ladefoged and Johnson (2006, p. 23), the one with the lower pitch (distinguishable in creaky voice) being called the first formant (F1) and the higher one (the one heard when whispering) the second formant (F2). The values of the frequencies of F1 and F2 are sufficient to distinguish most vowel contrasts in most languages.

The most important feature that characterizes a specific vowel is the formants, which are the resonant frequencies of the vocal tract. During the vowel articulation, the shape of the vocal tract remains relatively in constant shape so the formants do not change abruptly during a single vowel. We used this feature as the measure of vowel pronunciation accuracy. Vowels can be distinguished with sufficient accuracy by the first three formants [1]. But the first two of them play the more important role than the third. It is well known

that there is tight relation between a vowel and its F1 and F2 and that these are also closely tied to the shape of the vocal-tract articulators.

Table 2. 6 Average Formant Frequencies of F1 and F2 for American English Vowels for Men

Vowels	F1 (Hz)	F2 (Hz)
/i/	324	2322
/I/	427	2034
/e/	476	2089
/ɛ/	580	1799
/æ/	588	1952
/ɑ/	768	1333
/ɔ/	652	997
/ɒ/	497	910
/ʊ/	469	1122
/u/	378	997
/ʌ/	623	1200
/ɜ/	474	1379

Source: (Scherer, Lucas, Gratch, Rizzo, & Morency, 2016)

2.6 PRAAT Application

PRAAT is an application or computer software for speech analysis. It was designed, and continues to be developed by Paul Boersma and David Weenink from the Institute of Phonetic Sciences, University of Amsterdam. Wilson (2008) cited in Osatanandaa and WThinchanb (2021, p. 33) posited that

PRAAT can be used for more than simply plotting formants and proposed other applications based on measurements of duration, pitch, and intensity, in particular the distinctions of English. When a speech is recorded and entered into the PRAAT program, sound waves can detect the speech's features. Empowering the PRAAT software can provide accurate and full acoustic information.

Boersma & Weenink, 2013 cited in Jolayemi and Oyinloye (2019, p. 20), stated that the main highlight of PRAAT is that it is multi-task software, which is very handy for users of many categories such as: the phoneticians, phonologists, speech pathologists, voice coaches, and forensic experts. It can be excited by only sound files in the following formats: AIFC files, AIFF files, FLAC files, MP3 files, NeXT/Sun (.au) files, NIST files, and WAV files. The PRAAT software can be a great tool to help non-native speakers to improve their pronunciations by practicing supra segmental features Aramipoor and Gorjian (2018, p. 48). In this case, PRAAT software could be used to measure the formant of English vowels by the speaker.

According to Widayanti (2022, p. 77), with the PRAAT application it can be identified the value of pitch, formant and spectrogram of each sound recording. PRAAT providing spectral analysis (spectrograms), pitch analysis, formant analysis, intensity analysis, jitter, shimmer, voice breaks, cochlea gram, and excitation pattern. And also we can generate waveforms/oscillogram, wide and narrow band spectrograms, intensity contour and pitch tracks; make recordings, edit a recorded sound, and extract individual sounds for further

analysis; get information about pitch, intensity, formants, pulses and etc.; enhance certain frequency regions; segment and label words, syllables, or individual phonemes; put your work in graphic form ('draw a plot') for printing (EdUHK, 2024).

Figure 4. An example of PRAAT Fitur

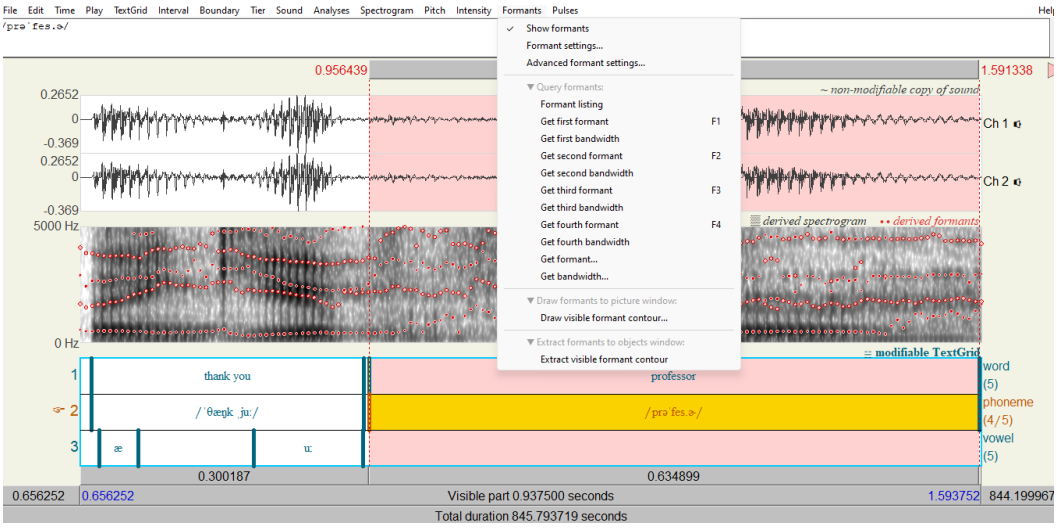
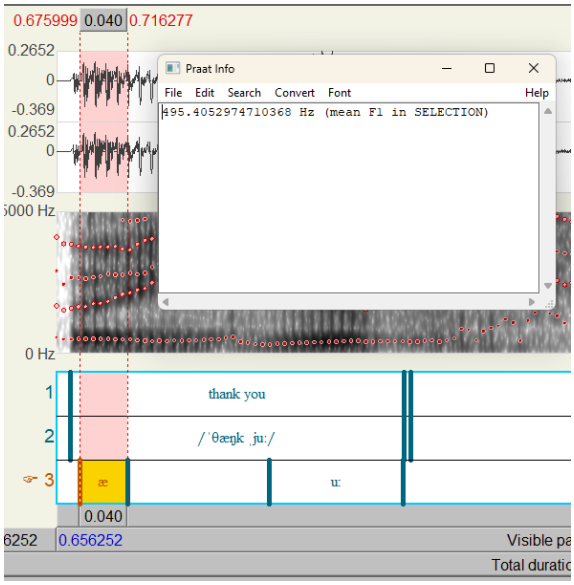


Figure 4 display the oscillogram on top, and the spectrogram in the bottom included the formant (the red line). An oscillogram is a visual representation of sound data that shows the relationship between time and amplitude. The vertical axis on an oscillogram represents amplitude while the horizontal axis shows seconds of time. Each point along the waveform represents relative amplitude in decibels – meaning, the tool allows you to see and measure how amplitude changes over time from loud to soft. Meanwhile, a spectrogram provides a view of amplitude and frequency over time (Wildife, 2022). Formants are the characteristic amplitude peaks in the spectrum of resonant sound sources. They result from the excitation of fixed resonant

chambers and are the most significant contributors to the timbre of tonal instruments. In speech, they are present below 5000 Hz and are usually "in-harmonic," meaning their frequencies are not integer multiples.

Figure 5. An Example formant of Vowel in PRAAT



Based on the figures above, the spectrogram displays the input data in oscillogram, the measurements made on the vowels involve the formant and as well as the duration of the vowel being uttered, in this case, the vowels measurement focuses more on the of the vowels uttered by the speaker (Widayanti, 2022).

2.7 The Relevance of YouTube, Anies Baswedan, and Pronunciation

In this modern era, many people have and use social media. With social media, people can easily get information. Carr and Hayes (2015, p. 8) defined social media as “Internet-based channels that allow users to opportunistically interact and selectively self-present, either in real-time or asynchronously, with both broad and narrow audiences who derive value from user-generated content and the perception of interaction with others". Individuals use social media to

express themselves, discuss their interests, connect with friends, and grow their careers. Different social media platforms are used for specific purposes.

With the rapid advancement of technology, there are numerous sorts of social media platforms that are now often used by people. Starting with YouTube, Twitter, WhatsApp, Facebook, Instagram, and many more. By using YouTube, people can get many impacts, where they can watch, like, share, comment, and upload their own videos. Such as a video of Anies Baswedan on a YouTube channel. Anies Baswedan is an Indonesian activist and politician who was Jakarta's governor from 2017 to 2022. Like explained earlier, he became a speaker about the experience of megacity Jakarta on the NISTH YouTube channel. NISTH is an abbreviation of the NTU (Nanyang Technological University) Institute of Science and Technology for Humanity. The video was published on March 2, 2023. We can watch the video on the NISTH YouTube channel for free and gain knowledge from what the speakers said. We can also listen to Anies Baswedan's speech in English and learn English from the video.

To deliver something or organized and accurate information, people need good language skills so that people can communicate effectively with others. Communicative efficiency can be guaranteed by correct pronunciation. So, Pronouncing the words correctly is one of the most essential things to do. Pronunciation is a little thing we need to pay attention to in order to avoid the habit of repeating the same mistakes because it probably causes misunderstandings and misinterpretations. When someone is able to say words

clearly and accurately, the message you want to deliver can be well received by the listener.

So, as a politician, is expected to have good verbal and non-verbal communication skills. And in this speech, Anies Baswedan spoke in front of many students and lecturers at NTU with the theme of megacity Jakarta. For that, Anies must be good at communicating everything clearly and be careful so that there is no misunderstanding. For this reason, correct and accurate pronunciation will be very helpful in delivering information. The writer would like to study more deeply about phonology and the accuracy of pronunciation.