

## A LITERATURE STUDY ON THE INTONATION OF COLLOQUIAL INDONESIAN

Hana Nurul Hasanah  
hana.nh@ui.ac.id  
Universitas Indonesia

### Abstract

*Standard Indonesian is the high variety used primarily in formal writings and formal occasions. On the other hand, the commonly used variety in spoken language by Indonesian is colloquial Indonesian. This paper attempts to present and discuss research results concerning Indonesian intonation and illustrates the general picture of colloquial Indonesian intonation. Considering the significant findings in the previous research, this paper concludes the possible future investigations of colloquial Indonesian intonation. Based on the literature study on Indonesian intonation, previous researches were limited to the intonation of colloquial Indonesian as spoken by speakers from various substrate languages. Thus, there is a need to confirm previous findings with data from more speakers, particularly speakers of the most influential variety, namely Jakarta Indonesian and constructed with controlled stimuli that also consider the role of pragmatic meaning on intonation pattern.*

**Keywords:** *Intonation, colloquial Indonesian language, literature review.*

### Abstrak

*Bahasa Indonesia baku merupakan ragam tinggi yang terutama digunakan dalam penulisan dan acara yang bersifat resmi. Di sisi lain, ragam yang pada umumnya digunakan dalam tuturan lisan oleh orang Indonesia adalah Bahasa Indonesia lisan sehari-hari. Artikel ini bertujuan untuk menyajikan dan mendiskusikan hasil-hasil penelitian mengenai intonasi Bahasa Indonesia dan mengilustrasikan gambaran umum dari intonasi Bahasa Indonesia lisan sehari-hari. Dengan mempertimbangkan temuan-temuan utama dalam penelitian terdahulu, artikel ini menyimpulkan kemungkinan penelitian lanjutan terhadap intonasi Bahasa Indonesia lisan sehari-hari. Berdasarkan kajian pustaka mengenai intonasi Bahasa Indonesia, penelitian-penelitian terdahulu masih terbatas pada intonasi Bahasa Indonesia lisan sehari-hari yang dituturkan oleh penutur dengan latar belakang bahasa yang beragam. Dengan demikian, terdapat keperluan untuk mengkonfirmasi temuan-temuan terdahulu dengan menggunakan data dari lebih banyak penutur, khususnya penutur ragam bahasa yang paling berpengaruh, seperti bahasa Indonesia dialek Jakarta. Selain itu, penelitian lanjutan perlu dikonstruksikan dengan stimuli terkontrol yang juga mempertimbangkan peran makna pragmatik terhadap intonasi.*  
**Kata kunci:** *intonasi, bahasa Indonesia lisan sehari-hari, kajian pustaka*

### INTRODUCTION

The production of words or phrases in natural speech is accompanied by melody and rhythm. Chains of pitch movements from mid-level to high and then to low level or other patterns are produced with particular tempo, and sometimes added with sudden pitch movements as we deliberately want to make a specific part more noticeable than others. Since intonation can express these all, we could simply regard intonation as the distinctive melodic pattern above lexical level. Intonation belongs to the realm of prosody; it is the melody of speech that is linguistically structured to convey pragmatic meanings of phrases or utterances as a whole (Ebing 1997: 13; Ladd 2008: 6).

Despite the essential use of intonation in spoken language, a systematic study

of Indonesian<sup>1</sup> intonation or other research on Indonesian prosody has not received much attention particularly by Indonesian scholars. Among the total of 20 scholars that had conducted research on Indonesian intonation and/or prosody<sup>2</sup> that Roosman (2007: 90-91) and Sustiyanti (2009: 5) separately mentioned in their papers, only eight are local Indonesians. Moreover, as to my knowledge, most of the authors did not exclusively concentrate their study on Indonesian intonation at length; only three authors did: Samsuri (1971), Amran Halim (1984), and Ewald Ebing (1997). Other authors discussed word accent, temporal stress (see Alwi et al. 2003: 19) or attempted to clear the dispute between stress, accent, and phrasal boundary in Indonesian.

Linguists who are non-native speakers of Indonesian, such as Dutch or English researchers, have considered Indonesian as a language that has weak stress. This conclusion has developed due to the misinterpretation of large pitch movements that are commonly found in colloquial Indonesian speech. These researchers have tried to construct stress rules for something that is cannot be regarded as stress (Goedemans and Zanten 2007: 59). The non-native linguists, who have proposed stress rules for Indonesian in the past, use stress in their native language, and subsequently they tend to perceive pitch movements in a foreign language, toward which they are not exposed on a daily basis, as stresses too (Zanten and Goedemans 2009: 220). However, a native speaker linguist, Halim (1984), has refused the existence of stress in Indonesian. Hence, to objectively examine and draw a non-bias conclusion on Indonesia intonation or prosody, any linguists should not associate or even force the rule of their native language in analyzing the use of pitch movements in Indonesian prosody. Furthermore, careful phonetic experimentation is also a necessity for conducting objective research towards a variety that is being used as koine in a nation that is the home for approximately 550 languages (Sneddon 2006: 2).

Previous studies that are presented in this paper were based on research conducted towards speakers of Indonesian that the structure of word prosody in their substrate language employ and not employ stress rules. A previous research on Indonesian prosody has proven that the distinct interpretation regarding pitch movements in Indonesian, whether based on production or perception experiment, is dominantly affected by the rule of stress in native Indonesian speakers' substrate language.

This paper aims to provide an overview of previous studies on Indonesian intonation and discuss to what extent results on Indonesian intonation pattern—

---

<sup>1</sup> This term refers to *Bahasa Indonesia* (BI), the standard variety or Standard Indonesian (SI).

<sup>2</sup> Research on Indonesian prosody started during colonialization, William Marsden in a book published in 1812 already discussed Indonesian word stress. I assumed the language that was being investigated was not current Indonesian but the high variety of Riau-Johore Malay. In addition, Alwi et al argued that Indonesian was based on this variety, but the language has somehow evolved, and currently, Standard Indonesian is notably different from the Riau-Johore Malay (2003: 15).

encompasses the interface between intonation and accent in natural speech—have been gained. The paper is organized as follows. The next section deals with the definition of *Bahasa* Indonesia; the standard variety and the colloquial one, Section 3 and 4 provide studies on the representative pattern (model) of Indonesian intonation, Section 5 explores related studies concerning word prosody and accent in Indonesian and other varieties that resemble them, Section 6 and 7 presents the conclusion and suggestion for further research.

### **THE STATE OF *BAHASA* INDONESIA**

Standard Indonesian (SI), as the high variety spoken in Indonesia, serves as a medium of instruction in formal education, national ceremony, formal circumstances, such as for radio and television news broadcast. Thus, SI is merely the language spoken by teachers, news anchors, and government staff. As for daily conversation, in a non-formal situation, people in the streets, pupils, or even educated Indonesians among their peers will express their feelings, ideas, and thoughts by utilizing a code that is not precisely the same as the SI that is found in formal situations. The code that is being used for daily conversation is the colloquial Indonesian, which is regarded as the low variety of Indonesian.

As to my knowledge, there is no standardized variety among the colloquial Indonesian varieties that are spoken in Indonesia. However, the most influential one is the variety spoken by educated/middle-class Jakartans. This variety is not Betawi Malay, but it is the non-standard/colloquial, Jakartan Indonesian, which also referred as colloquial Jakartan Indonesian (CJI). Betawi and Javanese Indonesian speakers are the ones that primarily shaped CJI. The main difference between the colloquial and the standard variety lies on the lexicon and morpho-syntax. As for the prosody, we may hear vowel lengthening to emphasize quality (occur in adjectives) or as a stalling device while the speaker thinks of what to say (see Sneddon 2006 for more detailed characteristics of CJI).

CJI has served as koine among immigrants of other language backgrounds that come and dwell in Jakarta and Bodetabek<sup>3</sup>, its surroundings area (Cikarang, part of Bekasi; Karawaci and Serpong, part of Tangerang; Sentul, Cibinong, part of Bogor; and Depok). The children of these immigrants, some of whom are monolingual (can

---

<sup>3</sup> Abbreviation for Bogor, Depok, Tangerang, Bekasi—the satellite cities of Jakarta.

only speak Indonesian), do not master the regional language of their elders<sup>4</sup>. The next generation of these immigrants speaks CJI with a very mild or even without regional accent. CJI is having an increasingly significant influence on varieties throughout Indonesia. It is also used by students and educated people in other cities, primarily by those who have at one time or another been to Jakarta (Sneddon 2006: 5). Even though there has not been any discussion of CJI intonation at length, its intonation is more or less characterized by the pronunciation of Betawi speakers (the native people) and Javanese speakers as a large number of them inhabit Jakarta.

## MELODIC MODEL FOR INDONESIAN INTONATION

This section presents a synthesized description of two notable works on Indonesian intonation, which respectively conducted by Halim (1974, 1984) and Ebing (1997). They have investigated the representative pattern (model) for Standard Indonesian that is used in spoken language. Their works are based on their doctoral research.

Although the term autosegmental-metrical (AM) phonology had not been popularized at the time he conducted his research, Halim's approach in analyzing intonation is in line with the AM theory, which views intonation as a string of contours that has distinct units. In Halim views, phonologically, an intonation group is set up by intonation phonemes, namely pitch level, accent, and pause. Moreover, certain combination of pitch levels refers to a contour, which each encompasses a pre-contour, occurs at the beginning of an intonation group and a primary contour (nuclear contour), which carries accented syllables to mark focused constituent. He distinguishes three pitch levels: low = 1, mid or neutral = 2, high = 3; three junctural phenomena: falling ( $\downarrow$ ), rising ( $\uparrow$ ), sustained ( $\_$ ), and two types of pauses: tentative pause ( $/$ ) and final pause ( $\#$ ). The pitch level for pre-contour is "2", while pitch level for nuclear contour is "3". In other words, an intonation group is shaped by the collective use of these intonation phonemes, whilst the contour type and the terminal juncture type for each group depends on the topicalization and focalization.

Halim identifies four possible pitch patterns derived from configuration of the three pitch levels: 211, 231, 232, and 233 (Halim 1974: 106). The phonetic realization based on topic-comment relation and focalization is exemplified in Table 1.

---

<sup>4</sup> I quoted this statement from Wouk (1996: 62) that was meant to refer to the children of immigrants in Jakarta only. However, I deliberately generalized Wouk's statement for a broader situation. To my knowledge and as part of the next generation of the children immigrants in Jakarta, nowadays, children who are monolingual or can only speak in Indonesian are found not only in Jakarta but also in its surroundings. Children or grandchildren of the immigrants in Jakarta have grown, and some have become parents and have moved to and dwelled in the satellite cities of Jakarta and their children are the next generation who use no regional language and can only speak Indonesian.

**Table 1. Halim’s melodic pattern for Indonesian intonation.**

Type	Melodic pattern	Example
Focalized topic	233 <sub>r</sub> 231 <sub>f</sub> (topic) (comment)	(1) <i>Doni sedang tidur.</i> 2-33 <sub>r</sub> / 2- 3 1 <sub>f</sub> # 'Doni is sleeping.'
Focalized comment	232 <sub>f</sub> 211 <sub>f</sub> (comment) (topic)	(2) <i>Kemarin mereka di sini.</i> 2-3 2 <sub>f</sub> / 2 11 <sub>f</sub> # 'It was yesterday that they were here.'

In Halim’s view, three distinct sentence modes—declarative, interrogative, imperative may share the same model. However, if we look over the examples provided by Halim (1974: 133-138), we may find that the mutual intonation pattern is likely to be applied to each of the sentence modes because each mode does not share the same lexical constituents. The questions were lexically marked by the interrogative pronouns like *ke mana* ‘where’, *apa* ‘what’; and the command was marked by particle *lah*. Therefore, I assume variations will follow if one construction, for instance declarative sentence, is being used to separately convey a statement and a question. To be specific, the patterns will display disagreement at the contour of “comment” in a sentence with focalized topic and at the final boundary tone.

Halim (1974, 1984) and Ebing (1997) both aimed at proposing intonation model for Indonesian. Both of them sought for a general trend inferred from speakers’ speech, which did not vary too much, in my assumption, due to the small number of speakers engaged. Despite the distinct substrate languages of the informants in Halim’s research and Ebing’s research<sup>5</sup>, five out of six melodic patterns proposed by Ebing (1997) can be associated with the patterns proposed by Halim (1974, 1984). Table 2. below displays Ebing’s pattern (left column) and its counterpart from Halim’s (right column).

Table 2. A comparison of Ebing’s and Halim’s melodic model.

RFL (rise fall low) or RFH-WRE (rise fall high-wide range-early timing)	231 <sub>f</sub>
RFH-NRE (rise fall high-narrow range early timing) and/or RFH-NRL (rise fall high- narrow range late timing)	232 <sub>f</sub>
R (rise)	233 <sub>r</sub>
RHP (rise high post-tonic)	233 <sub>f</sub> and/or 233 <sub>s</sub>
FRF (fall rise fall)	31 <sub>f</sub>
RFR (rise fall rise)	--

<sup>5</sup> Halim’s informants are himself and his wife, natives of South Sumatra; Ebing’s informants are natives of Riau, Yogyakarta, and Bandung.

From Ebing's and Halim's model as illustrated in Table 2., we may claim that Indonesian intonation patterns generated from distinct speakers of Indonesian share more similarities than discrepancies.

Halim has declined word accent in Indonesian, in terms of the function of pitch to indicate lexical contrast, but has acknowledged that in natural speech, accent does fall either at the penultimate (the default location) or at the final syllable. The position of such accent is predictable; determined by types of focalization. Syntactically, focalization in Indonesian is realized by word order, while melodically, this is indicated by boundary tone. A focalized topic is accompanied with a rise, and on the contrary, a focalized comment is accompanied with a fall. The emergence of boundary rising has caused a shift on accent placement from the penultimate to the final syllable. Thus, we may (temporarily) regard that nucleus placement in Indonesian is predictable—located at the primary contour of the focalized constituent or the nuclear accent. However, since focalization in Indonesian is mainly realized by word order, we need to question do speakers of Indonesian need to melodically re-emphasize certain constituents that are already clearly highlighted through word order and consequently, whether the relatively higher pitch and longer duration carried by a particular syllable as shown by Halim's mingograph tracing in his work could always be considered as an indication of accent.

Ebing (1997) respectively studied contrastive accentuation below the word level and on integral words in short-utterances. He studied the realization of narrow focus in Indonesian numerals, either in final or non-final position. Data used in his study. Data were elicited through question and answer as follow<sup>6</sup>:

Q: *Apakah Anda mengatakan / tujuh kali tiga / tambah lima?*

'Did you say seven times three plus five?'

A: *Bukan / saya mengatakan: / DUA kali tiga / tambah lima.*

'No, I said: two times three plus five.'

$$(2 \times 3) + 5$$

From this particular experiment, Ebing found that Indonesian listeners are heavily confused about the accent-lending and boundary-marking functions of the speaker's pitch movements (1997: 100). The perception experiment has shown that the listeners have a great tendency to think that an accent will occur on a word that is immediately

---

<sup>6</sup> Q-A sentences and English translation are copied from Ebing (1997: 98), boundary marker "/" and capitalized word indicating focus constituent are added by the author.

followed by a phrase boundary. For example, the last numeral (“five”) tend to be perceived as an accented word, regardless how it is intended. On the contrary, listeners also tend to interpret fluctuating pitch movements as boundary tone when they occur immediately after an accented word. For example, they tend to perceive pitch movement as a phrasal boundary when the first numeral is under focus. Their judgement is probably because the contrastive focus is realized by placing pitch accent at random position in the target word and, subsequently, contrastive focus is perceived without preference. Thus, it is not an easy task for Indonesian listeners to recognize phrase boundary-marking and contrastive accent during perception experiment. The finding indicates that even though prominence and boundary marking are discrete functions of Indonesian pitch movements, based on Ebing’s result, acoustic cues indicating the two features somehow have shown substantial intervention (Ebing 1997: 112).

As a conclusion, there is an indication that in Indonesian accent-lending pitch movement co-occurs with phrasal boundary-marking. Therefore, large pitch movements are often misperceived by the listeners. However, Ebing’s findings have not given sufficient explanation regarding the acoustic manifestation of focus/prominence and how boundary is marked in Indonesian as spoken by speakers with no-stress-rule substrate language. It is unclear which acoustic cues—duration or pitch—that presumably have caused listeners’ confusion in identifying boundary-marking phenomenon.

## QUESTION VERSUS STATEMENT-OPPOSITION

On the first decade of the twenty first century, investigations towards colloquial Indonesian were also concerning the melodic model but highlighted more on distinct patterns due to various functions of intonation such as for qualifying information.

Sugiyono’s article “Struktur Melodik Bahasa Indonesia (Melodic Structure in Indonesian)” reported findings on both production and perception experiments towards statement and echo question intonation in Indonesian<sup>7</sup> (Sugiyono 2007). His work has paid more attention to the acoustic issues on Indonesian intonation, particularly on  $F_0$  height distinction between intonation groups across intonation types. His data was generated from an acted speech produced by ten speakers (six males and four female). Target sentences shared similar lexical constituents and syntactic structure, namely SV, SVO, and SVOAdv.

The production experiment sought for the distinctive acoustic features for

---

<sup>7</sup> Sugiyono did not mention which varieties of Indonesian his participants speak.

Indonesian statement and echo question. Meanwhile, the perception experiment measured listeners' sensibility in identifying the two types of speech. The first experiment has shown that the discreteness between statement and echo question is determined by at least three acoustic features: fundamental frequency ( $F_0$ ) of the final note peak ( $P_a$ ),  $F_0$  of the final mark, and the position of  $P_a$  in an intonation contour. However, only two out of the three parameters are the main cues that are essential to enhance listeners' recognition, namely the height of  $F_0$  for the  $P_a$  and  $F_0$  of the final mark—both  $F_0$  for a question is higher than for a statement. Even though Sugiyono (2007) investigated the melodic marking of statement in contrast to echo question, he had no attempt to further investigate the acoustic manifestation of the old information stated in the echo question that is presumably under focus.

From his result, we gain knowledge regarding the typical intonation pattern for Indonesian statement and question. The first intonation group in a statement starts with an initial fall after which it remains level until the penultimate syllable and followed by a salient rise on the final syllable of the first intonation group to signal boundary (233<sub>i</sub>). The second intonation group starts with a higher  $F_0$  than the previous one, and pitch rises gradually until somewhere in between the end of the penultimate syllable and the beginning of the final syllable where pitch starts to rise saliently in the final syllable (final peak note) and ended with a fall (231<sub>i</sub>). His findings on intonation pattern of statement share similarities with Halim's pattern of statement with focalized topic.

The discreteness between statement and intonation is not exhibited over the entire sentence but falls on the last intonation group (Sugiyono 2007: 5-6), which is constituted by either a V or O or Adv, depending on the structure. A steep rise in question appears on the penultimate syllable and is followed by a configuration of fall and rise at the end. Meanwhile, for statement, a significant fall only appears after the final note peak. Moreover,  $F_0$  height for the final boundary in a statement is lower than its referent, while in a question, it is always higher than its referent and  $P_a$  of the previous constituent.

“Intonasi Kalimat Deklaratif dan Interogatif Konfirmatoris Bahasa Indonesia oleh Penutur Lampung” (Intonation of Declarative and Confirmatory Interrogative Sentences of Indonesian Language by Lampung Speakers) is an unpublished master's thesis by Sustiyanti. She investigated contours for each sentence constituents in Indonesian declarative and confirmatory interrogative modes. A total of 60 sentences corpus, which consisted of five types of target sentences in an SVOAdv pattern, were derived from an unspontaneous dialog spoken by four native Lampung Indonesian speakers. The target sentences were a declarative sentence mode and its corresponding

confirmatory interrogative sentences (lexically marked and unmarked).

Based on acoustic measurements she argues that utterances in declarative mode start with an initial rise (2 semitone), whereupon it remains level until the antepenultimate syllable where pitch rises again and reaches its  $P_a$  somewhere in between the penultimate and final syllable, then followed by a fall—final fall always appears in the last syllable. On the other hand, utterances in confirmatory interrogative mode start with a steep rise, and after that it remains level until somewhere in between the end of the antepenultimate and the beginning of the penultimate syllable where it starts to fall significantly to the beginning of the last syllable then followed by an abrupt rise and falls again at the coda of the last syllable. Thus, interrogative sentence intonation has relatively wider pitch range than declarative sentence intonation.

A clear distinction between the final pitch level of declarative and interrogative sentences was not explained verbally in her thesis. However, based on the intonation contour pictures, it can be observed that  $F_0$  height for the initial rise in interrogative sentence is higher than declarative sentence, and the final  $F_0$  for interrogative sentence is slightly higher than the lowest  $F_0$  found at the onset of the first word of the sentence (the subject). On the other hand, the final  $F_0$  for declarative sentence is either approximately the same or significantly lower than the lowest  $F_0$  found at the onset of the first word. In addition, in interrogative sentence, pitch reaches its  $P_a$  earlier than declarative sentence, it is somewhere in between antepenultimate and the penultimate syllable.

Her conclusion on the acoustic cues to distinguish declarative sentence from its corresponding confirmatory interrogative was drawn based on pitch movement in each constituent (Sustiyanti 2009: 96):

Table 3. Pitch movements in declarative and interrogative sentence modes.

Sentence mode	Pitch movement			
	Subject	Predicate	Object	Adverbial
Declarative	rise	level	rise-fall; fall; level	rise; final rise; initial flat; final fall
Interrogative	rise	level	level	final fall; rise-fall; initial flat; final fall-rise-fall

Sugiyono (2007) and Sustiyanti (2009) have shown an indication that sentence-final boundary rise is free; it can appear on the penultimate or the final syllable, or somewhere in between the antepenultimate and the penultimate or between the penultimate and the final syllable. Moreover, where pitch is going to have a steep final

rise is tied to its sentence mode—declarative has earlier final note peak compared to its corresponding interrogative. Their findings confirm my assumption that the contrastive signal for a pair of sentences that is segmentally and syntactically identical falls at the distinct of the  $F_0$  of final boundary tone, not at the  $F_0$  of the constituent before non-final boundary. A statement is indicated by a final boundary tone that is lower than its referent, and a question is indicated by a final boundary tone that is higher than its referent and the  $P_a$  of the previous constituents (Sugiyono 2007). Moreover, Sustiyanti (2009) also has shown an indication of distinct boundary tone configuration, such as fall tone for statement and rise-fall for question. Nevertheless, all are merely samples, and we should not hastily make any conclusions.

## PROMINENCE IN INDONESIAN

In a language that has no deterministic stress rules, such as colloquial Indonesian and some Indonesian regional languages (for example: Betawi Malay, Javanese), the issue concerning how focus is realized has attracted native and non-native linguists. Ebing (1997) has concluded that in Indonesian accent-lending pitch movement co-occurs with phrasal boundary-marking. Later, other linguists continue to investigate this phenomenon from an acoustic approach, and some took a comparative method. For example, Roosman (2007) compared the interaction of focus and intonation in disyllabic words in Toba Batak (TB) and in Indonesian variety that bears a resemblance to CJI, Betawi Malay (BM). The target words were under the condition of with and without focus at sentence-medial and sentence-final.

From Roosman's work (2007), it is noticeable that interference between focus and intonation affected the pattern of accent-lending pitch movements in BM and TB. The two varieties share common basic shape of pitch movements configuration "rise-fall", but for BM words under focus pitch rises and falls indeterminately. However, "rise-fall" is not obligatory to coincide in BM since rise may be absent for peak at pre-final, or only rise appears on the final syllable. Pitch rise at sentence-final needs more time to reach the centre of the accented (final) syllable. Meanwhile, the length and steepness of pitch rise are affected by the position of the word, where pitch rise is not as steep as the one in sentence-medial position. Thus, accent-lending pitch movement in a sentence-final focused word is perceived longer and less steep than the one in sentence-medial position.

As mentioned above, the position of pitch movement in BM is indeterminate even for words with focus. Zanten and Goedemans (2009) argued that for a language that has no distinctive stress the pitch cannot find a docking station on which it can always rely on any circumstances. We may conclude that this is the main reason why the position of pitch movement in Indonesian varies so much. Therefore, it would not be wise to consider any fluctuating pitch movement as an accent since there is no

Indonesian word whose one of the syllables will always serve as a docking station or be the one always stressed in a speech no matter where the word is placed in a sentence. One issue that is lack from their research is the acoustic manifestation of post focus constituents. As argued by other researchers that post focus constituents are affected phonetically in terms of post-focus compression (Jeon and Nolan 2017). In my opinion, one of the way to identify whether the fluctuating pitch movements are meant to indicate accent or not could be investigated by comparing the acoustic manifestation of its neighboring constituents.

One interesting fact that can be found from Zanten and Goedemans' result is that Indonesian speakers will use the prosodic pattern from their substrate languages in their production and perception of Indonesian. Since Toba Batak Malay speakers have lexical stress and it is the penultimate syllable that bears stress, they tend to produce penultimate syllable in Indonesian with a higher pitch so that it would be perceived louder than any other syllables. Meanwhile, their perception experiment has shown that Indonesian listeners are tolerant, as TB listeners tend to treat prominence produced by TB speaker and Javanese speakers differently. In a sentence produced by a TB speaker, a TB listener will perceive accent on the pre-final syllable; on the other hand, in a sentence produced by a Javanese speaker, TB listeners' judgement is not well determined, as they try to compromise with the speaker's daily use of pitch movements. TB listeners have fully understood that the Javanese prominence could occur anywhere or there is no prominence at all.

A more recent study by Athanasopoulou and Vogel (2016) has found that acoustic effects of focus on Indonesian trisyllabic words (serve as the object in SVOAdv pattern) are shown by (1) an increase in duration for the last two syllables and (2) a decrease in  $F_0$  for the final syllable of the target word. Pitch accent is not obligatory in the form of an increase in  $F_0$  since obtusion of a pitch at a particular point from the surrounding syllables is essential to determine the position of pitch accent Cruttenden (1997). These cues are simultaneously marking the end of a prosodic constituent that is followed by a non-final boundary. On the other hand, Indonesian trisyllabic words without focus condition have the following characteristics: (1) the pre-final syllable has the longest duration and (2) the final syllable has the shortest duration compared to the other two but has the highest  $F_0$ . The decrease and increase in duration and  $F_0$  are based on the comparison between focused and unfocused context. Therefore, although  $F_0$  on the final syllable is not as high as the one without focus, the highest  $F_0$  within the target words falls at the last syllable. Nevertheless, the experiment was not aimed to provide sufficient elaboration for the distinction in the pattern of pitch movements found between target words under focus and without focus.

However, there is a lack of evidence regarding interference between focus and intonation in CJI, especially from the acoustic approach. In addition, further studies

in CJI should consider different types of utterances (statement and question), the pragmatic meaning that lies behind the utterance, and also the speaker's substrate language.

## AN OUTLINE OF COLLOQUIAL INDONESIAN INTONATION

Currently, monolingual Indonesian speakers are centralized in Jakarta and its satellite cities. Previous research on colloquial Indonesian neither exclusively engage speakers from Jakarta as their informants, nor mainly study the intonation of the so-called most influential variety in Indonesia big cities—Jakartan Indonesian (CJI). Thus, what can be concluded as the general picture of colloquial Indonesian is limited to the intonation of colloquial Indonesian as spoken by speakers from various substrate language.

Pitch movements and intonation pattern in colloquial Indonesian, regardless speaker's substrate languages, are mostly tied to the sentence structure (SVOAdv) (see Sugiyono 2007 and Sustiyanti 2009). Whereas the first, second, or third constituent remains level, pitch rising has always been somewhere near the right edge of phrase-final, coinciding with boundary marking. It is proven in Betawi Malay (BM) that both phrase and sentence boundary marking affect the pattern of accent-lending pitch movements in focused constituents (Roosman 2007). The same phenomenon will likely to occur in colloquial Indonesian as BM is a variety that largely shapes colloquial Indonesian. Hence, nucleus placement in Indonesian presumably is not only determined by the type of focalization (with focalized topic or comment) as suggested by Halim (1974, 1984). In addition, works from the linguists illustrated above have shown that Indonesian speakers with different substrate languages behave differently at producing and perceiving accent as a focus marker. Consequently, it seemed unnecessary to construct one default rule for nucleus placement in colloquial Indonesian since the speakers of the language come from distinct substrate languages.

To summarize, there are three pitch levels (low, mid, high or 1, 2, 3) and three boundary tones (rise, fall, sustained) that can be used to describe the colloquial Indonesian intonation. The principal pattern for Indonesian intonation are either  $233_r / 231_f \#$  or  $232_f / 211_f \#$ . Moreover, the first, second or pre-final intonation groups may share common pitch level configuration as long as none of the intonation groups has a focalized constituent. However, the pattern is not applicable for a lexically unmarked question. It is proven that in unmarked question intonation, an utterance ends with a boundary rise and the  $F_0$  in its last contour of the intonation group is always higher than the  $P_a$  of its preceding constituents (Sugiyono 2007). Based on this result, a lexically unmarked question would not end with low pitch and a fall ( $31_f$  or  $11_f$ ).

## SUGGESTIONS FOR FURTHER RESEARCH

Indonesian linguists in the twenty-first century have paid more attention to the acoustic cues in Indonesian intonation. Both Sugiyono (2007) and Sustiyanti (2009) have reported an increase in  $F_0$  for the final constituent of declarative and interrogative sentence modes as in both conditions the expansion is to indicate boundary marking. However, both works have disregarded the acoustic manifestation of sentential focus. On the other hand, based on Roosman's (2007) investigation in BM, a focused constituent that is located before a phrase boundary has a variation on its accent-lending pitch pattern, since boundary rise has made pitch rise sharply. On the contrary, pitch rises gradually and has a lower height in focused constituent that is located before a sentence-final due to declination effect. Hence, in such a broad/integrative focus condition, we assume similar variations will likely appear in Indonesian intonation. We also expect  $P_a$  shifting due to a decrease  $F_0$  on the penultimate and the final syllable of a constituent that is located before a boundary. What is the pitch pattern within the focused constituent will be in colloquial Indonesian and to what extent boundary marking affects accent-lending pitch movement in colloquial Indonesian? Will the typical model in the absence of focus by Sugiyono (2007) and Sustiyanti (2009) remain the same if there is a focus? These are some problems that await further investigation.

Previous studies on colloquial Indonesian intonation have also illustrated the interference between intonation and focus; however, they have not provided comprehensive elaboration regarding the interaction between intonation type and pragmatic meaning. Therefore, there is a lack of empirical evidence that illustrates the interference between intonation pattern and pragmatic meaning. To query information is indeed represented by a question but not the other way around since question is also being used to express pragmatic meaning such as suspicion and surprise or to ask for confirmation of what has been heard. It is interesting to observe the interaction between pitch movements to indicate accent in focus constituents and pitch movements to indicate intonation meaning.

Listeners identify the existence of accents by hearing all or partially the three prosodic features—pitch, length, and loudness. Undoubtedly, pitch is the most consistently used prosodic features, with the other two factors playing a minor role which probably occurs in most intonation languages (Cruttenden 1997: 40). However, the account is not proven in Indonesian since the previous literature on prominence in Indonesian reported that tonal and temporal features are the most salient acoustic cues; both of which are equally manifested. Furthermore, linguists who have investigated the acoustic manifestation of prominence in Indonesian refuse to regard the acoustic phenomena found in focused words as an accent. This opinion is based on the following essential facts. First, there is no single word in Indonesian whose syllable, by default, regardless its location in a sentence will serve as the docking station of the fluctuating

pitch movement (see Zanten and Goedemans 2009). Next, pitch movements occur coincidentally with a major prosodic break which makes the so-called phrasal-accent consistently happens somewhere near the right edge of the phrase-final word (see Athanasopoulou and Vogel 2016). Thus, the acoustic manifestation of prominence in Indonesian is in line with the general theory concerning focal structure in languages without lexical stress, that Jeon and Nolan (2017) regard as prosodic phrasing phenomenon and phonetic events near prosodic boundaries.

We have noticed that previous works are limited to observe the realization of focus on the object of a sentence, which simultaneously is the constituent before the sentence final/non-final boundary. Hence, it is unwise to assume the function of pitch movements in a focused constituent is to exclusively indicate boundary tones, which will always occur with the same pattern regardless of the form of the focused constituents. By coincidence, both Zanten and Goedemans, as well as Athanasopoulou and Vogel, used target words that were the objects of verbal sentences used as the carrier in each experiment. In general, in the absence of topicalization, an object in Indonesian verbal sentence is located somewhere at sentence right, and it can be at sentence final (sentences used in Zanten and Goedemans' experiment) or phrase final (sentences used in Athanasopoulou and Vogel's experiment).

Next similarity found from previous research on Indonesian intonation lies on the participants involved in the research. Speakers and listeners involved are those whose substrate language has no deterministic stress rules, and those who are not familiar with "stress/accent phenomena" and at the same time, are accustomed to indicating prominence by word order, which can be perceived by pause/break. As a result, they have little knowledge of how to highlight important message using prosodic cues. Despite the fact that studies on Indonesian intonation are only relevant to a particular dialect or language variety, overgeneralization is an issue that every researcher should avoid. Nevertheless, based on Zanten and Goedemans findings, observing Indonesian intonation as spoken across distinct speakers' substrate languages is worth to study. For example, it is worth to study speakers of monolingual Indonesian and speakers with a defined stress rule in their substrate languages, such as Manado Malay or Toba Batak. Hence, in observing focalized constituents in colloquial Indonesian, at least there are three variables under concern: focus type (narrow or broad focus), the location of focalized constituent (near boundary or not), and the regional language of the speakers.

Further studies need to investigate the interaction between accent-lending pitch movements, phrase boundary tone, and pragmatic meaning and also the acoustic effect of their interaction on the post focus constituents. In particular, for the acoustic manifestation of colloquial Indonesian prominence, more observation is needed for prominence on compounds or two-word phrases since previous research mainly

observed focus on one content word. On the other hand, researchers should also provide a comprehensive description of Indonesian intonation by linking phonetic evidence with phonology by applying the auto segmental-metrical theory. Two prosodic constituents that can be utilized for intonational analysis are Phonological Phrase and Intonation Phrase. The two constituents have been utilized by Stoel (2007) for the description of Manado Malay intonation, a variety that is closely related to Indonesian and has a reasonably similar intonation system for statements (2007: 148).

## REFERENCES

- Alwi, H. et al. (2003). *Tata bahasa baku Bahasa Indonesia*. Third edition. Jakarta: Balai Pustaka.
- Athanasopoulou, A. & I. Vogel. (2016). The acoustic manifestation of prominence in stressless language. *Proceedings of Interspeech 2016*. doi: 10.21437/Interspeech.2016-1424.
- Cruttenden, A. (1997). *Intonation*. Cambridge: Cambridge Univ. Press.
- Ebing, E. (1997). *Form and function of pitch movements in Indonesian*. Leiden: Research School CNWS.
- Goedemans, R. & E. van Zanten. (2007). Stress and accent in Indonesia. In V. J. van Heuven & E. van Zanten (Eds.), *Prosody in Indonesian Languages* (pp. 35–62). Utrecht: LOT (Netherlands Graduate School of Linguistics).
- Halim, A. (1974). *Intonation in relation to syntax in Bahasa Indonesia*. Jakarta: Djambatan.
- Halim, A. (1984). *Intonasi dalam hubungannya dengan sintaksis Bahasa Indonesia [Intonation in relation to syntax in Indonesian]*. Jakarta: Djambatan.
- Ladd, D. R. (2008). “Introduction to intonational phonology” in D. R. Ladd, *Intonational phonology*. doi:10.1017/CBO9780511808814.
- Jeon, H.-S., & F. Nolan. (2017). Prosodic marking of narrow focus in Seoul Korean. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 8(1), 2. doi: <http://doi.org/10.5334/labphon.48>.
- Roosman, L. (2007). Melodic structure in Toba Batak and Betawi Malay word prosody. In V. J. van Heuven & E. van Zanten (Eds.), *Prosody in Indonesian Languages* (pp. 89–115). Utrecht: LOT (Netherlands Graduate School of Linguistics).
- Sneddon, J. N. (2006). *Colloquial Jakartan Indonesian*. Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, Australian National University.
- Stoel, R. (2007). The intonation of Manado Malay. In V. J. van Heuven & E. van Zanten (Eds.), *Prosody in Indonesian Languages* (pp. 117–150). Utrecht: LOT (Netherlands Graduate School of Linguistics).

- Sugiyono. (2007). Struktur melodik Bahasa Indonesia. *Kajian Linguistik dan Sastra*, 19(1), 1–13. Retrieved from <http://journals.ums.ac.id/index.php/KLS/article/view/4403>
- Sustiyanti. (2009). Intonasi kalimat deklaratif dan interogatif konfirmatoris Bahasa Indonesia oleh penutur Lampung (Master's thesis). Retrieved from <http://lib.ui.ac.id/>.
- Wouk, F. (1999). Dialect contact and koineization in Jakarta, Indonesia, *Language Sciences*, 21, 61–86. Retrieved from <http://udel.edu/~pcole/fieldmethods2010/Wouk.Jakarta.Ind.pdf>
- Zanten, E. v. & R. Goedemans. (2009). Prominence in Indonesian: Stress, phrases, and boundaries. *Wacana, Jurnal Ilmu Pengetahuan Budaya*, 11(2), 197-225.