Melodic and Temporal Structure of Indonesian Sentences of Sitcom "International Class" Actors: Prosodic Study with Experimental Phonetics Approach

Tri Sulistyaningtyas, Yani Suryani, Dana Waskita, Linda Handayani Sukaemi, Ferry Fauzi Hermawan

Abstract—The enthusiasm of foreigners studying the Indonesian language by Foreign Speakers (BIPA) was documented in a sitcom "International Class". Tone and stress when they speak the Indonesian language is unique and different from Indonesian pronunciation. By using the Praat program, this research aims to describe prosodic Indonesian language which is spoken by "International Class" actors consisting of Abbas from Nigeria, Lee from Korea, and Kotaro from Japan. Data for the research are taken from the video sitcom "International Class" that aired on Indonesian television. The results of this study revealed that pitch movement that arises when pronouncing Indonesian sentences was up and down gradually, there is also a rise and fall sharply. In terms of stress, respondents tend to contain a lot of stress when pronouncing Indonesian sentences. Meanwhile, in terms of temporal structure, the duration pronouncing Indonesian sentences tends to be longer than that of Indonesian speakers.

Keywords—Melodic structure, temporal structure, prosody, experimental phonetics, international class.

I. INTRODUCTION

As part of the culture in which humans play an important role, language also takes part in the role of humans because it serves as a communication tool continuing to grow in accordance with the development of human civilization itself.

The Indonesian language is now not only learned by Indonesian people. Evidently, the Indonesian language is now learned in 179 foreign universities located in 45 countries around the world. Of course it is a source of pride for us as Indonesian native speakers. In Indonesia, Indonesian teaching activities for Foreign Speakers (BIPA) have been conducted since the 1990s. This was done with the consideration that in the era of globalization, as the Indonesian language becomes increasingly important and the potential in the international arena increases [13].

The potential of the Indonesian language is supported by the geographical position of Indonesia, in which people of different nations are interacted with the indigenous people. Thus, the Indonesian language is also expected to be a bridge for other nations to increase their understanding of the people and culture of Indonesia. That fact has led many foreigners to be interested and attracted to learn the Indonesian language as a means to achieve various objectives, such as politics, economy, trade, education, art, culture, and travel.

The relatively large number of people who learn Indonesian impressed one of television station in Indonesia to broadcast a situation-comedy (sitcom) program named "International Class". This sitcom has been aired on Net TV since 2015. The International Class tells about people from different countries who are studying Indonesian in Jakarta. Among them are Abbas from Nigeria, Kotaro from Japan, and Lee from Korea. In terms of melody and temporal structure, the sentences uttered by Abbas, Kotaro, and Lee are interesting to be studied because their pronunciations are still affected by their native language. Therefore, this paper will explain the melodic structure and temporal structure uttered by Abbas, Kotaro, and Lee in the "International Class" sitcom.

II. PHONETIC EXPERIMENTAL

Experimental phonetics is a relatively young study. This study was recently developed in the 1940s when the spectrograph was discovered. Although it is clear that the object is a symptom of acoustic phonetics concrete, investigation of these objects originally could only be done by relying on the phonetics expert impression. Identification and their analysis of the object sounds were analyzed entirely based on the ability of the senses of hearing, sight, and awareness of organ activity of oneself when a sound is uttered [1].

According to Hayward [1], experimental phonetics includes a variety of research on speech using this instrument. That is, the instrument used to visualize some aspects of the speech that occurs. For example, if we use a tape recorder to repeat listening to a speech, it cannot be called a part of the experimental phonetics. But, if the tape recorder is connected to a computer and used to perform acoustic analysis of the speech, then that activity can be categorized as an experimental research. Therefore, experimental phonetics is often called instrumental phonetics.
Currently, many computer programs have been developed for the measurement of acoustic features such as Computerized Speech Research Environment (CRSE) and Praat. Thus, the study of phonetics in the impressionistic period, which was very limited by the ability and senses of an expert in phonetics, is now very open, due to the use of advanced technologies that are commonly used in this experimental approach. In this study, the authors use the Praat program to analyze the speech sounds pronounced by the actors of "International Class" sitcom.

III. PROSODY

Linguists have long discussed the theory of prosody. Generally, the end of the 1930s was regarded as a starting point of the development of the study of prosody. At that time, John R. Firth (1890-1960) from SOAS, University of London, powered the development of a new theory that was then known as phonological prosody [2]. Firth flipped the perspective of phonological study when it grew rapidly which was marked by the emergence of phonemic principle of Swadesh [10] and the pricing models of phonemes of Trubetzkoy [11] to see the speech from a different side to that which linguists had done at the time. If the classic phonology maps the object into a row of phonemes, prosodic phonology describes the object rather than as a separate segment, and instead as segments that make up a structure. A sound segmental separate unit is called phonemics, while relations systems or structures that weave one segment to other segments is called prosody.

Discussion on the prosodic aspects of Indonesian language has been conducted since the 1940s. According to Amran [9], even the recorded writings of Marsden [12] have initiated discussions about the pressure of the Indonesian word in his book A Grammar of the Malayan Language.

In linguistics, prosody is the rhythm, stress and intonation of an utterance. Prosody allows mirroring the various features of the speaker or the speech, the emotional state of the speaker, figure of speech, their irony or sarcasm, emphasis, contrast, focus, or other language elements that may be encoded by grammar or vocabulary selection. [1, p. 235] states that "Suprasegmental and prosodic features are general terms used to refer to patterned variation in pitch, force (stress), and duration – in other words, to the properties which spoken language shares with music".

According to [3], the components of phonetics in prosody are (i) variations in pitch that is determined by the repetition rate of vibration of the vocal cords, (ii) variations in loudness is determined by the intensity of the sound and the balance of spectral due to differences in the location of vowels, (iii) variations in the quality or timbre for precision articulator, and (iv) the timing variations due to fast and slow voice. Linguistic differences that depend on these parameters, among others, are tone, intonation, accent, stress, and rhythm.

IV. THE FEATURES OF SPEECH PROSODY

In addition to lexical items which are then assembled to form a lexical structure, a speech also has other elements that characterize the lexical structure i.e. how the lexical structure should be spoken. From the point of phonetics, the first element is called the segmental elements. The elements that characterize the segmental element are called suprasegmental or prosodic elements. Segmental element associates with single sounds, while the suprasegmental elements relate to other acoustic features sounds that accompany the single sound. Each has a segmental sound duration, frequency, duration and intensity of the so-called intrinsic, intrinsic tone, and intrinsic intensity. Every vowel has acoustic features called formants -usually a distinguished formant (F1) and formant two (F2) -which is determined by the shape and size of the oral cavity, pharynx cavity, and in some cases also the nasal cavity [1]. The suprasegmental element is the other side of the acoustic features that complement the segmental elements. Between the segmental and suprasegmental characteristics that are distinguished by the fact that the suprasegmental characteristics are determined by a comparison of the characteristics of the segment with the characteristics of other segments in sequence segments, while the segmental characteristics can be set without seeing the connection segment to other segments.

Collier [4] says that the characteristics of prosody have the function of demarcation, i.e. as modifiers in a speech. As modifiers among sentences, prosody marks lexical cohesion in a unit of information that is highlighted among other units. In this case, prosody borders serve as stresses so that the meanings of utterances become more transparent to the listener. This partition is called Perceptual Boundary Strength (PBS). Prosody can also be used for the inter-unit showing limited information, such as modifiers between words or phrases that can be understood by the listener. At the level of discourse, these modifiers have the same position as other modifiers and at a higher level of discourse structure, prosody becomes the modifiers, for example, for a change of subject in a monologue or marker turn-taking in a conversation. Van Heuven [5] classifies the functions of prosody characteristics of three kinds, namely (1) to give modifiers domain or part of speech (e.g., paragraphs, sentences, or phrases); (2) to give specific properties to the information displayed in the domain (such as a statement or question); and (3) to include certain constituents.

In discussions of utterance prosody characteristics, the melodic structure is separated from the temporal structure or rhythm [5]-[7]. The term melodic structure is also known as the intonation used to refer to a set of rules to characterize variations in tone lining a speech in a particular language, while a rhythm or temporal structure is a set of rules that determine the duration of the speech pattern.

V. MELODIC STRUCTURE

At the level of phonology, intonation is defined as the use of suprasegmental phonetic features to bring a pragmatic
significance at the level of the sentence or the post-lexical level (post-lexical or sentence-level) in the well-structured form linguistically (Ladd in [8]). Suprasegmental phonetic features that are intended to limit it, is the tone (pitch) with other functional perceptual characteristics in which distinct sounds form the sound frequency. At the level of phonetics, intonation is defined as a series of tones-usually one note per syllable-that characterize a sentence in the speech. In F0 dimensions, intonation is related not only to how high F0 of a segmental element, but also to how long and where a tone is played. Moreover, every tone in a speech is not seen in isolation, but rather a formation of a structure called a melodic structure, better known by the term intonation.

Acoustic manifestation of intonation is a complex sound which included a fundamental frequency and the frequency of harmony which is measured in Hertz or semitones. A semitone is a unit of the height difference between F0 as different tones of a key in the piano one to another. Ideally, intonation load variations of the whole tones of each syllable in the speech, but are considered to be inefficient because of the limitation threshold of perception. For simplification it is usually done, for example, just by describing the pitch changes in a particular domain. Units that are then called pitch movement. Pitch movement is then seen as its acoustic characteristics, namely the direction of the change of tone: up or down, tone size: large or small, the degree of change: a sharp or gradual, and its position within the syllable.

To describe the pitch, there is a distance measurement. The tone in the speech of a woman, who is definitely higher than the tone of the speech of a man, becomes a problem if that is the distance measured tone. For this reason, the measurement pitch in units of semitones (ST) is more appropriate than in Hertz (Hz) [6]. Thus, a high F0 originally in Hz is converted into units of semitones by applying the following formula:

\[
F_{ST} = \left( \frac{12}{\text{LOG}(2)} \right) \text{LOG}(F_{Hz}/F_{Ref})
\]

F (Hz) is the fundamental frequency measurements that is stylization and F (Ref) is the fundamental frequency is used as references. The Praat program can convert ST Hz to 100Hz easily with referents. However, this method is not selected. This study used F (Ref) of 130.7749 Hz, which is the fundamental frequency of the tone C in music so that the measurement results can then be compared with the tones in musical scales.

VI. TEMPORAL STRUCTURE

Van Heuven [5] gave limitation temporal structure as a set of rules that determine the duration of speech sounds and pauses in speech that were uttered in the language. Meanwhile, Nooteboom [6] referred to it as the rhythm of speech interpreted as the duration aspects of the speech-controlled by speakers (speakers controlled the aspect of speech timing).

Rhythm patterns relate to the provision of pressure parts of speech emphasized. Theoretically, this emphasis can be done with the tone of the media aspect, the aspect of the duration and intensity aspects. In both of the above limitations, the intensity aspect seems to be set aside so that rhythm is regarded as the temporal structure itself. The limit that is in accordance with the assessment given, it has been demonstrated that the pressure in the Indonesian language is noton-temporal [9], i.e. it does not affect the intensity of the pressure. Thus, besides the temporal organization of each syllable in a speech, the provision of a pause in the speech needs to be considered.

VII. METHOD

The method used in this research is a descriptive analysis method. This method provides an overview of data systematically and accurately. The data are collected by recording selected data objects as corpus of data because the data source of the research is in the form of spoken data taken from the Indonesian language utterances of actors in sitcom "International Class" when they were speaking. Furthermore, all data obtained were recorded on a data card, sorted and classified based on the sentences spoken. The data were analyzed using Prata program to assess the tone, stress, intonation, and pauses of the Indonesian sentences gathered. Sources of data were taken from a video of the sitcom "International Class" Season 2 Episode 1 Part 1. The data were analyzed in several stages. The first phase is to move the data recording in digital format sound waves, then utterances in accordance with what is desired to be analyzed was selected. Furthermore, the data segmentation is data which have been segregated into segments per syllable. After completion of data processing, it is followed by the measurement of the intensity, duration, and frequency. Measurement of acoustic features is done by adapting the theory of Instituut voor Perceptie Onderzoek (IPO)

VIII. DISCUSSION

A. Melodic Structure

In this study, we analyzed the sentences pronounced by Abbas from Nigeria, Kotaro from Japan, and Lee from South Korea. The sentences analyzed include declarative sentences, interrogative sentences, and imperative sentences.

1. Pronunciation Declarative Sentences

From the impressions that appear in "International Class" Session 2, is a dialogue between Lee and Abbas who meet the new students joining the International Class. The sentence pronounced by Lee is, "Kemarin Bu Rika dan Pak Budi sudah cerita (Yesterday, Mrs Rika and Pak Budi told a story)". Meanwhile, the declarative sentence uttered by Abbas is, "Saya ingin tahu (I want to know)." At the same time, Kotaro is daydreaming and remembering the story of his love with Ling-ling that ended. Kotaro says, "Hidup Kotaro pahit seperti the dalam gelas ini (Kotaro’s life is bitter like the tea in this glass)." The tone of speech in their pronunciation can be seen in Figs. 1-3.

When saying the phrase "Kemarin Bu Rika dan Pak Budi sudah cerita (Yesterday, Mrs Rika and Pak Budi told a story)", there is an additional voice / hoooh / in the word /
Rika / and / at /. The voice that emerged was apparently due to the habit of Lee’s speaking Korean, so that it appears as the sound / hoooh / . Judging from the melodic structure, there is a lot of emphasis in speech and tone up and down drastically. The sentence pronounced by Lee opened with a tone of 85.3 Hz or -7 ST and then rose sharply to 366.3 Hz tone or 18 ST. The tone continues to decline drastically and closed with the syllable / ta / word / story / tone 351.5 Hz or 17 ST. The tone on this sentence looks unique because in Indonesian intonation, the utterances of declarative sentences tend to fall when the tone is lower than the earlier part. However, in this sentence the alternative was the case.

Fig. 1 Contour pronunciation of declarative sentences by Lee

The sentence uttered by Abbas is a simple sentence that is not difficult to pronounce. This is evident from the tone that tends to rise and fall gradually. The first tone in this sentence is 147.1 Hz or 2 ST, and then drops to 135.9 Hz or 1 ST. However, the syllable / ya / in the word / i /, ringing unexpectedly rose to 252.6 Hz or 11 ST; after the fall and rise, gradually to end up at tone of 246.8 Hz or 11 ST. The tone on this declarative sentence was closed with a tone higher than the earlier tone.

Fig. 2 Contour pronunciation of declarative sentences by Abbas

The sentence spoken by Kotaro is full of pressure and fluctuates dramatically. The sentence starts with a 112.8 Hz tone or -3 ST and then down to 93.72 Hz or -6 ST. However, when speaking the syllables / dup / in the word / life /, the tone goes up sharply to 389.9 Hz or 19 ST and the falls again sharply to 114.1 Hz or -2 ST. The sentence spoken by Kotaro also rises up and down gradually and then closes on a high note at 300.6 Hz or 14 ST. Just like his friends, the tone of the end of a declarative sentence uttered by Kotaro is higher than the original tone.

2. Pronunciation of Interrogative Sentences

Fig. 3 Contour pronunciation of declarative sentences by Kotaro

A conversation between Lee, Abbas, and the new students in the sitcom International Class, is in interrogative sentence that can be analyzed as "Teman-teman, kalian murid baru kelasnya Pak Budi, kan? (Friends, you are a new student in the class Pak Budi, right?)" said Lee, and "Mereka berbicara apa? (They speak of what?)" said Abbas. Meanwhile, interrogative sentences are spoken by Kotaro when he remembers his failed love with Ling-ling, and he asks Ling-ling, "Kamu ngomong apa? (You say what?)" When pronouncing this sentence, Lee was able to correctly say the words, even though he looks as if he is struggling to say them. This sentence starts with the tone of 124 Hz or -1 ST, and rises...
up and down gradually until at the end to reach 185.5 Hz or 6 ST. It is seen that the end of the tone is higher than the early tone when the tone of the end of interrogative sentences in Indonesian is lower than the initial tone. The tone on this sentence begins at 102.9 Hz or -4 ST and then rises sharply to 201.7 Hz tone or 8 ST, and continues to rise again to reach the highest pitch of 481.9 Hz or 23 ST with the syllables / ra / word / bicara /. The tone of the speech goes up and then down to the end of the sentence at a 213.8 Hz tone or 9 ST.

Sentences spoken by Kotaro who is angry with Ling-ling for breaking his heart; he says, "Saya malu! (I am ashamed!)". The tone of each sentence of speech is shown in the Figs. 7-9.

Fig. 5 Contour pronunciation of interrogative sentences by Abbas

Fig. 6 Contour pronunciation of interrogative sentences by Kotaro

Fig. 7 Speech tone of imperative sentence by Lee

Fig. 8 Abbas’s speech tone in imperative sentence

3. Pronunciation of Imperative Sentences

After seeing a beautiful woman who has just joined the International Class, Abbas immediately shouts, "Kita akan ada murid baru! (We will have a new student!)". Lee addresses the matter directly with the response, "Sudah tahu, Abbas! (I know, Abbas)". Meanwhile, an intriguing sentence

Sentences spoken by Lee begin at 131.1 Hz tone or 0 ST and then rise sharply to 299 Hz tone or the 14 ST, with the sharpest rise recorded when he pronounces the word / idea /. This sentence ends at 286.6 Hz or 14 ST. The tone of speech for the end tone is higher than the initial tone, which is in keeping with the pattern of tone of speech for Indonesian imperative sentences. This sentence begins with a 209.8 Hz tone or 8 ST then rises sharply and causes a lot of pressure, especially when uttered / new /. The sentence ends at 342.5 tones or 17 Hz. From the tone of this speech, it can be seen that the end tone is higher than the initial tone, and that this speech tone is in accordance with Indonesian imperative sentence intonation.

Kotaro speaks this sentence out loud and places a lot of stress on the word / malu/ (shame). The sentence begins at 190.6 Hz tone or 7th ST and continues rising and falling and ends at 257.7 Hz tone or 12 ST. As an intriguing sentence, the speech tone of this sentence is in accordance with the
intonation for Indonesian imperative sentence in which the end tone is higher than the earlier tone.

Fig. 9 Speech tone of Kotaro’s imperative sentence

B. Temporal Structure

The temporal characteristics of the process description, not handled by the IPO; this process is done in a much simpler way. The process begins by segmenting speech on the segments forming the speech with the domain word.

In the declarative sentence, Lee uttered the sentence, “Kemarin Bu Rika sama Pak Budi sudah cerita,” lasting a duration of 5.01 seconds. However, when pronouncing the sentence, there is a noise / hooh / being a respite from the word / rika / and /sama / over 459.8 millisecond. In addition, the pronunciation of the word /cerita/ is longer than in similar sentences spoken by a native Indonesian. Lee spoke the word /cerita/ lasting for the duration of 1.6 seconds, as shown in Fig. 10. Meanwhile, Abbas utter a three-word declarative sentence, “Saya ingin tahu!” lasting for 1.2 seconds, as shown in Fig. 11.

The declarative sentence of eight words was uttered by Kotaro lasting for 5.8 seconds. This can be seen in Fig. 12. Meanwhile, in the pronunciation of interrogative sentences, the third of respondent recited with varying durations. However, in this study, he says the end of sentences with a longer pronunciation. This can be seen in Figs. 13-15.

In Fig. 15, the visible duration of the interrogative sentences spoken by Kotaro contain a lag of 0.59 seconds between words / kamu/ (you) and /bicara/ (talk); this pause occurs because the speaker is thinking for the right words to say.

In the imperative sentence pronunciation, the third respondent says the imperative sentence with varying durations. Lee pronounces an imperative sentence lasting a duration of 2 seconds, Abbas for a duration of 1.7 seconds, and Kotaro for 0.6 seconds, as seen in Figs. 16-18, respectively.

Fig. 16 shows Lee’s pronunciation of the word / Abase /, lasting for 1.3 seconds. The pronunciation of this word is much longer than for /sudah/ (already) and /tahu/ (know), each spoken for 0.7 seconds.

When reciting the imperative sentences, the last word spoken by Abbas lasted longer than the other words in the sentence. The word / baru / (new) in the sentence lasted for 0.6 seconds from the overall sentence that was spoken in 1.7 seconds.

Fig. 18 shows that Kotaro’s pronunciation is comparable to that of the speech pattern of a native Indonesian with a sentence duration of 0.67 seconds. The word / malu / (shame) is pronounced with a longer duration than the word / saya / (I), which lasted 0.4 seconds, despite both words being composed of four letters.

Fig. 10 Duration of the pronunciation of Lee’s declarative sentence
Fig. 11 Duration of the pronunciation of Abbas’s declarative sentence

Fig. 12 Duration of the pronunciation of Kotaro’s declarative sentence

Fig. 13 Duration of the pronunciation of Lee’s interrogative sentence
Fig. 14 Duration of the pronunciation of Abbas's interrogative sentence

Fig. 15 Duration of the pronunciation of Kotaro's interrogative sentence

Fig. 16 Duration of the pronunciation of Lee's imperative sentence
VIII. CONCLUSION

From the above discussion it can be concluded as follows:

1) In terms of melodic structure, when pronouncing declarative sentences, interrogative sentences, and imperative sentences, the participants of "International Class" exerted unnecessary pressure on sound, with the tones rising and falling dramatically. In addition, all the speakers inserted a pause mid-sentence when speaking in Indonesian, as they attempted to find the appropriate words to match their intention. The speech was also affected by the respective mother tongue of the speaker. When pronouncing words ending in a consonant, Lee from Korea always added tagged a vowel [e] to the end of Indonesian words. Moreover, Kotaro from Japan placed a lot of pressure on intonation when pronouncing sentences in Indonesian.

2) In terms of temporal structure, the speakers' pronunciation of declarative, interrogative, and imperative sentences lasted longer than that of the Indonesian native speakers.

REFERENCES

Tri Sulistyaningtyas is a lecturer at the Faculty of Fine Arts and Design of Institut Teknologi Bandung, Indonesia. She graduated from doctoral program Faculty of Cultural Sciences, Padjadjaran University in 2014. Her research of interest is sosiopragmatics and culture. Besides as lecturer and researcher, she is also a chief editor of Jurnal Sosioteknologi ITB.

Yani Suryani is a lecturer at the Faculty of Fine Arts and Design of Institut Teknologi Bandung, Indonesia. She is graduated from Faculty of Cultural Sciences, Padjadjaran University in 2013. Her research of interest is sociophonetics and culture. Besides as lecturer and researcher, she is also an editor of Jurnal Sosioteknologi ITB.

Dana Waskita is a lecturer at the Faculty of Fine Arts and Design of Institut Teknologi Bandung, Indonesia. He graduated from the Faculty of Arts, the University of Melbourne in 2005. His research of interest includes ELT annd CACDS (Corpus Assisted Critical Discourse Analysis. Besides a lecturer and researcher, he also actively participates in the Language Center of Institut Teknologi Bandung as an English instructor and the Academic Manager.

Linda Handayani Sukaemi is a lecturer at the Faculty of Fine Arts and Design of Institut Teknologi Bandung, Indonesia. She is graduated from Faculty of Cultural Sciences, Padjadjaran University in 2014. Her research of interest is cultural studies. Besides as lecturer and researcher, she is also an editor of Jurnal Sosioteknologi ITB.

Ferry Fauzi Hermawan is a lecturer at the Faculty of Fine Arts and Design of Institut Teknologi Bandung, Indonesia. He is graduated from Faculty of Cultural Sciences, Padjadjaran University in 2014. His research of interest is gender and sexuality. Besides as lecturer and researcher, he is also an editor of Jurnal Sosioteknologi ITB.