When Women Say *terserah* The Meaning Depends on The Intonation: Women's Prosody Analysis

Dzikri Ilma[1], Tengku Syarfina[2], Mahriyuni[3], Dedy Rahmad Sitinjak[4]

[1,2,3] Universitas Sumatera Utara

tengku.syarfina@usu.ac.id *

ABSTRACT

The research aimed to investigate two prosodic variations (intended meaning and unintended meaning) and to deepen the understanding of the emotional nuances conveyed by women's utterances of the word *terserah*, contributing insights into linguistic communication and emotional expression. This research employed a descriptive qualitative analysis design. Data were collected from women of different ethnic backgrounds and ages, all married to facilitate emotional expression during recording. Three speakers participated in this research: a 31-year-old Javanese woman (P1), a 35-year-old Banjanese woman (P2), and a 40-year-old Bataknese woman (P3). Prosodic elements were measured using PRAAT, a software tool used to analyze prosodic elements such as frequency, intensity, and duration. The findings revealed significant differences between intended and unintended meanings. Intended meaning was characterized by lower pitch, lower intensity, and longer duration with a descending tone. In contrast, unintended meaning exhibited higher pitch, higher intensity, and shorter duration with an ascending tone. In conclusion, the prosodic variations indicate different emotional expressions. However, contextual factors, individual psychological conditions, body language, and facial expressions also play crucial roles in determining the emotional meaning of verbal communication. Therefore, a comprehensive understanding of emotions requires an approach that considers all these aspects together.

INTRODUCTION

It has long been acknowledged that women's language often presents interpretive challenges, frequently leading to misinterpretation by men. One contributing factor is the inherent differences in linguistic features between women and men. This issue frequently arises in relationships, causing miscommunication and conflict. Tannen's (2005) theory on the differing communication styles of men and women underscores this phenomenon. According to Tannen (2005), women often use language to nurture relationships and foster understanding, whereas men use language primarily to convey information and address problems. Women typically communicate emotionally and employ implicit methods to express their ideas, either through speech or action, while men tend to communicate in a more logical manner. This complexity in understanding women's speech has given rise to jokes, often suggesting that men need a

specialized "dictionary" to decipher women's language. This issue has become a recurring topic in social discourse and humor.

In the *Kamus Besar Bahasa Indonesia* (KBBI), the word *terserah* is defined as "up to someone" or "as someone wishes," implying that a matter or decision has been left for another to determine. It can also mean returning to someone to inform or provide information. As an adjective, *terserah* can describe a situation or condition that is trivial or meaningless. *Terserah* is one of the words that men often find difficult to interpret, particularly in romantic relationships, where intonation significantly influences understanding. Numerous blogs, jokes, and memes explore the complexity of interpreting women's speech from psychological perspectives, but little attention has been given to prosodic approaches. This research addresses this gap by analyzing the prosody of women's utterances using phonetic-acoustic methods.

Effective communication between speaker and listener relies on the speech chain, which encompasses linguistic, psychological, and acoustic levels (Denes & Pinson, 1963). Various interrelated paralinguistic elements influence this process. The listener's perception of the speaker's speech can yield different interpretations due to the complexities inherent in the speech chain. Consequently, this study aims to document the acoustic properties of women's utterances and examine their intonation patterns by analyzing prosodic features, such as frequency, intensity, and duration, and their correlation with emotional expression.

Prosody encompasses suprasegmental aspects of speech, including pitch, duration, amplitude, and voice quality, which convey lexical, post-lexical, and paralinguistic meanings (Xu, 2019). These suprasegmental features enable humans to express meaning, emotions, and nuances through patterns of melody, word emphasis, and rhythm. For instance, descending intonation at the end of a sentence may indicate sadness or tenderness, while stress on specific syllables can emphasize meaning or importance in a sentence.

Heuven et al. (2007) identified four phonetic components of prosody that influence verbal expression during speech: variation in pitch, loudness, voice quality, and timing. Pitch variation, determined by the rate of vocal cord vibration, reflects emotional states or emphasizes information. Loudness, influenced by sound intensity, impacts clarity and precision. Voice quality relates to the articulation's clarity, affecting the sound's perception. Timing variations, such as speech rhythm acceleration or deceleration, provide emphasis or modify sentence meaning. These components collectively enhance the richness of verbal communication.

Individual pitch variation contributes to unique vocal characteristics, influenced by physiological differences in the larynx. Habitual pitch ranges from 50–250 Hz for men and 120–500 Hz for women (Heryono, 2019). Ladefoged (1996) noted that women's smaller, lighter vocal cords produce higher pitch ranges. Pitch, measured in Hertz (Hz), reflects the perceived highness or lowness of tone, while loudness, measured in decibels (dB), indicates sound intensity (Heryono, 2019). Intonation serves as a vehicle for emotional expression, encompassing states like happiness, boredom, or shock (Crystal, 1989).

The duration of speech also affects meaning. Duration, defined as the timing of articulatory sequences and acoustic signals, influences rhythm, structure, and intonation (Syarfina, 2009). It also aids in understanding phonetic and phonological processes, dialectal variations, and accent differences. Measured in seconds (s) or milliseconds (ms), duration provides insights into articulation and comprehension.

The relationship between intonation and emotion lies in how pitch patterns convey feelings. Intonation patterns vary with context and emotional expression, aligning with Schwartz and Pell's (2012) findings that synchronizing prosodic features such as speech rate, intensity, and duration enhances empathy and emotional involvement in conversations.

This research aims to contribute to understanding women's speech prosody, with practical applications for laypersons, particularly men, struggling to interpret women's language. The study focuses on analyzing the prosody of *terserah* using PRAAT software to measure frequency, intensity, and duration. Data collected through PRAAT will be analyzed for differences and

similarities. The research objectives are to (a) identify differences in prosody between two intonation patterns in women's utterances and (b) determine whether sound contours indicate distinct emotional expressions.

REVIEW OF LITERATURE

The theoretical framework of this research is grounded in prior studies that examine the relationship between prosody and emotion. Extensive research has been conducted on prosody, with studies ongoing for decades. For instance, Batalemba et al. (2024) found that the pronunciation of the word *anjing* exhibits variations in prosody, such as pitch, intensity, and duration. These variations influence the meaning or intent conveyed by the word, depending on the context in which it is spoken. The study concluded that the prosody of the word *anjing* not only reflects different emotions or feelings but also impacts how messages are interpreted in various situations.

Similarly, Chang et al. (2023) investigated how emotions affect the acoustics and perception of Mandarin tones. In their study, professional actors produced Mandarin tones expressing anger, fear, happiness, sadness, and neutrality. The findings demonstrated that emotions significantly influence the acoustics of Mandarin tones, varying based on specific tones and emotions. Emotional expressions also had a greater impact on Mandarin tone identification than the tones themselves. Additionally, syllables presented within a carrier phrase were identified more accurately for both tone and emotion, highlighting a complex interaction between lexical tones and emotions in Mandarin.

Aligned with these discoveries, Fonagy et al. (1963) explored the melodic patterns of ten different emotional attitudes, chosen somewhat arbitrarily. The findings suggest that the level of an intonation pattern conveys meaningful emotions, which can also be observed at the syllable level. The study analyzed records of conversations, dramas, radio plays, and experiments with actors, as well as vocal and instrumental musical compositions. Similarly, Bänziger et al. (2005) found that emotional expressions portrayed by actors show significant variation in the average level and range of fundamental frequency (F0) as a function of emotional arousal. However, evidence for qualitatively different contour shapes associated with specific emotions remains limited.

A recent study by Ertürk et al. (2024) examined gender differences in the perception and production of emotional prosody. The findings revealed no significant gender differences in the perception of emotional prosody, though differences were observed in prosodic production. Women tended to use higher pitches to express joy compared to sadness and neutral states. Both men and women varied their vocal intensity depending on the emotional state, and speech tempo was slower when expressing sadness than when expressing anger, joy, or neutrality. Additionally, Turkish speakers displayed similar abilities to perceive emotional prosody as speakers of other languages but utilized fluctuations in vocal intensity when producing emotional prosody.

The study of prosody and its role in conveying emotion spans multiple disciplines. According to Mozziconacci (2002), this research topic lies at the intersection of linguistics, psychology, neuropsychology, ethnology, cognition, artificial intelligence, information technology, acoustics, physiology, and phonetics. The present study contributes to this multidisciplinary field by advancing the understanding of communication through a phonetic-acoustic lens.

METHOD

The research employed a descriptive qualitative analysis design to observe women's utterances of the word *terserah*. In everyday language use, *terserah* is often employed to indicate that a decision or choice has been left to another person. However, the word *terserah* can sometimes convey nuances or meanings that are more complex than its lexical definition. For instance, depending on the intonation or expression used, *terserah* may reflect indifference,

apathy, or even subtle disagreement with a given situation or choice. Thus, while the basic meaning of *terserah* implies deference to someone else's decision, its usage in specific contexts may reveal deeper emotional nuances or attitudes.

Three speakers participated in this study, each representing different ethnic backgrounds and age groups: a 31-year-old Javanese woman (P1), a 35-year-old Banjanese woman (P2), and a 40-year-old Bataknese woman (P3). The speakers were all married, as the context of a romantic relationship was considered essential for providing emotional background and facilitating authentic expressions of the word *terserah*. Voice recordings were made in a quiet setting using a Sony ICD-PX470 voice recorder.

At the beginning of the data collection process, the researcher conducted simple interviews with the participants. The speakers were asked whether their utterances of *terserah* corresponded to its lexical meaning (intended meaning) or if they had ever used the word to convey a meaning beyond its lexical definition (unintended meaning). All three speakers confirmed that *terserah* could express both intended and unintended meanings, depending on the context.

The recordings captured utterances of *terserah* in two distinct emotional situations: stable and unstable emotional states. These scenarios were designed to reflect typical interactions between husbands and wives. The stable emotional state, referred to as the "harmonious" context, represented the intended meaning of *terserah*. Conversely, the unstable emotional state, characterized by tension in the relationship, represented the unintended meaning. By categorizing *terserah* into these two contexts, the data provided a framework for analyzing prosodic differences in the utterances.

Larrouy (2024) supports the notion that emotions can be inferred not only from the words spoken but also from how they are spoken. For example, low, soft speech may indicate sadness or disappointment, whereas loud and rapid speech may suggest anger or excitement. This highlights the importance of considering vocal characteristics such as pitch, volume, and speaking speed in interpreting emotions.

The recorded data were analyzed using PRAAT software, a widely used tool for phonetic analysis. According to Xu (2023), PRAAT enables the synthetic manipulation of pitch level, pitch range, and formant dispersion, which are essential for perception-oriented studies. In this study, the prosodic elements analyzed included pitch, loudness, and duration. These elements were examined to identify patterns in sound contours and to observe similarities and differences across the recordings.

The findings from the PRAAT analysis were interpreted in conjunction with theories and prior research on prosody and emotion in speech. The study aimed to uncover the emotional nuances conveyed by the prosodic features of *terserah* and how these features correlate with emotional states.

FINDINGS AND DISCUSSION
FINDINGS

The word *terserah* typically means "up to you" or "whatever," indicating that the speaker is allowing the listener the freedom to decide or that the speaker has no strong preference. However, the meaning of *terserah* can change significantly when spoken in a state of emotional upset, anger, or similar feelings. This phenomenon aligns with findings from research on the word "whatever" by Benus et al. (2007). Their study demonstrated that the prosody (intonation and pronunciation) of the word "whatever" reliably indicates the level of negativity in the context where it is used. For instance, whatever was perceived as more negative when used in situations involving harsh or demeaning language, such as in responses on the Jerry Springer Show or monologues containing strong language like "shit." Negative perceptions of whatever were also influenced by the lexical context in which it appeared.

Similarly, when a woman is emotionally upset, angry, or disappointed and responds with *terserah* to a choice or offer from her romantic partner, the underlying meaning may be more

complex than simply granting him the freedom to decide. In such situations, *terserah* might not just signify deference to the partner's decision but may also reflect dissatisfaction, displeasure, or even rejection of what is being offered or expected. Therefore, it is crucial to consider not only the literal meaning of the word but also the emotional context and underlying nuances of verbal communication.

To better understand these nuances, the utterance of *terserah* in different emotional contexts can be analyzed through a prosodic analysis approach. This approach observes variations in prosody—such as pitch, intensity, and duration—allowing for a deeper understanding of how prosodic differences manifest in different emotional states when the word *terserah* is spoken. Such an analysis sheds light on the complex interplay between linguistic expression and emotional states, providing valuable insights into verbal communication dynamics.

Table 1. Unintended meaning utterance

Participant	Prosody Analysis		
	Frequency	Intensity	Duration
P1	260.7 Hz	57.11 dB	0.69 s
P2	225.7 Hz	57.05 dB	0.75 s
P3	215.2 Hz	66.85 dB	0.93 s

The data analysis presented in Table 1 reveals that the average pitch is above 238.46 Hz, the average loudness exceeds 60.64 dB, and the average duration is 0.79 seconds. Overall, Table 1 demonstrates that when women utter the word *terserah* in a context reflecting unintended meaning, they tend to employ a higher pitch, slightly elevated loudness, and a shorter duration in their utterance. These findings suggest that prosodic variations—such as increased pitch and loudness coupled with reduced duration—play a significant role in signaling emotional nuance and contextual meaning in verbal communication.

Table 2. Intended meaning utterance

Participant	Prosody Analysis		
	Frequency	Intensity	Duration
P1	179.9 Hz	53.55 dB	0.91 s
P2	179.7 Hz	55.46 dB	0.79 s
P3	193.8 Hz	61.83 dB	1.33 s

Table 2 indicates that the average pitch is above 184.5 Hz, the average loudness is 57.93 dB, and the average duration exceeds 1.00 second. These results suggest that when women say *terserah* in the context of intended meaning, they tend to use a lower pitch, lower intensity, and a longer duration in their utterances. Comparatively, the findings in Table 1 and Table 2 highlight significant differences between unintended and intended meanings. Utterances with unintended meaning are characterized by higher pitch, higher intensity, and shorter duration, whereas those with intended meaning are marked by lower pitch, lower intensity, and longer duration.

To further discern significant variations, the utterance will be analyzed on a per-syllable basis, focusing on pitch and intensity. The word *terserah* consists of three syllables: ter-se-rah. Schirru et al. (2023) emphasized that linguistic datasets could be used to identify acoustic features critical to representing emotions, including signal acquisition, segmentation, construction, and encoding. In their analysis, words, syllables, phonemes (vowels and consonants), stress, and frequency tones were considered key variables.

To visualize the results, a subsequent table will display the soundwaves for each speaker's utterance of *terserah*, both for unintended and intended meanings. In the visual representation, the blue line will indicate pitch, representing the highness or lowness of the sound, while the green

line will denote intensity, reflecting the loudness or volume of the sound. To compare the pitch contour of both unintended and intended meanings, these will be marked with a red dashed line.

Analyzing the word syllabically allows for the identification of sound contour patterns within the utterance. Once these patterns are determined, the prosodic analysis will assess intonation and potential emotional expressions based on tonal variations, loudness, and duration in the utterances.
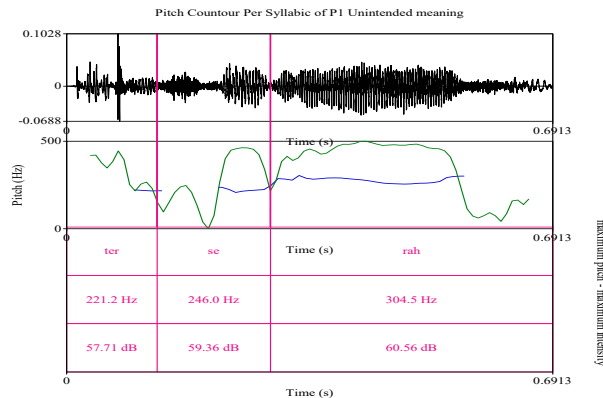

Figure 1. P1 Prosody changes at Unintended Meaning

Figure 1 illustrates the maximum prosodic element results for the syllables of the word *terserah* in the unintended meaning, as produced by the Javanese speaker, P1. The analysis reveals that the syllable ter exhibits a maximum pitch of 221.2 Hz and a maximum loudness of 57.71 dB. For the syllable se, the unintended meaning shows a maximum pitch of 246 Hz and a maximum loudness of 59.36 dB. These findings suggest that the syllable se in the unintended meaning is characterized by higher tones and greater volume compared to the other syllables.

The loudness contour for the unintended meaning displays a sharp rise in the middle, particularly during the syllable se, followed by a gradual decline. The final syllable, rah, in the unintended meaning is notable for having a maximum pitch of 304.5 Hz and a maximum loudness of 60.56 dB. This indicates that the syllable rah demonstrates a significantly higher pitch and a gradual rise at the end of the utterance, emphasizing its prosodic prominence in the unintended meaning context.
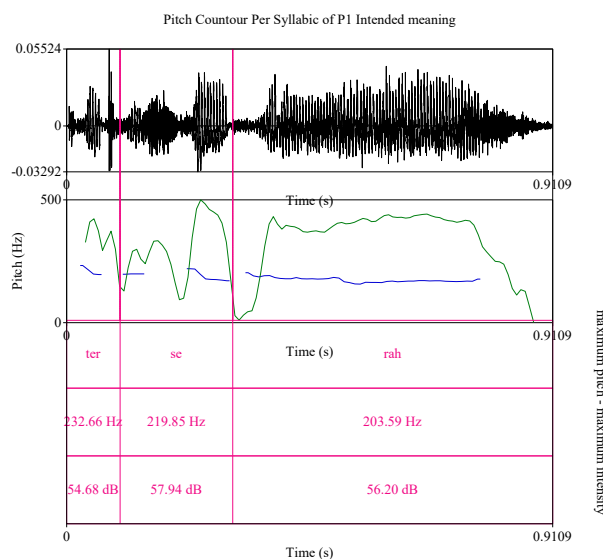

Figure 2.P1 Prosody changes at Intended Meaning

Figure 2 presents the maximum prosodic element results for the syllables of the word *terserah* in the intended meaning, as produced by the Javanese speaker, P1. The analysis shows

that the syllable ter has a maximum pitch of 232.66 Hz and a loudness of 54.68 dB. For the syllable se, the intended meaning exhibits a maximum pitch of 219.85 Hz and a maximum intensity of 57.94 dB. Higher tones are evident at the beginning of the utterance, particularly in the syllable ter. However, there is a sharp decline in pitch towards the end of this syllable, indicating a significant reduction in pitch and tone.

The final syllable, rah, in the intended meaning has a maximum pitch of 203.59 Hz and a maximum loudness of 56.20 dB. The prosodic pattern in the intended meaning is characterized by slightly lower tones and loudness, with a descending tone that concludes with the lowest pitch at the end of the utterance.

The prosody analysis of speaker P1 illustrates significant differences in pitch and loudness between the unintended and intended meanings. The unintended meaning features higher pitch and greater loudness compared to the intended meaning. Additionally, the pitch and loudness patterns vary between the two meanings: the unintended meaning displays a more dynamic pattern with greater variability, while the intended meaning follows a more stable and consistent pattern.
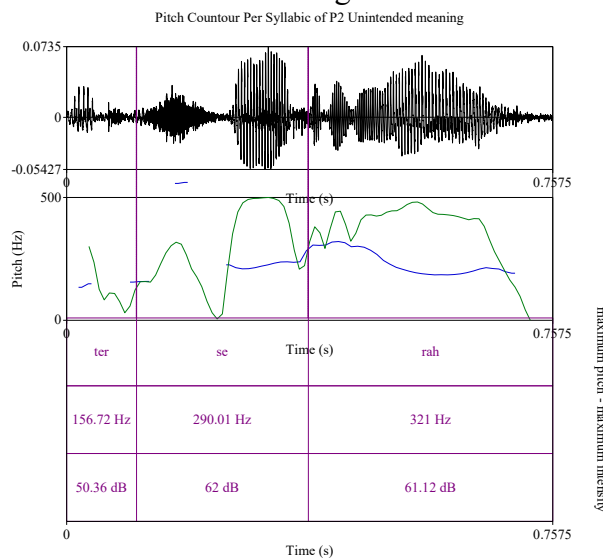

Figure 4. P2 Prosody changes at Intended Meaning

Conversely, Figure 4 presents the maximum prosodic elements for the intended meaning context. For the syllable ter, the analysis reveals a maximum pitch of 193.96 Hz and a loudness of 52.74 dB, resulting in a low and nearly flat pitch contour, which is similar to the unintended meaning. The intensity levels between the intended and unintended meanings show minimal differences. Similarly, the intended meaning of the syllable se exhibits a maximum pitch of 205.33 Hz and a maximum intensity of 58.46 dB, with both pitch and intensity remaining relatively consistent. For the syllable rah, the intended meaning has a maximum pitch of 203.15 Hz and a maximum loudness of 58.57 dB, maintaining a flat pitch contour and comparable intensity levels.

The prosody analysis for speaker P2 demonstrates a notable difference in pitch contour between the unintended and intended meanings. In the unintended meaning, the pitch contour features ascending tones in the middle syllable se. However, this pattern is absent in the intended meaning, where the pitch remains relatively flat. The loudness patterns for both meanings are nearly identical, with a descending tone toward the end of the utterance, reaching the lowest intensity.

The intensity analysis also correlates with speech duration. The duration comparison between the two contexts reveals only a 5-second difference, reflecting minimal variation in this prosodic feature.

Pitch Countour Per Syllabic of P3 Unintended meaning
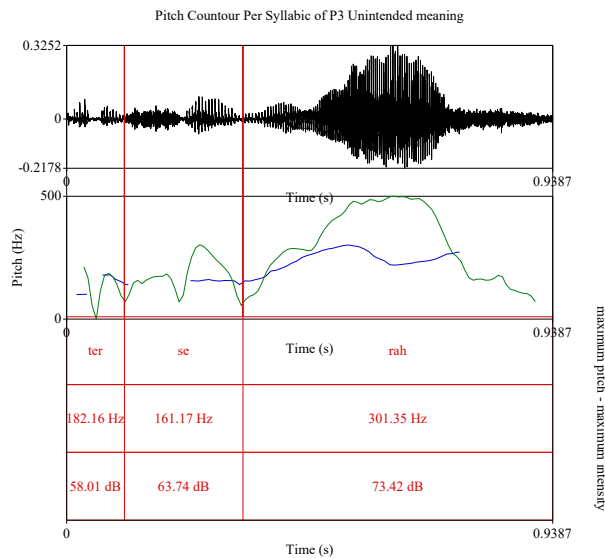
The prosody analysis of the Bataknese speaker, P3, reveals notable variations in pitch and intensity across the syllables ter, se, and rah. The syllable ter exhibits a maximum pitch of 182.16 Hz and an intensity of 58.01 dB, reflecting a lower pitch and moderate intensity. In contrast, the syllable se displays a maximum pitch of 161.17 Hz and an intensity of 63.74 dB, accompanied by a gradually rising pitch contour. The final syllable, rah, demonstrates a significant increase, with a maximum pitch of 301.35 Hz and an intensity of 73.42 dB, characterized by a higher rising pitch and a marked increase in intensity.
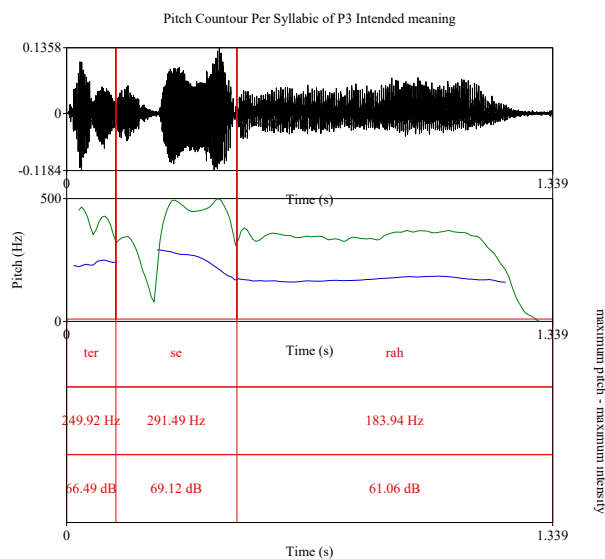

Pitch Countour Per Syllabic of P3 Intended meaning

Figure 6. P3 Prosody changes at Intended Meaning

The results presented in Figure 6 indicate the prosodic features of the intended meaning for the Bataknese speaker, P3. The syllable ter has a maximum pitch of 249.92 Hz and a loudness of 66.49 dB. For the syllable se, the maximum pitch reaches 291.49 Hz, with a maximum intensity of 69.12 dB. The pitch contour rises in the middle and gradually declines at the end, reflecting higher loudness for the intended meaning of the syllable se. The syllable rah demonstrates a maximum pitch of 183.94 Hz and a maximum loudness of 61.06 dB, with the pitch contour characterized by a steady decline. The intensity pattern for the intended meaning shows that the beginning of the syllable rah remains flat for a few seconds before sharply declining.

The prosody analysis of P3 highlights differences between the intended and unintended meanings. For the unintended meaning, the pitch contour exhibits ascending tones at the end of the utterance, whereas for the intended meaning, it begins with a standard tone, rises in the middle,

and slightly declines to flat tones. The intensity patterns also differ: for the unintended meaning, loudness is stable at the start of the syllable, rises and falls, and then rises again at the beginning of the third syllable before ending with a descending contour. In contrast, for the intended meaning, loudness begins with a sharp decline, rises sharply in the second syllable to form a stable pattern, and then sharply drops to the baseline at the end of the third syllable.

While the measurement data differ for each speaker, similarities in sound contour patterns are observed across the three participants, as evidenced by the prosody element data. Colorado (2002) noted that pitch has been recognized as a critical component of prosody since the 1950s, with duration and amplitude also identified as significant factors. Similarly, Oh et al. (2023) emphasized that pitch is the primary acoustic feature enabling listeners to differentiate emotions expressed by speakers. Syarfina et al. (2024) further explained that acoustic phonetics focuses on the physical properties of speech, including frequency, intensity, and duration, which are essential for analyzing and describing speech characteristics.

Based on these findings, the subsequent figures illustrate the pitch contours of the three participants, revealing consistent patterns for both unintended and intended meanings. These results underscore the crucial role of pitch as a key prosodic element in emotional communication.
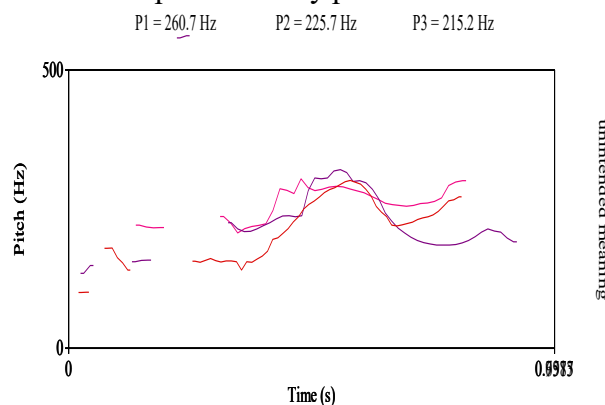
P1 = 260.7 Hz   P2 = 225.7 Hz   P3 = 215.2 Hz

Figure 7. Pitch Contour Unintended Meaning

Figure 7 illustrates the pitch contour for the unintended meaning of the word terserah. The analysis shows that P1 exhibits the highest pitch at 260.7 Hz, followed by P2 at 225.7 Hz, and P3 at 215.2 Hz. The pitch pattern for the unintended meaning reveals that the utterance begins with low tones, gradually rises to a peak in the middle, and then slowly descends, concluding with higher tones compared to the beginning. Based on the sound contour data, the rising tones at the end of the utterance suggest an emotion of "anger" or "cold anger."

This finding aligns with previous research highlighting the role of pitch in expressing anger. According to Rochman et al. (2008), anger is associated with distinct and consistent acoustic characteristics. When someone is angry, their voice exhibits an increased rate of speech, known as an elevated articulation rate. Additionally, the mean fundamental frequency (mF0) of the voice increases, resulting in a higher-pitched sound. Variability in pitch also rises, indicating greater fluctuations in the voice's frequency. Consequently, an angry person's voice tends to be faster, higher-pitched, and more varied than neutral or emotionless speech.

Moreover, many overt emotional expressions involve more than a single "pure" emotional state. Shield (2013) suggests that visibly angry individuals may also exhibit signs of fear, anxiety, sadness, or hatred. Emotional incidents are therefore subject to interpretation and reinterpretation by both the experiencer and observers. This implies that common emotions such as anger, happiness, and anxiety are often characterized by ascending tones at the end of an utterance. However, it is important to note that a high pitch at the end of an utterance is not exclusive to anger; it may also indicate other emotions such as happiness or anxiety.
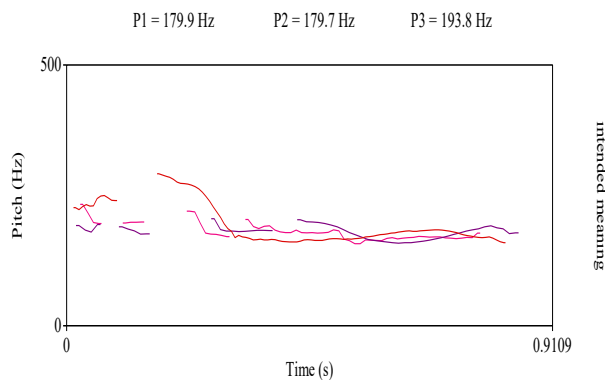
Figure 8. Pitch Contour Intended Meaning

Figure 8 illustrates the pitch contour for the intended meaning of the word terserah. P3, the Bataknese speaker, exhibits the highest pitch at 193.8 Hz. Interestingly, P3 uses a lower pitch for unintended meanings but a higher pitch for intended meanings. Although P3's utterance for the intended meaning has the highest pitch among the speakers, it remains lower than P3's pitch for unintended meanings. The sound contour pattern observed in speaker P3 is consistent with those of speakers P1 and P2. The second-highest pitch is from P1 at 179.9 Hz, which differs by only 2 Hz from P2's pitch at 179.7 Hz. This pattern indicates that the utterance begins slightly higher and gradually decreases toward the end. The longer duration of the intended meaning suggests that the word is pronounced slowly and smoothly, reflecting a stable emotion. Based on the sound contour data, the descending tones at the end of the utterance suggest an emotion of "sincerity" or "approval."

Each emotion can be characterized by a specific acoustic profile (Hammerschmidt et al., 2007). Emotions are often embodied in facial and bodily expressions, which can allow for objective assessment. However, emotions are also fleeting and may be ambiguously or partially expressed, creating room for subjectivity in observers' assessments of an individual's behavior (Shields, 2013). Changes in intonation within speech indicate emotional shifts and influence the perceived meaning of the speech. Understanding how someone feels, however, requires more than just listening to their tone of voice. Facial expressions, body movements, and the contextual situation must also be considered. Factors such as how loud or fast someone is speaking, whether they maintain eye contact, and the overall mood of the conversation are essential in interpreting emotions accurately and avoiding misinterpretations.

Some emotions exhibit fairly consistent acoustic features, while others are more variable (Chang, 2023). Understanding the true emotion expressed requires careful consideration of the context and the speaker's expressions. The findings of this research indicate that when women utter the word *terserah* with an unstable sound contour, it can be challenging to determine the precise emotion being conveyed.

In line with this, Ekberg et al. (2023) identified significant acoustic differences between surprise and other emotions. Interestingly, they found no acoustic distinctions between anger and happiness, which challenges most listeners' expectations of being able to differentiate between these emotions based on speech alone. Most studies on emotional speech focus on stylized emotions performed by actors and actresses. In such controlled settings, a few categories of emotions can be reliably identified by listeners, with consistent acoustic correlates. For instance, excitement is characterized by high pitch and fast speech, while sadness is associated with low pitch and slow speech. Hot anger involves over-articulation, a fast rate of speech, downward pitch movements, and elevated pitch overall. Cold anger, while sharing attributes with hot anger, is distinguished by a narrower pitch range (Colorado et al., 2002).

Colorado et al. (2002) further noted that emotions such as excitement, anger, fear, sadness, and sarcasm are marked by distinctive prosodic patterns. These include features like duration, mean fundamental frequency (f0 mean), f0 range, loudness, jitter (small frequency fluctuations),

spectral tilt, and accent shape. For example, excitement is characterized by fast duration, very high pitch, and loudness, while hot anger exhibits high pitch, a strong falling accent, and loudness. Fear shows jitter, sarcasm is marked by a prolonged accent with a late peak, and sadness is characterized by slow duration and low pitch. Cold anger also shares many attributes with hot anger but is distinguished by a lower pitch range. These prosodic patterns play a crucial role in identifying emotions through acoustic analysis of speech.

CONCLUSIONS

This research has demonstrated how prosodic elements significantly differentiate between the intended and unintended meanings of the word terserah in women's utterances. The study identified distinct emotional associations with voice contour patterns analyzed using the PRAAT application. Key findings indicate that higher pitch, shorter duration, and rising contours at the end of the utterance are associated with unintended meanings, conveying emotions such as "hot anger" or "cold anger." Conversely, longer duration, stable low-pitched contours, and descending tones at the end of the utterance are linked to intended meanings, reflecting emotions of "sincerity" and "approval."

However, this study is not without limitations. The small sample size of three participants from diverse backgrounds may limit the generalizability of the findings. Future research should expand the participant population to include a more representative and diverse sample, enabling a more comprehensive exploration of how intonation in utterances conveys emotions across different contexts and populations. Addressing these limitations and building on this foundation will advance the understanding of prosody's role in enhancing communication clarity and fostering effective interpersonal interactions.

REFERENCES

Bänziger, T., & Scherer, K. R. (2005). The role of intonation in emotional expressions. *Speech communication, 46*(3-4), 252-267.

Batalemba, I., Badrih, M., & Busri, H. (2024). *Kata "anjing" dalam Konten Youtube Asfmine: Kajian Prosodi dengan Pendekatan Fonetik Eksperimental Menggunakan Praat. Jurnal Onoma: Pendidikan, Bahasa, dan Sastra, 10*(1), 915-932.

Benus, S., Gravano, A., & Hirschberg, J. B. (2007). *Prosody, emotions, and…'whatever'.*

Chang, H. S., Lee, C. Y., Wang, X., Young, S. T., Li, C. H., & Chu, W. C. (2023). Emotional tones of voice affect the acoustics and perception of Mandarin tones. *Plos one, 18*(4), e0283635.

Colorado, D., Shih, C., & Kochanski, G. *Prosody and Prosodic Models*. In Proceedings of.

Crystal, David. (1989). *The Cambridge Encyclopedia of Language*. Cambridge University Press.

Denes and Pison. (1993). *The Speech Chain: The Physics and Biology of Spoken Language*. United States of America.

Departemen Pendidikan Nasional. (2008). *Kamus Besar Bahasa Indonesia Pusat Bahasa Edisi Keempat*. Gramedia Pustaka Utama.

Ekberg, M., Stavrinos, G., Andin, J., Stenfelt, S., & Dahlström, Ö. (2023). Acoustic features distinguishing emotions in Swedish speech. *Journal of Voice*. 1-10.

Ertürk, A., Gürses, E., & Kulak Kayıkcı, M. E. (2024). Sex-related differences in the perception and production of emotional prosody in adults. *Psychological Research, 88*(2), 449-457.

Fónagy, I., & Magdics, K. (1963). Emotional patterns in intonation and music. *STUF-Language Typology and Universals, 16*(1-4), 293-326.

Hammerschmidt, K., & Jürgens, U. (2007). Acoustical correlates of affective prosody. *Journal of voice, 21*(5), 531-540.

Heryono, H. (2019). Pengukuran pitch dan intensity diftong tertinggi menggunakan program PRAAT. *Jurnal Linguistik Komputasional, 2*(2), 47-55.

Heuven, V. J. V., & Zanten, E. V. (2007). Prosody in Indonesian languages: Concluding remarks.

    LOT Occasional Series, 9, 191-202.

Ladefoged, P. (1996). *Elements of Acoustic Phonetics*. The University of Chicago Press.

Larrouy-Maestri, P., Poeppel, D., & Pell, M. D. (2024). The Sound of Emotional Prosody: Nearly 3 Decades of Research and Future Directions. *Perspectives on Psychological Science*, 17456916231217722.

Mozziconacci, S. (2002). *Prosody and emotions. In Speech Prosody 2002*, International Conference.

Oh, C., Morris, R., Wang, X., & Raskin, M. S. (2023). Analysis of emotional prosody as a tool for differential diagnosis of cognitive impairments: a pilot research. *Frontiers in Psychology, 14*, 1129406.

Rochman, D., Diamond, G. M., & Amir, O. (2008). Unresolved anger and sadness: Identifying vocal acoustical correlates. *Journal of Counseling Psychology, 55*(4), 505.

Schwartz, R., & Pell, M. D. (2012). Emotional speech processing at the intersection of prosody and semantics. *PloS one, 7*(10), e47279.

Shields, S. A. (2013). Gender and emotion: What we think we know, what we need to know, and why it matters. *Psychology of Women Quarterly, 37*(4), 423-435.

Syarfina, T. (2009). *Ciri Akustik Bahasa Melayu Deli*. USU Press.

Syarfina, T. S., Zein, T. T., & Yusuf, M. (2024). The Exploration of Deli Malay Language Vowels: An Acoustic Phonetic Analysis. *Jurnal Arbitrer, 11*(1), 39-48.

Tannen, D. (2005). Conversational style: Analyzing talk among friends. Oxford University Press.

Xu, Y. (2019). *Prosody, tone, and intonation*. In The Routledge handbook of phonetics (pp. 314-356). Routledge.

Xu, Y. (2023). *Phonetics of Emotion*. In Oxford Research Encyclopedia of Linguistics.