

**A COMPARATIVE ANALYSIS OF STRESS AND INTONATION
BETWEEN INDONESIAN EFL PRE-SERVICE TEACHERS AND
NATIVE ENGLISH SPEAKERS**

A Thesis

**Presented to the Department of English Language Education as Partial
Fulfillment of Requirement to Obtain the Bachelor of Education Degree in
English Language Education**



By:

Furqan Arrahman

18322140

**ENGLISH LANGUAGE EDUCATION DEPARTMENT FACULTY OF
PSYCHOLOGY AND SOCIAL SCIENCES ISLAMIC UNIVERSITY OF
INDONESIA**

2025

APPROVAL SHEET

**A COMPARATIVE ANALYSIS OF STRESS AND INTONATION BETWEEN
INDONESIAN EFL PRE-SERVICE TEACHERS AND NATIVE ENGLISH SPEAKERS**

By

Furqan Arrahman

18322140



Approved on 18th July 2025

By

Supervisor

A handwritten signature in black ink, appearing to be 'Irma Windy Astuti', written over a horizontal line.

Irma Windy Astuti, S.S., M.Hum.

NIP. 0622160005

RATIFICATION SHEET

**A COMPARATIVE ANALYSIS OF STRESS AND INTONATION BETWEEN
INDONESIAN EFL PRE-SERVICE TEACHERS AND NATIVE ENGLISH
SPEAKERS**

By

Furqan Arrahman
18322140

Defended before the Board of Examiners on 21th July, 2025 and Declared

Acceptable.

Board of Examiners

Chairperson : Irma Windy Astuti, S.S., M.Hum.

First Examiner : Dr. Rizki Farani, S.Pd., M.Pd.

Second examiner : Willy Prasetya, S.Pd., MA.

Department of English Language Education

Faculty of Social and Cultural Sciences

Islamic University of Indonesia

Head of Department,



Wahyuni, S.Pd., M.L.S.T., Ph.D.

NIP.053310402

STATEMENT OF WORK ORIGINALLY

I honestly confirm that this thesis was originally written by me.

This thesis does not contain any work from others except those cited in quotation and references, as a scientific paper should.

Yogyakarta, 23th July 2025

Author,



Furqan Arrahman

18322140

MOTTO

“Pain is inevitable, suffering is optional.”

(Haruki Murakami)

DEDICATION

I wholeheartedly dedicate this thesis to, first and foremost, my parents; to my mother who never has any doubts I can achieve greatness, and to my father for being such an exemplar of perseverance and self-sacrifice in the face of adversities. To my siblings and relatives whose compassion I will never forget. Lastly, I dedicate this work to past and future self; let this be of assurance of your merit in times of uncertainties.

ACKNOWLEDGEMENT

All praise and gratitude belong to Allah, the Lord of all Worlds, for his blessings, guidance, good health, and peace of mind, which have enabled me and pushed me to complete this thesis. Peace and blessings be upon our beloved Prophet Muhammad. Humbly, I would like to extend my heartfelt appreciation, respect, and profound gratitude to:

1. My parents and siblings for their endless encouragement and support. Words fail to express just how much I depended on you to keep me tethered to the world beyond academia.
2. To my thesis supervisor, Ms. Irma Windy Astuti S.S., M. Hum, for your patient feedback and intellectual clarity, along with your understanding.
3. To my academic supervisor Ms. Banatul Murtafi'ah S.Pd., M.Pd.; your guidance never imposed, but always illuminated.
4. To the Department of English Education and staff, I am grateful for the academic shelter that allowed my inquiry to grow despite my endless shortcomings.
5. To the participants who lent their voices to this study. Your contributions shall not be mere data points to be measured. They are but a testament to your effort and dignity in learning.
6. To the friends, present, distant, and those scattered across time zones and time-thinned bonds, thank you for your echoes of laughter, and for leaving behind traces that quietly stayed.

TABLE OF CONTENT

COVER	i
APPROVAL SHEET	ii
RATIFICATION SHEET	iii
STATEMENT OF ORIGINALITY	iv
MOTTO	0
DEDICATION	1
ACKNOWLEDGEMENT	2
TABLE OF CONTENT	3
LIST OF TABLES	5
LIST OF FIGURES	6
LIST OF APPENDICES	7
ABSTRACT	8
GLOSSARY	9
CHAPTER I	11
INTRODUCTION	11
1.1 Background of the Study.....	11
1.2 Identification of the Problems.....	12
1.3 Formulation of the Problem.....	14
1.4 Objectives of the Study.....	14
1.5 Significance of the Study.....	14
CHAPTER II	16
LITERATURE REVIEW	16
2.1 Pronunciation in English Language Learning.....	16
2.1.1 Segmental Features of Pronunciation.....	16
2.1.2 Suprasegmental Features of Pronunciation.....	17
2.2 Common Phonological Tendencies Among EFL Teacher Trainees.....	17
2.3 Ways of Evaluating: Stress and Intonation.....	18
2.3.1 Manual Pronunciation Assessment.....	18
2.3.2 Machine-Assisted Pronunciation Analysis.....	19
2.4 Theoretical Framework.....	21
2.1. Theoretical Framework.....	21
RESEARCH METHODOLOGY	22
3.1. Research Design.....	22
3.2. Data Preparation.....	23
3.2.1. Setting and Participants.....	23
3.2.2. Types and Sources of Data.....	23

3.2.3. Ethical Considerations.....	25
3.3. Data Collection Technique.....	25
3.4. Data Analysis Technique.....	26
3.5. Trustworthiness.....	27
CHAPTER IV.....	28
FINDINGS AND DISCUSSIONS.....	28
4.1. Findings.....	28
4.1.1. Word Stress Patterns.....	28
Table 4.1.....	29
Normalized Syllable Durations and Mean Intensity of “Disputing,” “Immediately,” and “Obliged” in EFL Group.....	29
Table 4.2.....	30
4.1.2. Sentence Stress Patterns.....	31
Table 4.3. Mean Intensity and Normalized Duration of Content Words by EFL and NS.....	32
4.1.3. Intonation Patterns.....	34
Figure 4.1. Normalized Pitch Contours Four EFL Speakers.....	35
Figure 4.2. Normalized pitch contours for four NS speakers.....	36
4.2. Discussion.....	37
4.2.1. Word Stress Realization.....	37
4.2.2. Sentence Stress Patterns.....	38
4.2.3. Intonation Contours.....	38
CHAPTER V.....	40
CONCLUSION.....	40
References.....	42
APPENDIX 1: Syllable Duration and Intensity Listings of EFL and NS Group.....	47
APPENDIX 2: Word Duration and Intensity Listings of EFL and NS Group.....	55

LIST OF TABLES

TABLE 4.1. Normalized Syllable Durations and Mean Intensity in EFL Group	29
TABLE 4.2. Normalized Syllable Durations and Mean Intensity in NS Group	30
TABLE 4.3. <i>Mean Intensity and Normalized Duration of Content Words by EFL and NS</i>	32

LIST OF FIGURES

FIGURE 2.1. Theoretical Framework	21
FIGURE 4.1. Normalized pitch contours for four EFL speakers	35
FIGURE 4.2. Normalized pitch contours for four NS speakers	36

LIST OF APPENDICES

APPENDIX 1 Syllable Duration and Intensity Listings of EFL and NS Group	47
APPENDIX 2 Word Duration and Intensity Listings of EFL and NS Group	55

**A COMPARATIVE ANALYSIS OF STRESS AND INTONATION
BETWEEN INDONESIAN EFL PRE-SERVICE TEACHERS AND
NATIVE ENGLISH SPEAKERS**

Furqan Arrahman

18322140

ABSTRACT

This study examines the stress and intonation patterns of Indonesian EFL pre-service teachers through acoustic analysis using PRAAT. Four participants read a standardized passage, *The North Wind and the Sun*, with their recordings compared to those of four native American English speakers from online corpora. The analysis focused on word stress, sentence stress, and intonation, using selected utterances and target words of pedagogical and prosodic relevance. Intensity, normalized duration, and semitone-based pitch were extracted, with normalization applied for gender and speech rate differences. Findings show that while EFL speakers marked stress on key content words, their prominence cues were less consistent and more limited in range than those of native speakers. Intonation contours were also flatter, suggesting influence from L1 prosody and limited instructional exposure. The results underscore the need for targeted suprasegmental training and highlight Praat's value in pronunciation research.

Keywords: acoustic analysis, EFL, intonation, Praat, sentence stress, suprasegmentals, teachers, word stress

GLOSSARY

Acoustic Analysis

The investigation of speech features such as pitch, intensity, and duration through measurable physical signals using software like Praat.

Annotation

The process of labeling parts of a speech recording (e.g., syllables, words) to facilitate analysis of timing and prosodic features.

Content Word

A word that carries significant meaning and contributes to the core message of a sentence (e.g., noun, verb, adjective, etc.)

Forced Alignment

An automated technique for aligning a speech recording with its orthographic transcript at the word and phone level, used to generate time-aligned annotations (McAuliffe et al., 2017).

F0 (Fundamental Frequency)

The basic frequency of vocal fold vibration, perceived as pitch, and a primary parameter for intonation analysis (Johnson, 2012).

Intensity (dB)

The loudness of a sound signal measured in decibels, used as an acoustic correlate of syllabic and sentence stress (Sluijter & van Heuven, 1996).

Intonation

The pitch variation across an utterance that signals sentence modality, discourse structure, or speaker intention (Chun, 2002).

(Lexical) Stress

The syllable in a word that is pronounced with greater prominence through increased pitch, intensity, or duration (Field, 2005).

Normalization

Adjustment of acoustic data (e.g., pitch, duration, intensity) to a common scale so that comparisons across different speakers are valid despite physiological or stylistic differences.

Normalized Duration

A proportion-based duration measure expressed relative to a word or utterance's

total length, allowing comparison across speakers with different speech rates (Nolan, 2003; Xu, 1999).

NWS

Refers to The North Wind and the Sun passage.

Praat

A software tool for phonetic analysis that allows visualization and measurement of acoustic features including pitch, duration, and intensity (Boersma & Weenink, 2024).

Pitch

The perceptual correlation of F0 that reflects how high or low a voice sounds and is central to stress and intonation analyses (Nolan, 2003).

Polysyllabic

A word containing more than two syllables.

Semitones

A logarithmic unit that expresses pitch change relative to a speaker's median F0, enabling perceptually meaningful cross-speaker comparisons (Nolan, 2003; Johnson, 2012).

Sentence Stress

The pattern of prominence across words in a sentence, typically falling on content words to mark new or important information (Hahn, 2004; Cruttenden, 2006).

Suprasegmental Features

The prosodic elements of speech, such as stress, rhythm, and intonation, that operate over syllables, words, or entire utterances (Pennington & Richards, 1986).

TextGrid

A time-aligned annotation file format used in PRAAT for labeling phones (syllables) and words.

Wav (.wav)

A standard audio file format known for storing uncompressed and lossless audio data to retain higher audio quality.

CHAPTER I

INTRODUCTION

1.1 Background of the Study

In classroom contexts where English is taught and acquired as a foreign language, mutual intelligibility between interlocutors, namely teachers and learners, becomes an essential component of successful communication. This intelligibility depends not only on accurate articulation of segmental features, such as individual consonants and vowels, but also on the correct use of suprasegmental features, including word stress, sentence stress, and intonation. These prosodic elements are crucial in spoken language, as they function to convey emphasis, grammatical relationships, discourse structure, and speaker intent (Celce-Murcia, et al., 1996; Derwing & Munro, 2005).

For communication to be reciprocal and effective, both teachers and learners must develop a sufficient level of phonological competence. This includes the ability to perceive and reproduce sounds and patterns with clarity, fluency, and intelligibility (Saito & Lyster, 2012). However, existing studies show that non-native EFL teachers, particularly those in training, often face challenges in accurately modeling these features in their own speech. Bai and Yuan (2018), for instance, observed that Hong Kong teachers, while acknowledging the importance of pronunciation instruction, lacked confidence in their ability to teach it effectively. This often stemmed from limited training and weak proficiency in suprasegmental pronunciation.

Such issues are not limited to Hong Kong. In the Vietnamese and Malaysian contexts, similar trends have been reported. Teachers frequently resort to implicit feedback methods such as recasts, rather than explicitly teaching pronunciation patterns (Nguyen & Newton, 2020; Tran & Nguyen, 2020; Shah et al., 2017). These strategies may limit learners' awareness of their errors,

particularly when those errors involve features like sentence stress or rising and falling intonation.

This raises a pedagogical concern. If teachers struggle to model suprasegmental features effectively, learners may imitate incorrect patterns, potentially reinforcing fossilized errors over time. Jenkins (2000) emphasizes that teachers often serve as primary pronunciation models for their students. When teachers' speech significantly deviates from target norms, especially in areas such as prominence and intonation contour, it may compromise the clarity and effectiveness of the input learners receive.

Although the field of pronunciation teaching has gained greater attention in recent decades, many studies focus either on segmental issues or broad pedagogical beliefs. Research that systematically analyzes the actual suprasegmental output of EFL teachers remains limited. Moreover, few studies use objective acoustic tools to examine the speech of non-native instructors in training. Even where such research exists, it is largely situated in non-Indonesian contexts, such as Turkey (Arikan & Yilmaz, 2020; Canan & Oktay, 2021), where linguistic influences and training systems may differ significantly.

To address this gap, the present study focuses on pre-service EFL teachers in Indonesia and their production of stress and intonation, in contrast to read-aloud recordings by native speakers. Using Praat software, a phonetic analysis tool developed by Boersma and Weenink (2021), the study investigates stress and intonation patterns of both groups' reading of *The North Wind and the Sun*, as recommended by the International Phonetic Association (1999). This controlled reading task enables a focused and replicable analysis of the participants' spoken output, offering insight into their current level of proficiency and the typical prosodic patterns they output.

1.2 Identification of the Problems

While there is a growing body of research on the pronunciation challenges faced by EFL learners, there remains a noticeable lack of studies that specifically examine the suprasegmental aspects of pre-service teachers' speech using empirical tools. Many investigations tend to include prosody only as part of broader diagnostic categories, often relying on subjective teacher feedback or classroom observation. This makes it difficult to pinpoint specific patterns or quantify the extent of the problem.

Several external factors contribute to this gap, particularly in the Indonesian EFL context. One key issue is the limited inclusion of pronunciation, especially suprasegmental features such as intonation, word stress, and rhythm, in national curriculum frameworks and teacher education syllabi (Astawa et al., 2017; Dewi & Pratiwi, 2020). As pronunciation is often embedded implicitly within speaking skills without explicit instructional focus, many pre-service teachers complete their training with minimal exposure to the pedagogical and acoustic nuances of prosody. Additionally, a persistent exam-oriented teaching culture discourages the prioritization of features not directly assessed in standardized tests (Marcellino, 2008), leading to a lack of both awareness and motivation to develop suprasegmental proficiency.

Resource-related constraints also play a role. Limited access to training in technological tools like PRAAT or Montreal Forced Aligner, and a general scarcity of phonetics lab infrastructure in many teacher-training institutions, further impede efforts to introduce empirical analysis into pronunciation instruction (Darmawan, 2019). These structural challenges contribute to the continued reliance on auditory or impressionistic evaluation methods, which may be insufficient for addressing subtle prosodic deviations.

Additionally, the Indonesian context has received limited attention in terms of instrument-based analyses of EFL pronunciation. Although a few studies (e.g., Utami & Yulia, 2022) have touched on stress placement errors among

prospective teachers, they typically use impressionistic methods and do not offer acoustic evidence that could inform data-driven pedagogy.

1.3 Formulation of the Problem

Based on the above considerations, this study aims to answer the following research question:

- How does the (word and sentence) stress and intonation-related pronunciation patterns of EFL teachers compare to those of native speakers (NS), as analyzed through Praat?

1.4 Objectives of the Study

This study has the following objectives:

1. To compare the patterns of suprasegmental pronunciation of EFL teachers in comparison to native speakers (NS) specifically in word stress, sentence stress, and intonation,
2. To analyze and describe the acoustic properties of these features using Praat software.

1.5 Significance of the Study

The study is expected to provide insights into the pronunciation development of Indonesian EFL teacher trainees, with a specific focus on suprasegmental features. It offers practical relevance for curriculum designers, pronunciation trainers, and teacher educators who aim to enhance the communicative competence of future language instructors.

By employing Praat to document the acoustic features of stress and intonation in a controlled context, this research moves beyond anecdotal or perceptual evidence, contributing to a more rigorous understanding of interlanguage phonology. The findings may also support the integration of more targeted suprasegmental instruction within teacher education programs.

CHAPTER II

LITERATURE REVIEW

2.1 Pronunciation in English Language Learning

Brazil et al. (1980) conceptualized pronunciation as a component of a system through which referential meaning is expressed; an especially important component that allows for interactional dynamics of a communication process. Such a component is referred to by Pennington and Richards (1986) as the articulation of each phoneme using the stress and intonation patterns of the target language. They also identified three factors that are involved in the process of mastering pronunciation, namely perceptual, articulatory, and interactional, which were then categorized into three types of features: segmental, voice-setting, and suprasegmental. Such factors work together in that the degree to which a learner is able to integrate each factor into their speech determines whether or not their speech is considered intelligible by the listeners. Accordingly, to avoid miscommunication and irritation, teachers generally agree to incentivize learners to articulate “differentiable English phonemes” (Norrish, 1983; Harmer, 1991).

2.1.1 Segmental Features of Pronunciation

Segmental features are the sounds of an individual’s speech which can be compartmentalized into sounds, and those sounds can be divided into minimal units. These units are referred to in phonetic terminology as phonemes. Consequently, distinguishing between each phoneme within the confines of a target language’s phonetic inventory and the various forms in which they appear in some syllables or words is detrimental to the acquisition of that target language (Pennington & Richards, 1986). This is due to the proclivity of some phonemes to alter meaning depending on their use and or position (Demirezen, 1987).

2.1.2 Suprasegmental Features of Pronunciation

Pennington and Richards (1984) described suprasegmental or prosodic features as the levels to which stress and intonation differ relative to each syllable, word, phrase, and speech of longer stretches. Moreover, suprasegmental features are also associated with the phenomenon of coarticulatory; the compounding and intermingling of phonemes in a fluent speech. In other words, suprasegmental features are beyond mere articulation of phonemes and have more to do with other elements of speech that create meaning. For instance, one of the elements is stress; the amount of effort, or lack thereof, put into producing syllables (Pennington & Richards, 1984). This may involve the levels of pitch, intensity, or omission of syllables.

Brown and Yule (1983) regarded this tweaking of syllables as the way of a speaker to denote what the listener is required to attend to. Furthermore, Dirven and Oakeshott-Taylor (1984) made the case that syllables or words of high importance, that is, containing a high amount of information, are articulated in a precise manner, while the ones that are predictable and can be inferred from context are articulated in a weakened, shortened, and loose manner. This notion – words in an utterance carry different levels of “information units” – dates back to as early as late 1960s (Halliday, 1967).

2.2 Common Phonological Tendencies Among EFL Teacher Trainees

In the process of learning a target language, learners constantly evolve their understanding and hypothesis of the systems of the language, although there comes a point when the learning process stagnates, resulting in “fossilization” (Selinker, 1972). This phenomenon is generally referred to as interlanguage, although some researchers have more nuanced views of such a concept. For

instance, Corder (1981) used the term interlanguage to mean the ever-evolving language systems that learners use in the process of learning. Richards (1974), however, viewed interlanguage as a completely separate system that mediates L1 and L2. Smith (1994) specifically described it as the systematic behaviors of learners.

Learners who share a common L1 tend to develop similar interlanguage patterns and thus encounter similar pronunciation difficulties. This tendency has been documented in several EFL teaching contexts. Hunt-Gómez and Navarro-Pablo (2020) found that Spanish EFL teachers frequently substituted unfamiliar vowels like /ə/ and /ʌ/ with more familiar ones, and displayed a tendency to de-voice final segments. Likewise, Arikan and Yilmaz (2020) reported that Turkish EFL teachers struggled with sounds such as /θ/, /ð/, and /ŋ/, alongside issues like vowel insertion and shortening. In a related study, Kapranov (2019) observed that Norwegian EFL teacher trainees often mispronounced /ʌ/, /ð/, and /z/, all of which are phonemes absent in their L1.

Although these examples highlight segmental difficulties, they underscore a broader principle: learners' phonological systems are constrained by their native language and learning environment. This principle also applies to suprasegmental features, including stress and intonation, which have been shown to be affected by similar L1 influence (Jenkins, 2000; Mennen et al., 2012). Indonesian EFL pre-service teachers may thus exhibit prosodic patterns distinct from native norms, shaped by their shared linguistic background and pedagogical exposure. These phenomena, however, are not necessarily indicative of individual failure but reflect shared developmental trajectories.

2.3 Ways of Evaluating: Stress and Intonation

2.3.1 Manual Pronunciation Assessment

There have been some methods used in either predicting or evaluating learner's errors that are done with the involvement of human evaluators such as language instructors or trained assessors. These conventional, non-technological methods focus on analyzing and comparing the phonemes in L1 and L2, while others focus on the sounds learners produce. One such practice of comparing the mother language and the target language is mainly used in Lado's (1957) Contrastive Analysis Hypothesis. Supported by behaviorist views, CAH aims to predict errors that learners would make based on the phonological differences between the two languages, as they account for the bulk of the errors (Dulay et al., 1982).

Despite gaining prominence at the time, CAH was criticized for not taking into account learners' strategies of dealing with the differences between the languages. CAH was then divided into two forms: strong form or Lado's original CAH and weak form where explanation of learner errors takes place based on similarities and dissimilarities between the two languages (Wardaugh, 1970).

The other practice involves the identification and evaluation of learner errors of an interlanguage at a certain point in time by comparing them to the target language (Corder, 1973; Ellis, 1994). It is important, however, to distinguish between errors and mistakes, as errors are systematic and are the results of learner's ignorance regarding the nature of the target language rather than irregular and caused by factors such as tiredness or lapse of memory (Crystal, 1992). Unsystematic errors or mistakes were regarded by Corder (1974) as of insignificant use to learners or instructors. Corder's Error Analysis was described as an attempt to take into account the unexplained and unpredictable errors using the CAH (Dulay et al., 1982).

2.3.2 Machine-Assisted Pronunciation Analysis

In conjunction with the more traditional ways of assessing pronunciation difficulties, newer methods with the help of tools and software have been becoming more prominent in recent years. Tools like speech recognition software and acoustic analysis programs have been increasingly employed to analyze both individual sounds (segmental features) and aspects such as stress and intonation (suprasegmental features).

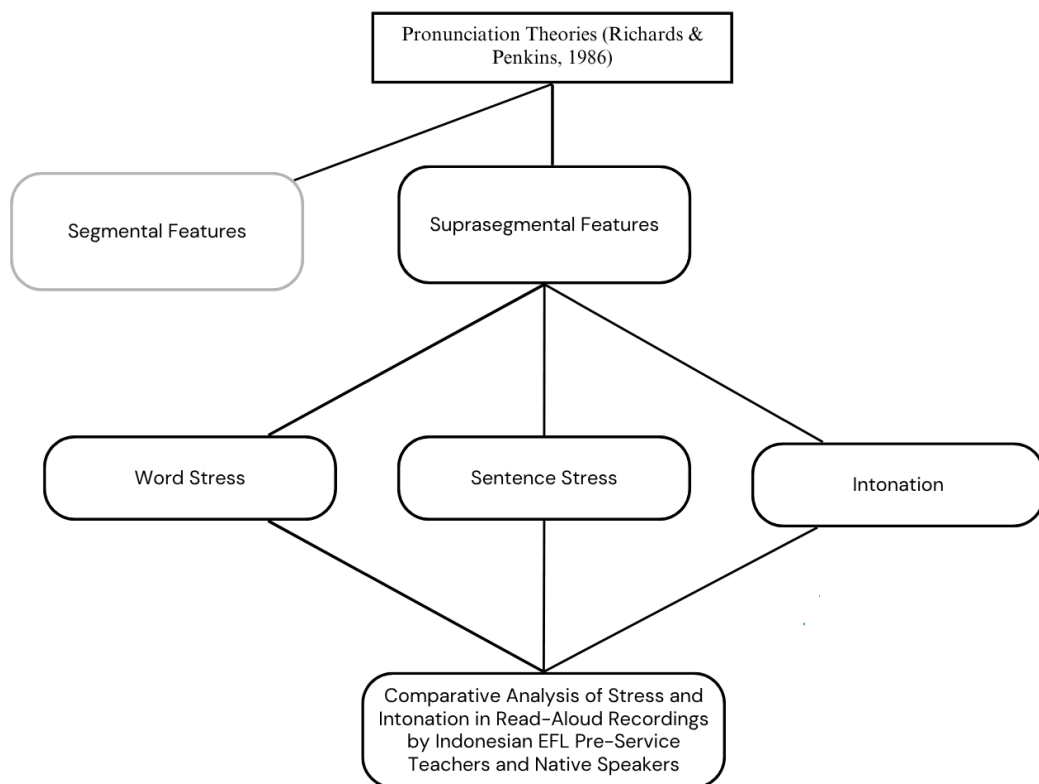
In one case, Evers and Chen (2020) conducted a quasi-experiment using an Automatic Speech Recognition (ASR) system called “Speechnote – Speech to Text” to determine the effects of peer feedback on participants. The participants were divided into two test groups; one received peer feedback and the other without it. It was found that the group receiving peer feedback outperformed the control group in aspects such as comprehensibility and spontaneous speech.

In a similar vein, Benitez-Correa et al. (2020) explored the potential of Audacity, an audio recording and editing software, to improve pronunciation in English as a Foreign Language (EFL) learners. In their study, they focused on the pronunciation of regular past-tense verbs among advanced EFL students at Universidad Técnica Particular de Loja (UTPL) in Ecuador. The study found that Audacity not only helped improve the learners' pronunciation but also created a relaxed and autonomous learning environment.

Acoustic analysis software such as Praat (Boersma & Weenink, 2024) has also become standard in phonological research. Its ability to visualize pitch contours and stress distribution has led to its use in studies examining suprasegmental features with precision, offering a robust method of phonological assessment suitable for EFL learners.

2.4 Theoretical Framework

Richards and Perkins (1986) classified the features in pronunciation, two of which, segmental and suprasegmental features, are the focus of this research. As a means of analysing word stress, sentence stress, and intonation patterns within English as a Foreign Language (EFL) participants and native speakers' recordings, this study adopts a diagnostic reading passage commonly used in phonetic analysis: The North Wind and the Sun (NWS), as recommended by the International Phonetic Association (1999). This passage has been widely implemented due to its phonetic diversity, structured nature, and prior use in cross-linguistic studies (e.g., Deterding, 2006). Its length and design make it suitable for acoustic segmentation and Praat analysis, which will be employed in analyzing participants' stress and intonation pronunciation (Boersma & Weenink, 2024).



2.1. Theoretical Framework

CHAPTER III

RESEARCH METHODOLOGY

3.1. Research Design

This study employs a comparative acoustic-phonetic analysis design to examine suprasegmental pronunciation features, specifically, word stress, sentence stress, and intonation, in the speech of Indonesian pre-service EFL teachers (Johnson, 2012). The analysis centers on controlled read-aloud data and compares participants' production against a set of native-speaker models to gather deeper insights into prosodic patterns of each group of participants and not to judge either one.

While qualitative in its interpretive orientation, the study integrates quantitative techniques through precise acoustic measurements, including normalized pitch in semitones and syllable duration, to ensure systematic and replicable comparisons. This approach enables the identification of prosodic tendencies without relying on impressionistic judgments. Praat software (Boersma & Weenink, 2024) is used to extract and visualize relevant acoustic parameters, and the data are processed to account for variation in speech rate and pitch range across speakers.

The methodological structure reflects the recommendations of prior empirical studies in phonetics and second language pronunciation, where segment selection, duration normalization, and semitone-based pitch scaling were used to compare speaker groups across gender and proficiency levels (e.g., Mennen et al., 2012; Rattanavich, 2013).

By targeting specific utterances and acoustic features, this design allows for focused, interpretable comparisons that illuminate the prosodic profiles of

each speaker group. The outcome contributes to understanding how intelligibility-relevant suprasegmental features manifest in EFL teacher speech.

3.2. Data Preparation

3.2.1. Setting and Participants

The participants in this study are four pre-service EFL teachers enrolled at the Islamic University of Indonesia. They were selected purposefully due to their stage in teacher education (i.e., having undergone a microteaching course and were ready to participate in a mandated teaching internship as part of their training); they represent a population tasked with modeling intelligible English speech in future classrooms. This makes their suprasegmental competence, particularly in stress and intonation, of pedagogical significance (Jenkins, 2000; Hahn, 2004).

The relevance of this group lies in their dual role as language learners and future instructors. Prior research indicates that pre-service teachers often lack confidence and training in pronunciation, especially suprasegmental features (Bai & Yuan, 2018; Murphy, 2014). Their pronunciation habits are likely to be transferred to learners, given their position as primary language models (Derwing & Munro, 2005). Thus, their speech provides a valuable lens for the state of EFL teachers' prosodic awareness.

The small sample size reflects the qualitative, exploratory nature of the study, which emphasizes depth over generalizability. As noted by Patton (2002) and Merriam and Tisdell (2016), purposeful sampling in qualitative research enables a rich, context-sensitive analysis of language behavior, especially in focused domains such as interlanguage phonology.

3.2.2. Types and Sources of Data

The primary data comprises two sets of audio recordings: (1) four recordings of Indonesian pre-service EFL teachers reading *The North Wind and the Sun* (NWS), recorded, monitored, and processed directly by the researcher, and (2) four native speaker recordings sourced from phonetic archives, including one from <https://www.phonetics.expert/english> (Phonetics.expert, n.d.) and three from <https://usefulenglish.ru/phonetics/listening-for-american-accents> (Useful English, n.d.). Each recording features a reading of the same diagnostic passage, which is reproduced in full below:

The North Wind and the Sun were disputing which was the stronger when a traveler came along wrapped in a warm cloak. They agreed that the one who first succeeded in making the traveler take his cloak off should be considered stronger than the other. Then the North Wind blew as hard as he could, but the more he blew the more closely did the traveler fold his cloak around him; and at last the North Wind gave up the attempt. Then the Sun shone out warmly, and immediately the traveler took off his cloak. And so the North Wind was obliged to confess that the Sun was the stronger of the two. (IPA, 1999)

The passage, originally introduced by the International Phonetic Association (IPA), has long served as a standardized instrument for phonetically balanced speech sampling (International Phonetic Association, 1999). It includes a broad distribution of phonemes, diverse sentence structures, and a range of prosodic contexts suitable for the analysis of stress and intonation. Its use is further supported by pedagogical and phonetic research (e.g., Celce-Murcia et al., 1996; Deterding, 2006), particularly in L2 settings where consistent, comparable spoken data are required for acoustic analysis.

3.2.3. Ethical Considerations

Before the data collection, participants were asked to provide informed consent. They were fully informed about the nature and purpose of the study, and were assured that their identities would be kept confidential. Permission was granted for the recordings to be used in analysis and for academic reporting.

3.3. Data Collection Technique

The primary data for this study consisted of read-aloud recordings of NWS passage, a phonetically balanced diagnostic passage recommended by the International Phonetic Association (1999). Each group of recordings, including that of Indonesian EFL pre-service teachers, were made in a controlled environment using a high-quality microphone setup to minimize background noise and ensure signal clarity suitable for acoustic analysis.

The speech files were saved in .wav format and accompanied by plain-text transcriptions, formatted to align with Montreal Forced Aligner (MFA) requirements. Following recording, alignment was performed using MFA (McAuliffe et al., 2017), which generated TextGrid annotations marking phoneme- and word-level boundaries. These aligned annotations enabled consistent extraction of relevant speech segments with a baseline accuracy of 10 milliseconds.

To optimize the efficiency and relevance of analysis, three representative words and utterances from the passage were selected. These items were chosen due to their inclusion of polysyllabic content words and phrase-level contrastive focus, making them particularly well-suited for analyzing suprasegmental features (Hahn, 2004; Deterding, 2006).

3.4. Data Analysis Technique

The acoustic analysis was conducted using Praat (Boersma & Weenink, 2024). Prior to measurement, each utterance underwent two forms of normalization to support cross-speaker comparison:

1. Pitch normalization: To account for gender-related vocal differences, pitch tracking was calibrated using a floor of 50 Hz and ceiling of 300 Hz for male speakers, and 100 Hz to 500 Hz for female speakers. This range was selected based on established norms in F0 analysis literature (Xu, 1999; Nolan, 2003). Such calibration ensured robust pitch extraction across participants without compromising data quality due to octave errors or unstable tracking.
2. Temporal normalization: Since each participant read at varying speeds, time normalization was applied. The timeline of each utterance was expressed as a percentage of total utterance duration, allowing segmental comparisons independent of absolute timing (Byrd & Mintz, 2010). This allowed the plotting of prosodic contours (such as pitch in semitones) on a comparable scale.

Pitch values were then converted into semitones relative to each speaker's median F0. This transformation enabled meaningful comparisons across speakers with different baseline pitches. The semitone conversion follows the formula: $[12 \times \log_2 (F0 / F0 \text{ Median})]$. This logarithmic scale reflects the human ear's perception of pitch differences more accurately than raw F0 in hertz (Johnson, 2012), making it especially suitable for suprasegmental analysis.

Duration was measured in two forms:

- Absolute duration (ms): The actual time taken to articulate each word or syllable, providing raw data for temporal analysis.

- Normalized duration (%): Each segment's duration expressed as a percentage of the entire utterance's length, used to account for differences in speaking rate (Nolan, 2003).

These metrics were extracted from the TextGrids and manually cross-referenced with waveforms to ensure accuracy. Analyses focused on polysyllabic content words and clause-final pitch movements, selected based on their prosodic salience in previous EFL research (Hahn, 2004; Chun, 2002). Tabulated results were visualized using spreadsheet software to facilitate cross-speaker comparisons.

This analytical framework enables clear, objective observations about word stress and intonation patterns across speaker groups while accounting for physiological and temporal variability.

3.5. Trustworthiness

To ensure credibility, the research follows a systematic process of data collection and acoustic verification. Each recording is reviewed and measured consistently within Praat. The analysis process is transparent and replicable, which supports the study's dependability and confirmability. Following the guidance of Lincoln and Guba (1985), the accuracy of the findings is also supported by clear documentation of procedures and alignment between the participants' actual speech and the interpretations provided.

CHAPTER IV

FINDINGS AND DISCUSSIONS

4.1. Findings

This chapter presents the acoustic findings comparing the suprasegmental features of four Indonesian EFL pre-service teachers and four American English native speakers. The analysis focuses on three key prosodic components: word stress, sentence stress, and intonation, using data derived from a controlled read-aloud of *The North Wind and the Sun* (NWS). All data were processed in Praat (Boersma & Weenink, 2024) and analyzed using normalized values to allow meaningful cross-speaker comparison.

4.1.1. Word Stress Patterns

To evaluate participants' realization of word-level stress, three polysyllabic content words from NWS passage were analyzed: *disputing*, *immediately*, and *obliged*. These words were selected due to their lexical stress contrast and pedagogical relevance for EFL learners (Hahn, 2004; Field, 2005). For each target word, intensity and syllable duration were measured and normalized as indicators of prominence.

Table 4.1

Normalized Syllable Durations and Mean Intensity of “Disputing,” “Immediately,” and “Obliged” in EFL Group

Speaker Group	Target Word	Syllable	Expected Stress	Mean Normalized Duration (%)	Mean Intensity (dB)
EFL	Disputing	dis-	<input type="checkbox"/>	34.23	74.49
		-put-	<input checked="" type="checkbox"/>	40.32	77.54
		-ing	<input type="checkbox"/>	25.45	79.91
	Immediately	im-	<input type="checkbox"/>	8.44	75.21
		-me-	<input checked="" type="checkbox"/>	22.70	78.46
		-diat ely	<input type="checkbox"/>	68.86	80.83
	Obliged	o-	<input type="checkbox"/>	12.74	77.70
		-blig ed	<input checked="" type="checkbox"/>	87.26	75.56

Table 4.2

Normalized Syllable Durations and Mean Intensity of “Disputing,” “Immediately,” and “Obliged” in NS Group

Speaker Group	Target Word	Syllable	Expected Stress	Mean Normalized Duration (%)	Mean Intensity (dB)
NS	Disputing	dis-	<input type="checkbox"/>	35.11	71.70
		-put-	<input checked="" type="checkbox"/>	38.26	76.3
		-ing	<input type="checkbox"/>	26.63	75.76
	Immediate ly	im-	<input type="checkbox"/>	15.60	76.53
		-me-	<input checked="" type="checkbox"/>	28.40	78.95
		-diately	<input type="checkbox"/>	69.26	77.12
	Obliged	o-	<input type="checkbox"/>	9.75	75.38
		-bliged	<input checked="" type="checkbox"/>	90.25	74.71

As shown, native speakers consistently produced greater intensity and longer relative duration on the lexically stressed syllables, in alignment with expected English stress patterns. EFL participants, while approximating stress placement, often displayed narrower contrasts in intensity and timing, suggesting reduced perceptual salience of stress cues.

These findings mirror those of Field (2005), who noted that Indonesian learners often produced reduced acoustic prominence in stressed syllables, likely due to L1 prosodic transfer or limited exposure to explicit pronunciation instruction. Hahn (2004) similarly emphasized that intelligibility improves when learners accurately signal primary stress through intensity and duration.

4.1.2. Sentence Stress Patterns

This section presents the analysis of sentence stress as produced by Indonesian EFL pre-service teachers compared with native speakers of American English. Three sentences were selected from *The North Wind and the Sun* passage because they each contain multiple content words that are syntactically and prosodically significant. These sentences allowed for a more focused comparison without needing to analyze the entire passage, which aligns with previous research on prosody (e.g., Hahn, 2004; Cruttenden, 2006).

To measure sentence stress, this study looked at two acoustic features: mean intensity (in decibels) and normalized syllable duration (as a percentage of the total sentence duration). These two measurements have been used in several studies (e.g., Sluijter & van Heuven, 1996; Hahn, 2004) as reliable indicators of stress, especially in controlled read-aloud tasks.

Table 4.3.*Mean Intensity and Normalized Duration of Content Words by EFL and NS*

Group	Word	Mean Intensity (dB)	Normalized Duration
Sentence 1			
EFL	disputing	77.24	21.33%
	stronger	74.71	15.53%
NS	disputing	75.40	18.75%
	stronger	75.63	20.10%
Sentence 2			
EFL	blew	80.04	9.89%
	hard	74.95	11.28%
NS	blew	78.24	11.43%
	hard	77.52	12.04%
Sentence 3			
	warmly	78.25	10.41%
EFL	immediately	80.23	14.35%
	traveler	77.59	9.10%
	cloak	71.00	7.60%
	warmly	75.14	13.68%

NS	immediately	77.56	13.52%
	traveler	76.11	10.90%
	cloak	74.38	9.70%

In the first sentence, both speaker groups placed stress on *disputing* and *stronger*, which are two key content words in the clause. However, the EFL speakers showed a bigger difference in emphasis between the two, stressing *disputing* more strongly both in intensity and duration. Meanwhile, native speakers showed a more balanced pattern, suggesting that they followed sentence-level stress patterns more consistently observed in English L1 data. This supports previous claims that native English speakers tend to place nuclear stress toward the end of declarative sentences (Cruttenden, 2006).

In Sentence 2, both groups highlighted the words *blew* and *hard*, but once again the EFL participants showed more difference between the two. *Blew* was given much more intensity, which might reflect a tendency to overemphasize certain action verbs. Native speakers, on the other hand, distributed the stress a bit more evenly. This could indicate that native prosodic intuition leads to a tendency to produce an acoustically more stable prominence across an utterance.

The third sentence showed the biggest contrast between the groups. EFL speakers seemed to put a lot of emphasis on *immediately* and *warmly*, which are adverbs, while giving less prominence to *traveler* and especially *cloak*. Native speakers spread the stress more evenly across all four content words, which might reflect a relatively higher control over broader prosodic phrasing. According to Chun (2002), this kind of wide stress domain helps listeners better understand the information structure of a sentence.

4.1.3. Intonation Patterns

To examine intonation contours, semitone-normalized pitch values were extracted from one utterance “*The North Wind and the Sun were disputing which was the stronger*” for all eight speakers. This sentence contains two prominent lexical stress targets (“disputing” and “stronger”), and is often used in

suprasegmental analysis due to its clause complexity and pitch movement potential (Deterding, 2006; Chun, 2002).

The pitch data were first normalized in semitones relative to each speaker's median F0 to account for gender- and physiology-based pitch differences (Nolan, 2003; Johnson, 2012). Temporal normalization was then applied by aligning pitch values across fixed intervals of 10ms, mapped to a shared time axis (0 to 100%). This approach allows for direct comparison across recordings of different durations (Xu, 1999; Byrd & Mintz, 2010).

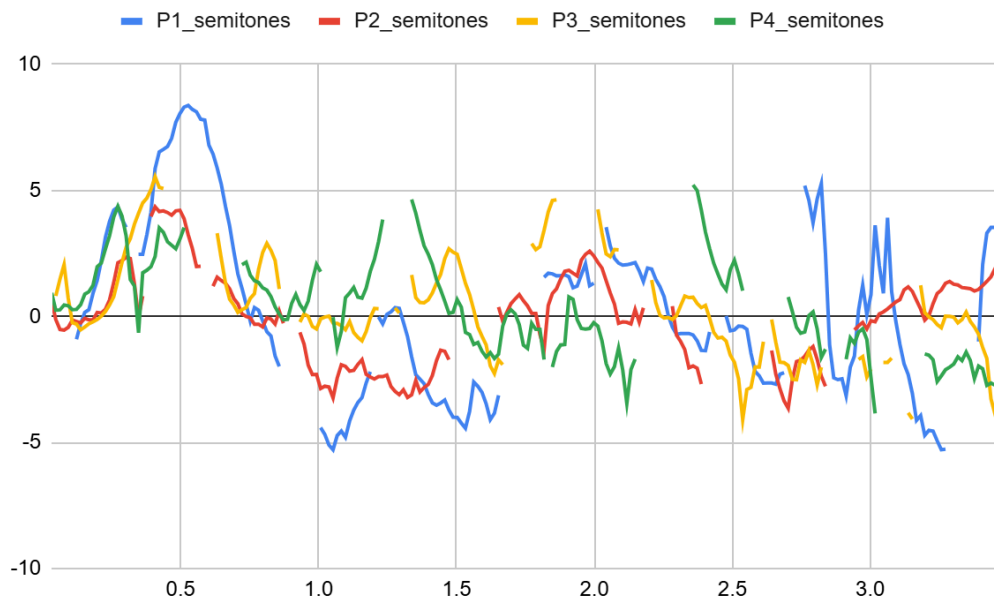


Figure 4.1. Normalized Pitch Contours Four EFL Speakers

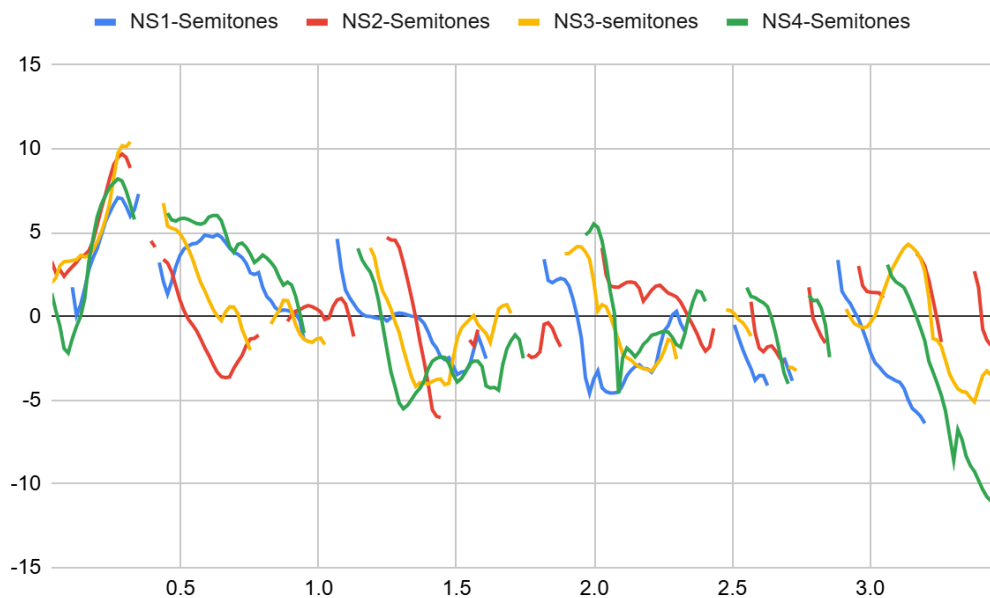


Figure 4.2. Normalized pitch contours for four NS speakers

Visually, native speaker contours (NS1–NS4) showed more gradual rises and controlled falls across the utterance, particularly at clause junctures. For instance, NS3 and NS2 displayed clear pitch peaks at “disputing” followed by a downward glide, aligning with expected declarative intonation (Cruttenden, 2014). The native speaker group exhibited both broader pitch ranges and more dynamic contours, consistent with intonation phrasing frequently observed in L1 English discourse. This aligns with previous findings by Mennen et al. (2012) that second-language speakers tend to underutilize pitch excursions due to limited prosodic awareness.

In contrast, EFL contours exhibited greater pitch instability. P3 showed a sharp rise toward the middle of the utterance, with less consistent fall afterward. P1 and P2 produced flatter contours, indicating limited pitch range modulation. These tendencies suggest that, while some EFL participants approximated native-like patterns, others struggled to reproduce modulated pitch movements. Notably, while native contours reflected expected patterns for broad-focus declaratives, EFL contours were either overly compressed or disjointed, echoing

Chun's (2002) observation that prosodic training is often insufficient in EFL contexts.

4.2. Discussion

This chapter interprets the acoustic findings of suprasegmental production by Indonesian EFL pre-service teachers in comparison with American English native speakers. The discussion synthesizes results from word stress, sentence stress, and intonation pattern analyses to evaluate the prosodic tendencies exhibited by the EFL participants. The implications for intelligibility and pedagogical development are also explored.

4.2.1. Word Stress Realization

The word stress results revealed that both EFL and native speakers generally produced stress on the expected syllables of polysyllabic target words such as *disputing*, *immediately*, and *obliged*. However, native speakers consistently demonstrated clearer acoustic prominence through longer syllable duration and stronger intensity on stressed syllables. For example, in *immediately*, native speakers marked the second syllable (-*me*-) and final syllable (-*diately*) with more consistent and proportionate duration and intensity values, reflecting clearer stress contours.

In contrast, the EFL group showed signs of overgeneralization and acoustic imbalance. Stress was sometimes misplaced or rendered with insufficient contrast between syllables. These findings align with Hahn (2004), who emphasized that L2 learners often fail to produce sufficient acoustic differentiation for primary stress, reducing the perceptual salience needed for intelligibility. Field (2005) similarly noted that Indonesian EFL speakers may exhibit stress placement that appears correct but lacks the duration and intensity cues that define native-like prominence. This suggests an underlying influence of L1 timing patterns and limited prosodic training in EFL contexts.

4.2.2. Sentence Stress Patterns

The sentence stress results suggest that while EFL speakers do apply stress to appropriate content words, they tend to narrow their focus, often exaggerating one or two items and underemphasizing others. This could make their speech consistent or less easily processed in extended discourse. Native speakers showed a more even stress distribution, especially in longer sentences. These differences might come from different levels of prosodic awareness and training, or simply from exposure to diverse English speech input patterns over time (Mennen et al., 2012; Rattanaovich, 2013).

This section also highlights the importance of teaching sentence-level stress, not just word-level stress, in EFL contexts. Without proper training or awareness, learners might overlook how stress contributes to meaning and clarity at the sentence level (Hahn, 2004; Jenkins, 2000).

4.2.3. Intonation Contours

The pitch contour analysis revealed a clear difference in dynamic range and phrasing between the two groups. Native speakers produced broader, smoother intonation contours with consistent rises and falls aligned to prosodic units. These patterns conform to documented patterns of declarative sentence intonation, where pitch typically peaks at stressed syllables and gradually falls toward sentence-final elements (Cruttenden, 1997; Deterding, 2006).

EFL participants, however, displayed more erratic pitch trajectories. Some produced compressed contours with minimal pitch range, while others showed abrupt rises and drops inconsistent with documented norms in L1 English discourse. These deviations suggest limited prosodic control and possible interference from L1 intonation patterns. As Mennen et al. (2012) pointed out, L2

speakers often underutilize pitch movement due to reduced prosodic awareness and training.

Furthermore, the pitch normalization procedure conducted via semitone conversion relative to each speaker's median F0 enabled meaningful comparison across genders and speaking styles. The findings reaffirm that "native-like" intonation involves not only hitting appropriate pitch targets but also managing pitch transitions and phrase-final declinations, areas where EFL speakers still require support.

CHAPTER V

CONCLUSION

5.1. Conclusions and Implications

The combined findings from stress and intonation analysis indicate that while Indonesian EFL pre-service teachers demonstrate partial acquisition of English prosody, their production lacks consistency and acoustic clarity. Their prosodic realizations are sometimes intelligible but may not consistently align with discourse-level patterns typically observed in proficient L2 or native-speaker models. This reinforces Jenkins' (2000) proposition that suprasegmental features are central to mutual intelligibility in international contexts and that teacher models are critical in shaping learner outcomes.

Given the relatively systematic misalignment of stress distribution and intonation control, the results point to a need for explicit prosody instruction in EFL teacher education. Tools such as Praat offer valuable visual and auditory feedback that help learners identify data-informed pitch contours, duration shifts, and intensity variation, thereby increasing metacognitive awareness of prosody (Levis & Pickering, 2004; Derwing & Munro, 2005). Research has shown that focused instruction on suprasegmental elements can significantly improve both the comprehensibility and confidence of L2 speakers (Hahn, 2004; Saito & Lyster, 2012). Embedding these tools in a scaffolded learning environment, paired with opportunities for guided practice and reflection, can enhance learners' ability to interpret and produce prosodic cues aligned with discourse needs.

In broader terms, the findings advocate for a shift in pedagogical priorities within pronunciation curricula, from a segmental-heavy approach to a more balanced framework that values prosody as communicatively meaningful. As Chun (2002) and Jenkins (2000) note, intelligibility and listener comprehension are shaped not only by accurate articulation but also by how speech is timed,

emphasized, and modulated. By equipping future teachers with both theoretical knowledge and practical strategies for teaching stress and intonation, EFL programs can better prepare them to model intelligible speech and diagnose prosodic issues in their own learners.

5.2. Limitations and Suggestions for Future Research

This study analyzed a small sample of recordings and focused on three key utterances from a controlled passage. While this allows for targeted analysis and cross-speaker comparability, it does not capture the full range of prosodic behavior in spontaneous or interactive speech. The reliance on read-aloud data, though methodologically efficient, may have influenced the naturalness of participants' intonation.

Additionally, while semitone and duration normalization helped control for speaker differences, future studies could integrate perceptual testing or listener-based intelligibility ratings to triangulate acoustic findings. Expanding the speaker pool and including longitudinal data would also help trace developmental progress in prosody acquisition.

References

- Arikan, A., & Yilmaz, S. (2020). Pronunciation problems of Turkish pre-service EFL teachers. *International Journal of Contemporary Educational Research*, 7(1), 1–10.
- Astawa, N. L. P. N. S., Artini, L. P., & Nitiasih, P. K. (2017). Needs analysis of academic-English speaking material in promoting 21st century skills. *International Journal of Language and Literature*, 1(1), 13–20. <https://doi.org/10.23887/ijll.v1i1.9616>
- Bai, B., & Yuan, R. (2018). EFL teachers' beliefs and practices about pronunciation teaching. *ELT Journal*, 72(4), 398–407. <https://doi.org/10.1093/elt/ccy012>
- Benitez-Correa, C., Gonzalez-Torres, P., & Vargas-Saritama, A. (2020). Audacity and pronunciation improvement in EFL learners. *International Journal of English Linguistics*, 10(5), 1–10. <https://doi.org/10.5539/ijel.v10n5p1>
- Boersma, P., & Weenink, D. (2024). *Praat: Doing phonetics by computer* (Version 6.3.01) [Computer software]. <http://www.praat.org>
- Brazil, D., Coulthard, M., & Johns, C. (1980). Discourse intonation and language teaching. *TESOL Quarterly*, 14(4), 431–440. <https://doi.org/10.2307/3586487>
- Brown, G., & Yule, G. (1983). *Teaching the spoken language*. Cambridge University Press.
- Byrd, D., & Mintz, T. (2010). *Discovering speech, words, and mind*. Wiley-Blackwell.
- Canan, S., & Oktay, A. (2021). Pronunciation challenges in Turkish EFL teacher education. *Journal of Language and Linguistic Studies*, 17(1), 45–60. <https://doi.org/10.17263/jlls.851022>
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge University Press.
- Chun, D. (2002). *Discourse Intonation in L2: From theory and research to practice*. John Benjamins.
- Corder, S. P. (1973). *Introducing applied linguistics*. Penguin. [Cambridge University Press & Assessment](https://www.cambridge.org/9780521224850)
- Corder, S. P. (1981). *Error analysis and interlanguage*. Oxford University Press.

- Cruttenden, A. (2006). *Gimson's pronunciation of English* (7th ed.). Arnold.
- Cruttenden, A. (2014). *Intonation* (3rd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511805107>
- Crystal, D. (1992). *An encyclopedic dictionary of language and languages*. Blackwell.
- Darmawan, I. G. N. (2019). Integrating language lab-based pronunciation teaching in EFL contexts: Challenges and potentials. *JEELS (Journal of English Education and Linguistics Studies)*, 6(2), 223–248.
- Demirezen, M. (2007). The /æ/ and /ə/ contrast as a case of fossilized pronunciation error for Turkish English teachers. *Journal of Language and Linguistic Studies*, 3(1), 83–97. <https://dergipark.org.tr/en/pub/jlls/issue/9925/122836Home+4Home+4Home+4>
- Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, 39(3), 379–397. <https://doi.org/10.2307/3588486>
- Dewi, R. S., & Pratiwi, K. D. (2020). A review of pronunciation teaching in Indonesian EFL contexts. *Language Circle: Journal of Language and Literature*, 14(2), 201–211. <https://doi.org/10.15294/lc.v14i2.23269>
- Deterding, D. (2006). The North Wind and the Sun: A short passage for the phonetic analysis of English speech. *Journal of the International Phonetic Association*, 36(2), 187–196. <https://doi.org/10.1017/S0025100306002544>
- Dirven, R., & Oakeshott-Taylor, J. (1984). *Listening comprehension (Part I)*. TBL.
- Dulay, H., Burt, M., & Krashen, S. (1982). *Language two*. Oxford University Press.
- Evers, V., & Chen, Y. (2020). Peer feedback and pronunciation improvement using ASR. *Journal of Educational Technology & Society*, 23(2), 45–58.
- Field, J. (2005). *Intelligibility and the listener: The role of lexical stress*. *TESOL Quarterly*, 39(3), 399–423. <https://doi.org/10.2307/3588487>
- Gussenhoven, C. (2004). *The Phonology of Tone and Intonation*. Cambridge University Press.

- Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, 38(2), 201–223. <https://doi.org/10.2307/3588378>
- Halliday, M. A. K. (1967). Notes on transitivity and theme in English: Part 2. *Journal of Linguistics*, 3(2), 199–244. <https://doi.org/10.1017/S0022226700016613> Cambridge University Press & Assessment
- Harmer, J. (1991). *The practice of English language teaching*. Longman.
- Hunt-Gómez, C., & Navarro-Pablo, M. (2020). Pronunciation errors in Spanish EFL teachers. *Revista Española de Lingüística Aplicada*, 33(1), 123–140. <https://doi.org/10.1075/resla.18034.hun>
- International Phonetic Association. (1999). *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge University Press.
- Jenkins, J. (2000). *The phonology of English as an international language*. Oxford University Press.
- Johnson, K. (2012). *Acoustic and auditory phonetics* (3rd ed.). Wiley-Blackwell.
- Kapranov, O. (2019). Pronunciation difficulties of Norwegian pre-service EFL teachers. *Scandinavian Journal of Educational Research*, 63(3), 345–360. <https://doi.org/10.1080/00313831.2017.1415961>
- Levis, J., & Pickering, L. (2004). Teaching intonation in discourse using speech visualization technology. *System*, 32(4), 505–524. <https://doi.org/10.1016/j.system.2004.09.002>
- Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. University of Michigan Press.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic Inquiry*. Beverly Hills, CA: Sage Publications, Inc.
- Marcellino, M. (2008). English language teaching in Indonesia: A continuous challenge in education and cultural diversity. *TEFLIN Journal*, 19(1), 57–69. <https://doi.org/10.15639/teflinjournal.v19i1/57-69>
- Mennen, I., Schaeffler, F., & Docherty, G. (2012). Cross-language differences in fundamental frequency range: A comparison of English and German. *Journal of the Acoustical Society of America*, 131(3), 2249–2260. <https://doi.org/10.1121/1.3681915>

- McAuliffe, M., Socolof, M., Mihuc, S., Wagner, M., & Sonderegger, M. (2017). Montreal Forced Aligner: Trainable text-speech alignment using Kaldi. *Interspeech 2017*, 498–502.
- Murphy, J. (2014). Intelligible, comprehensible, non-native models in ESL pronunciation teaching. *System*, 42, 258–269. <https://doi.org/10.1016/j.system.2013.12.007>
- Nguyen, T. T., & Newton, J. (2020). Pronunciation teaching in Vietnam: Teachers' beliefs and practices. *RELC Journal*, 51(1), 59–70. <https://doi.org/10.1177/0033688219883130>
- Nolan, F. (2003). Intonational equivalence: An experimental evaluation of pitch scales. *Speech Prosody 2002, Aix-en-Provence*, 259–262.
- Norrish, J. (1983). *Language learners and their errors*. Macmillan.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Sage Publications.
- Pennington, M. C., & Ellis, N. C. (2000). Cantonese speakers' memory for English sentences with prosodic cues. *Modern Language Journal*, 84(3), 372–389.
- Pennington, M. C., & Richards, J. C. (1986). Pronunciation revisited. *TESOL Quarterly*, 20(2), 207–225. <https://doi.org/10.2307/3586541>
- Phonetics.expert. (n.d.). *American English: North Wind and the Sun*. Retrieved June 4, 2025, from <https://www.phonetics.expert/english>
- Rattanaovich, S. (2013). A comparison of the English reading comprehension and moral judgment of students taught through drama-based and conventional instruction. *English Language Teaching*, 6(10), 72–84. <https://doi.org/10.5539/elt.v6n10p72>
- Rattanaovich, S. (2013). The effects of teaching stress and intonation on the English pronunciation of Thai university students. *English Language Teaching*, 6(6), 1–16. <https://doi.org/10.5539/elt.v6n6p1>
- Richard, J. C. (1974). Error analysis: Perspectives on second language acquisition. *Longman*.
- Saito, K., & Lyster, R. (2012). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /ɪ/ by Japanese learners of English. *Language Learning*, 62(2), 595–633. <https://doi.org/10.1111/j.1467-9922.2011.00639.x>

- Sluijter, A. M. C., & van Heuven, V. J. (1996). Acoustic correlates of linguistic stress and accent in Dutch and American English. *Proceedings of the Fourth International Conference on Spoken Language Processing (ICSLP 1996)*, 630–633. https://www.isca-speech.org/archive/icslp_1996/i96_0630.html
- Sluijter, A. M. C., & van Heuven, V. J. (1996). Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America*, 100(4), 2471–2485. <https://doi.org/10.1121/1.417955>
- Useful English. (n.d.). *Listening practice: American accents – North Wind and the Sun*. Retrieved June 4, 2025, from <https://usefulenglish.ru/phonetics/listening-for-american-accents>
- Utami, R. B., & Yulia, Y. (2022). Investigating stress placement among Indonesian EFL pre-service teachers. *Indonesian Journal of Applied Linguistics*, 12(1), 47–56. <https://doi.org/10.17509/ijal.v12i1.45299>
- Xu, Y. (1999). Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics*, 27(1), 55–105. <https://doi.org/10.1006/jpho.1999.0092>

**APPENDIX 1: Syllable Duration and Intensity Listings of EFL and NS
Group**

Speaker	Group	Word	Syllable	Stress (expected)	Duration (sec)	normalized Duration (%)	Intensity (dB)			Total Duration
P1	EFL		dis	Unstressed	0.1799	28.60%	75.78			0.629
		disputing	-put	Stressed	0.2599	41.32%	79.92			
			-ing	Unstressed	0.1892	30.08%	80.73			
			im-	Unstressed	0.06	6.99%	76.57			0.858
		immediately	me-	Stressed	0.169	19.70%	80.69			
			-diately	Unstressed	0.629	73.31%	82.45			

		obliged	o-	Unstressed	0.06	9.84%	75.93			0.61	
			bliged	Stressed	0.55	90.16%	75.31				
P2	EFL		dis	Unstressed	0.21	35.35%	71.24			0.594	
		disputing	-put	Stressed	0.254	42.76%	80.79				
			-ing	Unstressed	0.13	21.89%	83.43				
				im-	Unstressed	0.06	6.98%	74.63			0.86
		immediate	me-	Stressed	0.171	19.88%	75.78				
			-diatel	Unstressed	0.629	73.14%	79.84				
			y								
		obliged	o-	Unstressed	0.08	15.09%	78.75			0.53	

			bliged	Stressed	0.45	84.91%	77				
P3	EFL		dis	Unstressed	0.28	39.05%	73.99			0.717	
		disputing	-put	Stressed	0.268	37.38%	73.4				
			-ing	Unstressed	0.169	23.57%	81.17				
			im-	Unstressed	0.08	10.53%	70			0.76	
		immediately	me-	Stressed	0.22	28.95%	79				
			-diately	Unstressed	0.46	60.53%	81.79				
			obliged	o-	Unstressed	0.07	11.48%	76.35			0.61
				bliged	Stressed	0.54	88.52%	75.18			

P4	EFL		dis	Unstr essed	0.28	33.90%	76.96			
		disput ing	-put	Stres sed	0.329	39.83%	76.06			0.826
			-ing	Unstr essed	0.217	26.27%	74.34			
			im-	Unstr essed	0.05	9.28%	79.66			
		imme diatel y	me-	Stres sed	0.12	22.26%	78.38			0.539
			-diatel y	Unstr essed	0.369	68.46%	79.24			
		oblige d	o-	Unstr essed	0.09	14.54%	79.8			0.619
	bliged	Stres sed	0.529	85.46%	74.75					
NS1	NS		dis	Unstr essed	0.22	30.64%	73.65			0.718

		disput ing	-put	Stres sed	0.29	40.39%	80.84			
			-ing	Unstr essed	0.208	28.97%	75.56			
			im-	Unstr essed	0.03	5.57%	76.3			
		imme diatel y	me-	Stres sed	0.16	29.68%	83.12			0.539
			-diatel y	Unstr essed	0.349	64.75%	81			
		oblige d	o-	Unstr essed	0.06	9.84%	75.65			0.61
			bliged	Stres sed	0.55	90.16%	76.77			
NS2	NS		dis	Unstr essed	0.18	36.73%	72.86			0.49
		disput ing	-put	Stres sed	0.2	40.82%	79.94			

			-ing	Unstr essed	0.11	22.45%	77.37			
			im-	Unstr essed	0.3	41.78%	73.52			
		imme diatel y	me-	Stres sed	0.17	23.68%	76.7			1.099
			-diatel y	Unstr essed	0.629	87.60%	79.84			
		oblige d	o-	Unstr essed	0.05	10.00%	75.48			0.5
			bliged	Stres sed	0.45	90.00%	76			
NS3	NS		dis	Unstr essed	0.22	34.98%	67.15			0.629
		disput ing	-put	Stres sed	0.23	36.57%	66.67			
			-ing	Unstr essed	0.179	28.46%	71.32			

			im-	Unstr essed	0.05	7.25%	75.57			
		imme diatel y	me-	Stres sed	0.2	28.99%	75.6			0.69
			-diatel y	Unstr essed	0.44	63.77%	69.77			
		oblige d	o-	Unstr essed	0.06	12.02%	72.92			0.499
			bliged	Stres sed	0.439	87.98%	74.26			
NS4	NS		dis	Unstr essed	0.27	38.08%	73.17			0.709
		disput ing	-put	Stres sed	0.25	35.26%	77.75			
			-ing	Unstr essed	0.189	26.66%	78.8			
			im-	Unstr essed	0.05	7.81%	80.76			0.64

		immediate	me-	Stressed	0.2	31.25%	80.4			
		-diately		Unstressed	0.39	60.94%	77.9			
	obliged	o-		Unstressed	0.049	7.12%	77.49			0.688
		bliged		Stressed	0.639	92.88%	71.81			

APPENDIX 2: Word Duration and Intensity Listings of EFL and NS Group

Sp eak er	Wo rd	Int ens ity(dB)	Dur atio n(ms)	Nor mal ise d Dur atio n	Se nte nce Dur atio n(s)
P1	dis puti ng	79. 4	63 0	21. 72 %	2.9
	str on ger	75. 21	47 0	16. 21 %	
P2	dis puti ng	80. 18	60 0	17. 54 %	3.4 2
	str on ger	78. 3	58 0	16. 96 %	
NS 1	dis puti ng	78. 22	72 0	20. 40 %	3.5 3
	str on ger	77. 93	78 0	22. 10 %	
NS 2	dis puti ng	77. 76	49 0	16. 23 %	3.0 2
	str on ger	75. 48	58 0	19. 21 %	

P1	blew	83.29		180	9.23%	1.95
	hard	76.28		210	10.77%	
P2	blew	81.58		240	9.72%	2.47
	hard	76.04		290	11.74%	
NS1	blew	79.58		260	11.76%	2.21
	hard	78.43		250	11.31%	
NS2	blew	80.26		190	9.74%	1.95
	hard	78.4		270	13.85%	
P1	warmly	78.52		550	10.17%	5.41
	immediately	81.42		860	15.90%	
	traveler	78.15		370	6.84%	
	clerk	69.72		500	9.24%	

P2	war mly	77. 97		62 0	12. 16 %	5.1
	im me diat ely	79. 03		86 0	16. 86 %	
	tra vel er	77. 03		52 0	10. 20 %	
	clo ak	72. 24		44 0	8.6 3%	
NS 1	war mly	77. 85		56 0	13. 76 %	4.0 7
	im me diat ely	81. 61		54 0	13. 27 %	
	tra vel er	76. 61		44 0	10. 81 %	
	clo ak	77. 46		38 0	9.3 4%	
NS 2	war mly	75. 42		48 0	13. 37 %	3.5 9
	im me diat ely	76. 68		50 0	13. 93 %	

	tra vel er	76. 11		45 0	12. 53 %	
	clo ak	73. 65		41 0	11. 42 %	
P3	dis puti ng	76. 89		72 0	21. 49 %	3.3 5
	str on ger	75. 46		50 0	14. 93 %	
P4	dis puti ng	72. 47		84 0	24. 56 %	3.4 2
	str on ger	69. 86		48 0	14. 04 %	
NS 3	dis puti ng	68. 68		63 0	18. 10 %	3.4 8
	str on ger	74. 57		72 0	20. 69 %	
NS 4	dis puti ng	76. 92		71 0	20. 29 %	3.5
	str on ger	74. 53		64 5	18. 43 %	
P3	ble w	76. 4		21 0	10. 50 %	2

	hard	71.44		220	11.00%	
P4	blew	78.89		200	10.10%	1.98
	hard	76.03		230	11.62%	
NS3	blew	71.05		250	10.33%	2.42
	hard	77.09		300	12.40%	
NS4	blew	82.07		340	13.88%	2.45
	hard	76.17		260	10.61%	
P3	warmly	78.52		460	8.49%	5.42
	immediately	81.42		760	14.02%	
	traveler	78.15		570	10.52%	
	clerk	69.72		390	7.20%	

P4	war mly	77. 97		55 0	10. 83 %	5.0 8
	im me diat ely	79. 03		54 0	10. 63 %	
	tra vel er	77. 03		45 0	8.8 6%	
	clo ak	72. 3		27 0	5.3 1%	
NS 3	war mly	73. 55		65 0	13. 54 %	4.8
	im me diat ely	72. 87		70 0	14. 58 %	
	tra vel er	74. 31		51 0	10. 63 %	
	clo ak	70. 86		46 0	9.5 8%	
NS 4	war mly	73. 75		73 0	14. 04 %	5.2
	im me diat ely	79. 09		64 0	12. 31 %	

	tra vel er	77. 4		50 0	9.6 2%	
	clo ak	75. 56		44 0	8.4 6%	

