

## THE ACOUSTIC FEATURES OF EMPHATICITY IN INDONESIAN

**Sugiyono**

Pusat Bahasa Indonesia, Jakarta

Email: sh\_sugiyono@indo.net.id

### ABSTRACT

*This study describes prosodic features which marked the emphatic meaning in Indonesian from Indonesian speaker whose mother tongue is, for instance, Javanese, Sundanese, Betawi, Bataknese, Bugisnese, Minangnese, Ma'anyanese, Balinese, and Mandailing. This study is conducted by applying experimental phonetic approach that is by selecting imperative sentences Tutup pintu! 'Close the door!' as a target sentence. The data are collected by asking a subject to commend and to repeat the same command to a child or someone who is of the same age of his child to close the door. All collected data are measured in terms of its frequency, intensity, and duration. Those three features are then analyzed statistically to find out the significance differences of acoustic feature difference on every level of utterance emphaticity. The results of this study shows that frequency, intensity, and duration are significant markers which differentiate emphaticity level. In terms of frequency, the higher the emphaticity level of the utterance is, the higher the frequency of the base pitch, final pitch, and pitch range of that utterance will be. From the sound intensity point of view, the emphaticity of utterance is marked by base intensity and intensity range. The utterance with high emphaticity level is marked by high base intensity and wider range intensity. From duration point of view, the utterance emphaticity is marked by the duration of all vowels. The height of emphaticity level is marked by the length of time needed by the vowels to be uttered. The emphaticity markers show meaningful correlation.*

**Keywords:** *acoustic phonetics; experimental phonetics; speech perception; and prosodie feature*

### ABSTRAK

*Penelitian ini memaparkan ciri-ciri prosa yang menandai makna empati dalam bahasa Indonesia yang misalnya diujarkan oleh penutur Indonesia yang bahasa ibunya adalah bahasa Jawa, Sunda, Betawi, Batak, Bugis, Minang, Ma'anya, Bali dan Mandalin. Penelitian ini menggunakan pendekatan fonetik eksperimen dengan memilih kalimat-kalimat imperatif 'Tutup Pintu.' Pengumpulan data dilakukan dengan meminta subjek untuk memberi perintah dan mengulangi perintah yang sama pada seorang anak atau seseorang yang seusia untuk menutup pintu. Semua data yang dikumpulkan diukur menyangkut frekuensinya, intensitas dan durasinya (lama). Kemudian, ketiga sifat ini dianalisis secara statistik untuk menemukan signifikansi perbedaan ciri akustik pada setiap tingkat empati ujaran. Hasil penelitian ini menunjukkan bahwa frekuensi, intensitas dan durasi merupakan*

penanda yang signifikan yang membedakan tingkat empati. Berkaitan dengan frekuensi, semakin tinggi tingkat empati ujaran, semakin akan tinggi frekuensi nada dasar, nada akhir dan tingkat nada ujaran. Dari aspek intensitas bunyi, empati ujaran ditandai oleh intensitas dasar dan tingkat nada. Ujaran dengan tingkat empati tinggi ditandai oleh intensitas dasar tinggi dan intensitas tingkat nada yang lebih lama. Dari aspek durasi, empati ujaran ditandai oleh durasi semua vokal. Tingginya tingkat empati ditandai oleh lamanya waktu yang diperlukan untuk mengucapkan vokal. Penanda empati menunjukkan korelasi yang signifikan.

**Kata Kunci:** fonetik akustik, fonetik eksperimen, empati ujaran.

## 1. Introduction

The relationship between stimulus and response is always interesting and has never been completely studied. Even, closely observed, every linguistics study is not other than a study of symbol which is related to stimulus and response of the verbal symbols. If stimulus and response could be well connected by speaker, then the symbol delivered by this speaker will be well accepted by the hearer. If stimulus could not be well responded, either verbally or behaviorally, then stimulus is often repeated until the expected response is achieved. In this paper, the repetition of stimulus in order to get response is called emphaticity.

Basically, emphaticity is an effort to provide emphatic content for stimulus in order to get certain response. *To emphasize* means *to show or state that something is particularly important or worth giving attention to* or *to make something more obvious*. If lexicon used in utterance is unchanging, emphaticity will be carried out by acoustic features, either by frequency, intensity, or duration. In some languages, emphaticity is carried out by the increase of the utterance frequency, intensity, or duration. In some other languages, this emphaticity may be also marked by the decrease of those three features. The basic question is that how does speaker carry out emphaticity within his utterance and how does hearer perceive emphaticity in someone's utterance. This study describes prosodic features

which marked the emphatic meaning in Indonesian from Indonesian speaker whose mother tongue is, for instance, Javanese, Sundanese, Betawi, Bataknese, Bugisnese, Minangnese, Ma'anyanese, Balinese, and Mandailing. This study is conducted by applying experimental phonetic approach that is by selecting imperative sentences *tutup pintu* as a target sentence.

## 2. Research Method

The data are collected by asking a subject to command and to repeat the same command to a child or someone who is of the same age of his child to close the door. The acoustic features of the utterances are measured by using the last version of Praat program (2007). All collected data are measured in terms of its frequency, intensity, and duration. Those three features are then analyzed statistically to find out the significance differences of acoustic feature difference on every level of utterance emphaticity. correlations between a certain acoustic feature and other acoustic features. This correlation is analyzed by using biphariat correlation with Pearson Product Moment (PPM) which shows correlation index (R) and its significance.

The subject is asked to recommend the child once more because the child has not closed the door yet. Then, again the subject is asked to command the child one more time since the child still has not closed the door yet.

The target sentence which is uttered by

the subject for the first time is the sentence without emphatic label or [-emphatic], the target sentence which is uttered for the second time is the utterance with a single emphatic content with [+emphatic] symbol, and the target sentence which is uttered for the last time is the utterance with double emphatic content with [++emphatic] symbol.

From 44 subjects (16 males and 28 females), there are 264 utterances collected. All utterance is analyzed into three steps: sound segmentation which aims to measure the duration of segmental sounds of the utterance; sound intensity measurement of the utterance; pitch stylization which aims to measure the significant pitch that forms the utterance.

### 3. Finding and Discussion

#### 3.1 Acoustic Features Measurement And Significances

The acoustic features of the utterances are measured by using the last version of Praat program (2007). All collected data are measured in terms of its frequency, intensity, and duration. Those three features are then analyzed statistically by using SPSS version 12 in order to find out the significance of the measurement differences, the focus is the significance of acoustic feature difference on every level of emphatic utterance.

The measurement of the utterance frequency is preceded by close copy analysis in order to find out distinctive pitch within the utterance. From this distinctive pitch is then measured (1) base pitch, the first distinctive pitch frequency within the utterance ; and (2) final pitch, the last distinctive pitch frequency within the utterance, (3) lower pitch, the lowest distinctive pitch within the utterance; (4) upper pitch, the highest distinctive pitch within the utterance. Those acoustic features are firstly measured by Hertz (Hz), but then converted into semitone (st) unit by using c natural frequency of piano (130.7749 Hz) as its reference. With semitone unit, the frequency analy-

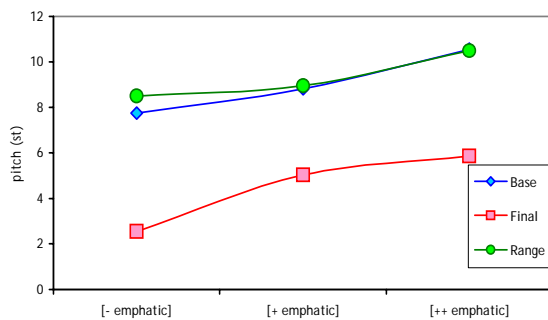
sis is comparable, not only from one pitch to another pitch within the utterance melodic structure, but also as pitch which is known by people, pitches in music. Based on the measurement attained, then the pitch range, pitch excursion, and pitch changes tendency will also be measured.

The utterance intensity is measured by using decibel (dB) unit. This measurement is done after simplifying or removing intensity points which are not distinctive within the utterance. Based on utterance intensity, the things measured are (1) base intensity, the first distinctive intensity within the utterance; (2) final intensity, the last distinctive intensity within the utterance; (3) lower intensity, the lowest distinctive intensity within the utterance; and (4) upper intensity, the highest distinctive intensity within the utterance. Based on the measurement attained, then the utterance intensity range and the intensity changes tendency will also be measured.

Duration is firstly measured by second. However, in order to simplify the analysis, this unit is then converted into millisecond (ms). The duration measurement is preceded by segmenting the utterance into singles utterances. The next duration measurement is focused on vowel duration since vowel is the core of syllable. Therefore, for the target utterance *tutup pintu* is identified based on four vowels duration, they are /u/ as the first vowel, /u/ as the second vowel, /i/ as the third vowel, and /u/ as the fourth vowel.

#### 3.2 Frequency Significance

Frequency in an utterance, either in base pitch, final pitch, lower pitch, upper pitch, or pitch range becomes a very significant marker of utterance emphaticity ( $P < 0.01$ ). Generally, the pattern of utterance frequency difference is that [-emphatic] is lower than the [+emphatic] utterance frequency and those two frequencies are lower than [++emphatic] utterance frequency.

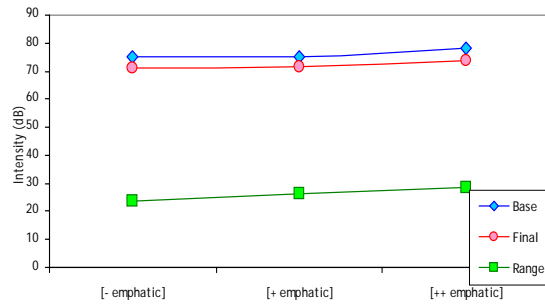


Picture 1: Base pitch, Final Pitch, and Pitch Range based on emphaticity level

Base pitch of the utterance without emphatic is around 7.75 st (221.29 Hz). On [+ emphatic], this base pitch increases less than 1.1 st to become 8.81 (227.72 Hz) which then on [++emphatic] increases less than 1.3 st to become 10.55 st (251.23 Hz). On every emphaticity change, it increases pitch from 1.1 st to 1.3 st with a meaningful significance. A similar significance difference is also found in final pitch comparison. The final pitch of the utterance [- emphatic] is around 2.55 st (157.84 Hz) which becomes 5.03 st (187.20 Hz) on [+ emphatic] utterance and increases into 5.87 st (199.64 Hz) on [++ emphatic] utterance.

This is shown by graphic on Picture 1. On that picture, it can be seen a significant difference between base pitch and final pitch. On every emphaticity level, final pitch is much lower than base pitch, the mean point is 9.04 st ratio 4.48 st. Both pitches have strong and significant positive correlation. In this case, an evidence found is that the higher the base pitch is, the higher the final pitch is ( $r=0.667$ ;  $p<0.01$ ).

The comparison between the highest pitch range and the lowest pitch range within utterance shows that emphaticity has significant impact to the pitch range ( $p<0.01$ ). The pitch range before obtaining emphatic stress is shorter than the pitch range of [+emphatic] utterance and both pitch ranges are much shorter than the pitch range of [++emphatic]



Picture 2: Base, Final, and Intensity Range based on emphaticity Level

utterance. The measurement shows that the pitch range of the utterance without emphaticity is 8.51 st, the pitch range of the utterance with emphaticity is 8.96 st, and the pitch range of the utterance with double emphaticity is 10.49 st. Although the average difference of those three pitch ranges is around 0.99 st, this difference has a very high significance.

Unlike pitch range, the trend of pitch changes within utterance shows different tendency. Graphic on Picture (2) shows that the trend of pitch change of the utterance without emphaticity is (5.20 st) bigger than the trend of pitch change of the utterance with emphaticity, either with single emphaticity (3.78 st) or with double emphaticity (4.68 st). Besides, it is also interesting to find out that the trend of utterance with double emphaticity is slightly higher than the trend of utterance with single emphaticity though both do not show significant differences. Based on *post hoc* analysis, the difference pattern of this trend follows the pattern that the trend of utterance without emphaticity is higher than the trend of utterance with emphaticity. If this trend is related to the base pitch, it seems that the higher the base pitch is, the trend formed will not be very big. The trend of utterance with emphaticity is smaller because its base pitch is higher.

### 3.3 Intensity Significance

Role of intensity as emphaticity marker is not as significant as frequency. Intensity is sig-

nificantly found only on the initial (base) intensity realization, on the highest intensity, and on the intensity range ( $P < 0.01$ ). The final intensity realization ( $p=0.20$ ), the lowest intensity ( $p=0.995$ ), and the intensity trend ( $p=0.74$ ) do not show meaningful differences.

The base intensities within [-emphatic] and [+emphatic] utterances tend to be the same ( $p=0.99$ ), but both of them are much lower than the base intensity of [++emphatic] utterance ( $p=0.04$ ). From this point of view, generally the speakers will increase their base intensity for about 3.2 dB out of the average rate 75.1 dB in order to express their highest empathy. In this case, the base intensity of males tends to be 1.2 dB lower than females.

The final intensity based on emphaticity level does not show any significant differences ( $p=0.20$ ). If it is compared to the base intensity, the final intensity is much lower with more or less 4 dB ( $p < 0.01$ ) difference. It can be seen that there is a parallelization between final intensity and final pitch, that is either final intensity or final pitch is always lower than its base point. On the previous studies (Sugiyono, 2003) the negative excursion of the final pitch marks modus, the final intensity which is lower than its base intensity may mark a certain modus. The comparison between base pitch and final pitch as well as base intensity and final intensity will be analyzed by using acoustic feature trend analysis.

The average rate of the utterance intensity range is around 26.3 dB. This intensity range becomes a very significant difference of the utterance emphaticity ( $p=0.01$ ). The intensity range of [-emphatic] utterance is around 23.8 dB, the intensity range of [+emphatic] utterance is around 26.48 dB, but the intensity range of [++emphatic] utterance is around 28.7 dB. Post hoc analysis shows that the intensity ranges of [-emphatic] and [+emphatic] utterances have similar tendencies, although both ranges are smaller than the intensity range of [++emphatic] utterance.

From the above analysis, it can be seen that the sound intensity has not changed much at the first command repetition. It seems that the repetition is done with consideration that the hearer does not listen to or does not pay attention to the utterance well. The stress of emotion appears at the second repetition. Emotion implied because it seems that the speaker believes that the hearer has listened and understood the command well, but he/she does not want to respond the command intentionally. Therefore, the base intensity and the intensity range of [++emphatic] utterance are significantly different from the base intensity and the intensity range of the other utterance emphaticity levels. However, difference intensity feature is not valid for male utterance. Male speakers do not apply intensity feature as a media of emphaticity within their utterances. With or without emphatic utterance are both carried out by the same level of emphaticity, the base intensity is about 75.3 dB and the intensity range is about 22.1 dB.

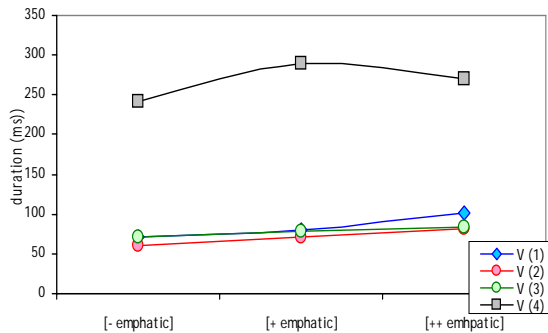
The utterance intensity trend does not show a significant difference ( $p=0.74$ ). The average rate of intensity trend is 4.0 dB, the intensity trend of [-emphatic] utterance is 4.0 dB, the intensity trend of [+emphatic] utterance is 3.4 dB, and the intensity trend of [++emphatic] utterance is 4.6 dB.

### 3.4 Duration Significance

A significant difference can be seen on the duration of the last /u/ vowel of the word *pintu* which is located at the final utterance. The next picture shows that the duration of the first vowel, /u/ at *tutup* and /i/ at the first syllable of *pintu* is more or less 77.6 ms, while the duration of the last vowel is about 266.9 ms. In relation with emphaticity, the vowel length difference could not be claimed yet as emphaticity marker, since the previous study (Sugiyono, 2003) found that there is always an extension on the last syllable of an utterance which functions as a modifier.



In relation with emphaticity, vowel duration has a huge significance difference. Generally, it can be said that vowel or syllable is the core of a syllable wherever position it takes and it can also be an emphaticity marker in Indonesian ( $p < 0.01$ ).

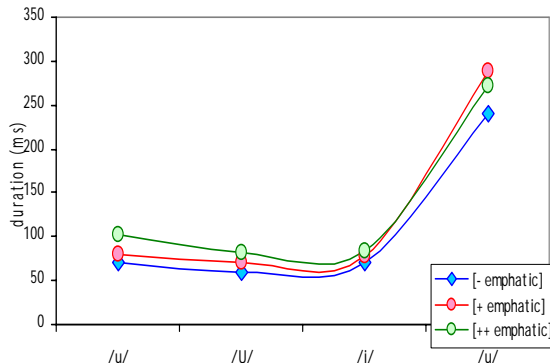


Picture 3: Vowel Duration within Utterance based on emphaticity Level

The duration of vowel /u/ at the first syllable of *tutup* on [-emphatic] utterance is around 70.6 ms which then increase to 80.6 ms on [+emphatic] utterance, and claim up again to 101.7 ms on [++emphatic] utterance. /u/ vowel at the second syllable of *tutup* by most of subjects is carried out as [U] whose duration is around 60.4 ms on [-emphatic utterance], which then increase to 70.3 ms on [+emphatic] utterance, which then claim up to 81.9 ms on [++emphatic] utterance. The duration of /i/ vowel at *pintu* is around 71.3 ms on [-emphatic] utterance, which becomes 78.7 ms on [+emphatic] utterance, which then increase to 83.0 ms on [++emphatic] utterance. The initial duration of /u/ vowel at *pintu* is 240.9 ms on [-emphatic] utterance, which becomes 288.8 ms on [+emphatic] utterance, but then slightly decrease to 270.9 ms on [++emphatic] utterance.

The significant difference pattern of vowel duration which is summarized from the post hoc analysis strengthens the difference pattern of the previous acoustic features that is vowel duration of [-emphatic] utterance tends to be

similar to vowel duration of [+emphatic] utterance. Both durations are smaller than vowel duration of [++emphatic] utterance.



Picture 4: The Comparison of Vowel Duration Linearly within Utterance

Syntagmatically, the vowel duration within utterance shows similar tendency or tends to be shorter at the first three syllables, which then extend at the last syllable on all levels of emphaticity. As it can be seen at Picture (4) above, generally, the duration of [-emphatic] utterance tends to be shorter than the duration of [+emphatic] utterance, and both durations are shorter than the duration of [++emphatic] utterance. A significant exception is that the duration of /u/ vowel at the last [+emphatic] utterance is longer or even also longer than the duration of /u/ vowel on [++emphatic] utterance. This fact implies that the extension of vowel at the last syllable within utterance is always carried out, but the quantity of the duration length does not marked the emphaticity difference individually.

### 3.5 Acoustic Features Correlation

The next question arises is that is there are any correlations between a certain acoustic feature realization and other acoustic features. This question will convey a further analysis that is an interdependency correlation between one acoustic feature and others. The analysis of biphariat correlation with Pearson

Table 1: Correlation of Significant Acoustic Marker within Utterance

\*\* Significant correlation on 0,01 level (2-peak).  
 \* Significant correlation on 0,05 level (2-peak).

Product Moment (PPM) shows that correlation index (R) and its significance is completely features, such as base intensity, intensity range, and all vowel duration. A similar correlation is

	Base pitch	Final Pitch	Pitch Range	Base intensity	Intensity Range	Vowel 1	Vowel 2	Vowel 3	Vowel 4
Base pitch		0,657**	0,123*	0,232**	0,235**	0,354**	0,291**	0,381**	0,162**
Final Pitch	0,657**		-0,273**	0,050	0,230**	0,126*	0,240**	0,311**	0,193**
Pitch Range	0,123*	-0,273**		0,124*	0,181**	0,353**	0,124*	0,035	0,167**
Base intensity	0,232**	0,050	0,124*		0,197**	0,260**	0,136*	0,215**	-0,313**
Intensity Range	0,235**	0,230**	0,181**	0,197**		0,365**	0,314**	0,168**	0,176**
Vowel 1	0,354**	0,126*	0,353**	0,260**	0,365**		0,390**	0,318**	0,229**
Vowel 2	0,291**	0,240**	0,124*	0,136*	0,314**	0,390**		0,214**	0,234**
Vowel 3	0,381**	0,311**	0,035	0,215**	0,168**	0,318**	0,214**		0,112
Vowel 4	0,162**	0,193**	0,167**	-0,313**	0,176**	0,229**	0,234**	0,112	

the pitch range will be. However, the correlation between final pitch and pitch range is a negative correlation. The higher the final pitch is, the shorter the pitch range will be. On the contrary, the lower the final pitch is, the wider the pitch range will be.

Base pitch, final pitch, and pitch range are positively correlated with other acoustic

features, such as base intensity, intensity range, and all vowel duration. A similar correlation is shown by the height of the voice duration. Therefore, the importance of frequency, intensity, and duration are significant markers which differentiate emphaticity level. Specifically, it can be concluded as follows.

1. Base pitch, final pitch, and pitch range are significant markers of emphaticity. The higher the emphaticity level of the utterance is, the higher the frequency of the base pitch, final pitch, and pitch range of that utterance

- will be.
2. From the sound intensity point of view, the emphaticity of utterance is marked by base intensity and intensity range. The utterance with high emphaticity level is marked by high base intensity and wider range intensity.
  3. From duration point of view, the utterance emphaticity is marked by the duration of all vowels. The height of emphaticity level is marked by the length of time needed by the vowels to be uttered.
  4. The emphaticity markers show meaningful correlation.

### BIBLIOGRAPHY

- Ebing, Ewald. 1997. *Form and Function of Pitch Movements in Indonesian*. Leiden: Research School CNWS.
- Fujimura, Osamu dan Donna Erickson. 1999. "Acoustics Phonetics". Dalam Hardcastle, William J. and John Laver. 1999. *The Handbook of Phonetics Sciences*. Oxford: Basil Blackwell.
- Mozziconacci, S.J.L. 1998. *Speech in Variability and Emotion: Production and Perception*. Disertasi Eindhoven University of Technology.
- Odé, Cecilia and Vincent J. van Heuven (eds.). 1994. *Experimental Studies of Indonesian Prosody (Semaian 9)*. Leiden: Rijksuniversiteit te Leiden.
- Sugiyono. 2002. "Pitch Marker of Interrogativity in Indonesian: Psychoacoustics Approach". Paper of *International Symposium of Linguistics and Speech-Hearing Sciences 2002*, Kuala Lumpur, 21–22 October 2002.
- Sugiyono. 2003. *Pemarkah Prosodik Kontras Deklaratif dan Interogatif Bahasa Melayu Kutai: Kajian Fonetik Eksperimental dan Psikoakustik*. Thesis.
- Van Heuven, Vincent J. van dan Ellen van Zanten. 1994. "Effects of Substrate Language on the Localization and Perceptual Evaluation of Pitch Movement in Indonesian". Dalam Odé, Cecilia dan Wim Stokhof (eds.). 1997:63—80.
- Van Zanten, Ellen dan Vincent J. van Heuven. 1994. "Effect of Word Length and Substrate Language on the Temporal Organisation of Words in Indonesian". Dalam Odé, Cecilia dan Wim Stokhof (eds.). 1997:201—216.