

The Perception of Tone and Focus in Mandarin by Indonesian Learners: A Case Study

Hana Nurul Hasanah
Faculty of Humanities, Universitas Indonesia
Email: hana.nh@ui.ac.id

Abstract

In a tone language, the interface between tone, intonation, and focus will affect the pitch height and contour of tones. Previous perceptual studies revealed the potential conflicts in perceiving pitch variations at lexical and post-lexical levels that were experienced by either native listeners or listeners who speak Mandarin language as a second or foreign language. Rarely we find research in Indonesia that provides evidence for Mandarin language learners' perceptual ability at a post-lexical level. This paper investigated how well learners with distinct first language (L1) background identify tones that are affected by the realization of focus and the presence and location of focus in distinct intonation types. Perceptual experiments were conducted towards two groups of listeners: Mandarin learners with Indonesian L1 and learners with a tone language L1 background (Hakka or Hokkien). Their identification accuracy (IA) rate in recognizing the tone type for the last syllable with a narrow focus was compared with their IA in identifying the location of focus. In general, identifying tone type was easier than identifying focus position for both groups. However, the Mean from each group showed that learners with a tone language L1 were slightly better than the other group. Results exhibited more similarities between the two groups of listener, which indicates that L1 background only has a mild effect on the perceptual ability of Indonesian learners of Mandarin as a foreign language.

Keywords: Indonesian learners of Mandarin, perceptual experiment, post-lexical level

1. INTRODUCTION

In a tone language such as Mandarin (*Putonghua* 普通话), one of Chinese language (*Hanyu* 汉语) families, pitch is primarily used to distinguish lexical meaning. There are four primary lexical tones in Mandarin: (1) *Yinping* (阴平), a high-level tone; (2) *Yangping* (阳平), a mid-rising tone; (3) *Shangsheng* (上声), a low-dipping tone; (4) and *Qusheng* (去声), high-falling tone (Chao 1968:26; Yang 2016:64). These four lexical tones are generally referred to as Tone 1, Tone 2, Tone 3, Tone 4; and this paper will use the symbol T1, T2, T3, and T4 to refer to each tone respectively. Chao (1968:25-26) coined the five-point scale (1 to 5) to illustrate the turning point (only for T3), the starting and ending points for each tone produced in isolation. The numerical points 1 to 5 represent "low", "half-low", "middle", "half-high", and "high". Thus, the distinctive pitch patterns for T1 to T4 respectively are high-level (55); high-rising (35);

low-dipping (214); and high-falling (51).

Other linguistic functions for pitch in Mandarin, as also found in non-tone languages, are to mark prominence of certain intended meaning and to qualify the information presented in an utterance. In natural speech, the various uses of pitch in Mandarin inevitably lead to the interaction within lexical tones, and interaction between tone, focus, and intonation. The influence of intonation upon tone may take different forms (Lehiste 1970:100). In Mandarin, the F_0 height of a lexical tone is affected by the manifestation of pitch at syntactic level. A rise or a fall of the intonational band (final or non-final boundary tone) will raise or lower the F_0 height of a given tone at final syllable, but boundary-tone effect does not occur on other syllables in the phrase or clause (Norman 1988:149).

Many languages use specific variations of pitch and duration to indicate prosodic domain of focus, which could be expressed in an entire constituent or only to make one syllable of a phrase or a sentence is more prominent than others (Van Heuven 1994:15). The acoustic manifestation of focus in Mandarin statement and question, as far as pitch is concerned, has been investigated by Liu and Xu (2005). As mentioned earlier, pitch use at syntactic level affected the F_0 height of a lexical tone rather than its contour and the same trend is also occurred for a narrowly focused constituent. However, based on Liu and Xu's results the acoustic effect of focus also affects the pitch of its neighbouring tones. The realization of focus in a certain constituent has made pitch range of the particular focused word expanded and has affected the post-focus constituents by means of compressing and lowering the pitch range but the pre-focus constituents principally unaffected (Liu & Xu 2005:85). In addition, this particular condition of focus in Mandarin is found both in questions and in statements.

According to a more recent research, there is a tendency that the expansion of F_0 range is in line with the characteristic of a given tone and each tone demonstrates distinct F_{0peak} . For narrowly focused syllables with tones that do not have a turning point (T1, T2, T4), the F_0 peak for each tone is significantly high ([+RAISEH]). On the other hand, we can see a lower F_0 trough ([+LOWERL]) for narrowly focused syllables with T3 (Chen & Gussenhoven 2008:743; Lin & Li 2011:1248). Thus, to perceive pitch variations at word and sentence levels is not an easy task even for Mandarin L1 listeners. Previous perceptual experiments revealed that acoustic manifestation of lexical tone affects native Mandarin listeners in understanding the speaker's meaning; they encountered confusion due to the simultaneous use of pitch for tones and focus (Yuan & Shih 2004, Yuan 2004). Findings from previous research bring up the question how well do foreign learners of Mandarin perceive these simultaneous uses of pitch at sentence level?

Based on listening exercise result in the classroom, I found that learners often failed to recognize T4 that occurred in the final syllable of a question. These learners' substrate language is Indonesian or a non-tonal Indonesian regional language. I assumed the failure arose from the distinct realization of T4 in isolation with the one they heard in question intonation. The falling tone of T4 in this particular context was not realized from high to low point (51) due to the final rise effect in question. Their limited knowledge regarding the acoustic manifestation of tones at sentence level presumably caused them misperceived the identity of lexical tones they heard within stretches of utterances. Learners tend to consider the pitch movements of the last syllable they hear

merely as the contour of the lexical tone carried by the final syllable. Nevertheless, their failure was also assumed because L1 interferences since the primary use of pitch in Mandarin and in Indonesian are different.

A comparative perceptual study between L2 listeners with a tone language L1 and L2 listeners of Mandarin with a non-tone language L1 showed that the L2 listeners with a non-tone language L1 exhibited better sensitivity to intonation than to Mandarin lexical tones; on the contrary, listeners with a tone language L1 were more sensitive to lexical tones (Liang & Van Heuven 2007). Therefore, listeners' success in recognizing pitch use in Mandarin utterances is not only related to the listeners' knowledge to Mandarin prosody, but also to their sensitivity towards the various uses of pitch in an utterance.

Mandarin is one of the tone languages that is widely spoken whether as first, second, or foreign language. In Indonesia, Mandarin language is taught as a foreign language. The first language of Indonesian learners of Mandarin in this country consists of two general groups, namely Indonesian language and tone language background, i.e. Hokkien, Hakka, and Teochew language.¹ Previous perceptual studies regarding perception of Mandarin prosodic features have been conducted, but the experiments, mostly concentrated on the perception of native listeners or L2 listeners. On the other hand, research conducted towards Mandarin language learners as a foreign language, usually were regarding learner's production (Guo & Tao 2008, Hasanah 2011, Yang 2016). Recent studies on pitch perception by foreign language learners were exclusively for the perception of Mandarin lexical tones at lexical level towards Mandarin learners in European countries (Ding, Jokisch, & Hoffmann 2010; Chen & Kager 2011; Gao 2016).

Based on Liang and Heuven's result (2007), listener's difficulty during perception was not only due to acoustic reason but also due to their L1 background. Nevertheless, there is still a lack of knowledge regarding the perceptual ability of foreign language learners which also considers the learner's L1 background. In this sense, this study involves two types of Mandarin learners in Indonesia, the one with a tone language L1 background and the one with non-tone language L1,² and investigates their perception of two discrete linguistic use of pitch in Mandarin—to mark lexical meaning and to mark sentence prominence. This study aims to reveal the similarities and differences that may occur between two groups of learners based on their accuracy rate for various perceptual tasks. However, this study is not intended to seek whether the distinction that lies between the two groups are significant or not.

In particular, this study focuses on how lexical tones are perceived by learners when the F_0 of the tones are being congruent or incongruent with the boundary tones and at the same time carry final focus. Moreover, this study also focuses on the identification of focus and the location of focal prominence in statement and question intonation. It is expected that learners with a tone language background will demonstrate better perceptual ability for identifying tone types. Since the sensitivity of listeners on recognizing sentence focus is related with the listener's L1 stress rules,

¹Hokkien and Hakka are languages used in China and member of Han language family. Teochew is a sub-dialect of Hokkien.

²None of their L1 background is a stressed language.

whether the L1 has a deterministic stress rules or not,³ both groups of learners may have similar accuracy rate in identifying sentence focus.

By utilizing perceptual experiment, this research is expected to provide evidence on what sort of context that will contribute to learner's difficulty when recognizing tone types of a syllable with narrow focus and identifying sentence focus produced in statements and questions. Thus, language teachers may have some clues on how to minimize the misunderstanding/failure in comprehending native speaker's utterances during listening comprehension. In addition, they may use findings from this research to anticipate mistakes that might occur during speech production by learners. On the other hand, by comparing the perceptual ability of two types of learners, we may earn some knowledge on whether L1 background plays role on the future success of learners in learning Mandarin, especially on learners' listening and speaking skills.

2. METHODS

2.1. Materials

Ten pairs of simple sentences with minimally contrasted lexical tone in the final position (T2 vs T4) were constructed in order to investigate learner's perceptual ability to differentiate/identify the tone types. Simple sentences were used in this study, instead of spontaneous speech, in order to control the intonation types and the investigated tone types. Each sentence was equal in length (eight syllables) and ended with an open syllable. The meaningfulness of sentences used for this experiment had been consulted with a native Mandarin speaker. Each pair of simple sentences was produced with two types of intonation, namely, statement and the corresponding syntactically unmarked yes/no question; and also produced with broad focus and narrow focus in the middle and in the final position. The targeted words for medial focus were nouns of locality: *shangbian* 'top', *xiabian* 'bottom', *zhongjian* 'middle', *youbian* 'right', and *zuobian* 'left', and the targeted words for final focus were semantically meaningful mono-syllable words either in tone 2 or tone 4. All sentences were transcribed in Hanyu Pinyin Romanization. Number 1, 2, 3, or 4 at the end of each syllable were used to indicate its lexical tone type. Emphasized words were printed in bold. The following sentences are examples of the materials used in this research:

- 1) Ta1 shuo1 shang4bian de zi4 du2 bai2
'He/she said the top character is read *bai* (white)'
(no focal prominence; as if to answer such question "ta1 shuo1 shen2me?").
- 2) Ta1 shuo1 shang4bian de zi4 du2 bai4?

³Stress is also marked by tonal and temporal use but always accompanied by a docking station; the place where the accented part or the focused part is located. Listeners who are exposed by this particular linguistic use in their native language, are trained to recognize the location of pitch movements used for stress marking and to identify the most prominent syllable (Van Zanten&Goedemans 2009). Consequently, they will implement this rule when recognizing focus constituent in other languages. On the other hand, listeners with a stressless language L1, are less sensitive to distinguish pitch as prominence marker from pitch as boundary markers and they may consider any steep rise and fall as accent (Van Heuven 1994:19).

‘He/she said the top character is read *bai* (fail or lose in a battle/contest)’ (no focal prominence; to confirm the whole information of what the person just heard).

- 3) Ta1 shuo1 shang4bian de zi4 du2 **bai2**
(with focal prominence in final position; to emphasize the character is pronounced as “*bai2*” with rising tone, not “*bai4*” with falling tone).
- 4) Ta1 shuo1 shang4bian de zi4 du2 **bai2?**
(with focal prominence in final position; to confirm certain information whether the top character is pronounced as “*bai2*” with rising tone just like what the person just heard).
- 5) Ta1 shuo1 **shang4bian** de zi4 du2 bai2.
(with focal prominence in middle position; to emphasize that “the top” character and not the left or the right or any other position that is pronounced as *bai2*).

In total, 120 sentences were designed (10 simple sentences x 2 Tones x 2 Intonation x 3 Focus). Forty (40) of the 120 sentences had syllables that carried focus in the final position of the sentence. In other words, during tone-type identification test 40 sentences served as target sentences and the remaining 80 sentences served as filler sentences.

2.2. Recording and Stimuli Preparation

One native Mandarin speaker who is also our faculty native teacher acted as the informant. She was instructed to read the sentences naturally and asked to avoid any exaggerated emotional prosody. She was asked to read each sentence twice. Each context was explained to the informant beforehand. I accompanied her during the whole recording session.

2.3. Listeners

Twelve (12) students who studied Chinese as their major at a public university in Indonesia were selected and voluntarily participated in this experiment. All listeners learn Mandarin as a foreign language and none of them speak the language at home or at their community as a second language.

Listeners consisted of two groups as follow: (1) group 1: six students (varying level) with a tone-language L1 background, either Hokkien, Hakka, or Teochew; (2) group 2: six students (6th semester students) with a non-tone language (Indonesian)⁴ L1 background. Sixth semester students were chosen for this experiment as they were in their final period of learning the language, thus, had passed the process of pronunciation

⁴Their substrate languages are either colloquial Jakarta Indonesian or Indonesian variety that is a non-tonal language and has no deterministic stress rule.

learning and were regarded to have sufficient exposure or at least familiar to distinct types of utterances.

2.4. Procedure

The identification task was conducted in a quiet classroom. 120 stimuli were presented to the listeners using a media player and a loud speaker. Each stimulus was played twice with a 3 seconds-length interval. In total, the time used for the identification task was approximately 60 minutes. A five-minute break was given in between the two types of identification.

In the first identification task, tone-type identification, listeners were requested to identify the tone-type for the final syllable. However, they were not informed that the last syllables would be either tone 2 or tone 4, thus, they were instructed to write the numbers 1, 2, 3, or 4 as their answers. In the second identification task, focus-position identification, the same 120 stimuli previously presented for tone-type identification test were once again played to the listeners in different order. They were requested to write a dash symbol (-) for utterance recognized as sentence without focal prominence, write "M" for utterance recognized as a sentence with medial focus, and write "F" for utterance recognized as a sentence with focal prominence in the final position. All answers were written in an answer sheet provided by the experimenter.

3. RESULTS AND DISCUSSION

3.1. Tone-type Identification

For this particular test, each tone type was randomly presented in ten pairs of simple utterances that syntactically identical but had different intonation types, i.e. statement and question. In order to understand how well learners' perception on tones that were affected by the realization of focus, we calculated the percentage of accurate identification by listeners. The rate is called Identification Accuracy (IA) rate (Xu & Mok 2012:102).

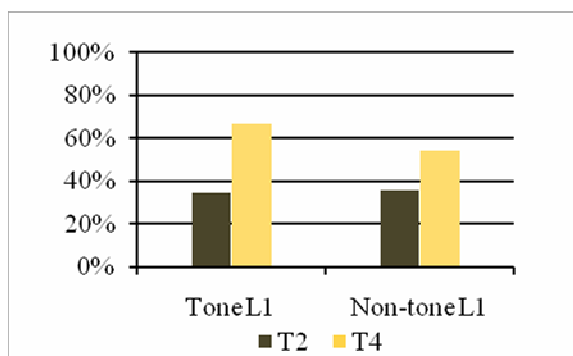


Figure 1. Average IAs of tone-type identification

For Tone-type Identification task, IA rate was the percentage of correct identification by listeners towards the tone types given in various contexts of utterance. High IA rate was gained due to the high degree of easiness in recognizing tone type of the target utterance. On the other hand, a low IA rate was a result of learners' confusion when identifying the tone type which made them failed to identify the tone type

correctly. The rate or the percentage itself is the Mean of correct responses for each identification tasks in varying contexts by each type of listeners.

Tone language L1 listeners were expected to gain higher IA for identifying tone type for the final syllable of a focused constituent in statement and question. Although they did better than the non-tone language L1 listeners, in general, the average rate of IAs for T2 and T4 in various contexts by both groups were relatively similar. Figure 1 illustrates the average IAs⁵ for this task.

Based on the percentage shown in figure 1, in general, listeners with tone language L1 background were slightly better than listeners with non-tone language L1 background for Tone-type Identification task. However, both groups of listeners have a common degree of easiness as well as difficulties. T2 was not easy to recognize for both groups regardless the intonation types. The IAs for both groups were less than 40%, respectively were 34.17% and 35.83%. On the other hand, T4 was easier to recognize by both groups especially for T4 in question intonation.

Figure 2 below illustrates listeners' accuracy for identifying the tone types of syllables with final focus (FF) in statement (S) or question (Q). For practicality, the horizontal axis in figure 2 uses symbols such as T2, T4; S, Q; and FF to respectively indicating tone types, intonation types, and focus location. For example, T2SFF refers to tone 2 in an utterance that was produced with statement intonation and final focus, T4QFF refers to tone 4 in an utterance that was produced with question intonation and final focus, and so on.

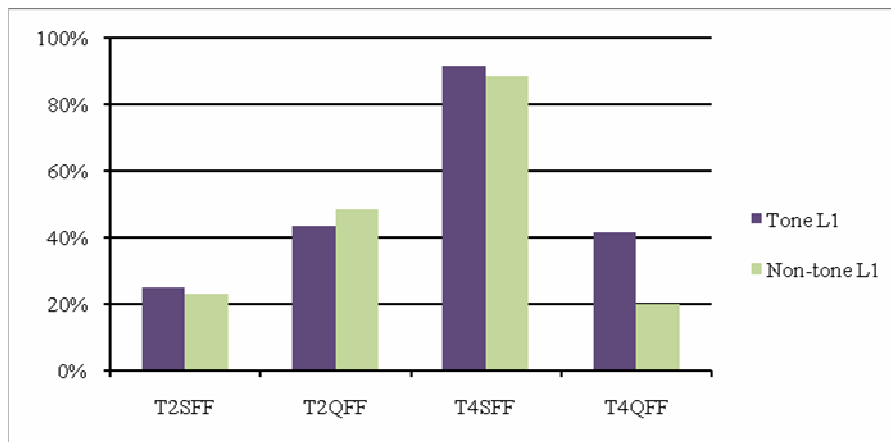


Figure 2. IAs of T2 and T4 in varying contexts by both types of listeners

Both types of listeners had high IAs for recognizing T4 in a statement with narrow focus at the end (T4SFF). Listeners from both groups had nearly excellent IA, 91.67% (tone language L1 background) and 88.33% (non-tone language L1 background). However, recognizing T4 in a question with final focus (T4QFF) was much more difficult for them; the IAs were lower than 50%. Moreover, this seemed to be the most difficult task for listeners with non-tone language L1 background. One

⁵Average IAs in figure 1 respectively refer to the Mean calculated from each IA of T2 and T4 with final focus that were found in two intonation types.

listener from this group failed in recognizing T4 with final focus in question intonation. Therefore, their IA for recognizing T4QFF was the lowest (20%) among other contexts. This result proves that rising boundary tone of question intonation and acoustic manifestation of narrow focus have greater prosodic strength in the final position of an utterance, as once suggested by previous literature (Yuan 2011).

While the average IA⁶ for T4SFF was the highest, the average IA for the contrasting pair (T2SFF) was very low; it was as low as 21.6% (for each group's response see table 1 in section 3.3). The varying degree of easiness in identifying tone types above lexical level was related to the global F₀ employed at the sentential level. High F₀ peak of final syllables with a narrow focus and the downward trend found in statements resembled Mandarin T4 characteristics—a high-falling tone. Consequently, T4SFF was much more noticeable than T4 at any other condition and could be easily recognized by the listeners.

The confusion encountered by the listeners in recognizing T4 in question intonation and T2 in statements seemed to arise due to the incongruence of F₀ encodings. F₀ employed to encode intonation meaning and the one employed to encode lexical meaning were asymmetrical. The upward trend and peak raising in T4QFF syllable were apparently competing for each other. As a result, rising boundary tone to indicate question intonation presumably had hindered the distinctive acoustic cue of T4, the steep fall, to be realized. This particular interaction had made T4 perceived more likely as a high-level tone accompanied by a slight fall. Therefore, listeners mostly misperceived T4 in syllables with final focus produced in question as T1.

A previous study has proven that a constraining semantic context (narrow focus) significantly improves question intonation identification mainly for sentences with T4 in the final position (T4QFF) (Liu, Chen, & Schiller 2016:1059). While based on their work, this particular context was helpful for intonation-type, on the other hand, based on this study, the same specific context was not beneficial for tone-type identification. Moreover, narrowly focused T4 syllables produced in question intonation were the second hardest condition to be well-perceived by the listeners regardless their L1 background. This is noticeable from the low IAs found in both groups, particularly for the non-tone language L1 listeners.

Based on the above explanations, the use of F₀ above lexical level hindered the acoustic identity of lexical tone. The interference of F₀ used as intonation marker and F₀ used as lexical tone marker had contributed to learner's confusion when recognizing the lexical tones. We may conclude with a caution that any rise or fall in any syllables heard by the listeners was considered exclusively as a distinctive acoustic cue to mark lexical meaning.

3.2. Focus-position Identification

For this part of identification test, listeners with tone language background also performed better than the other group. Nevertheless, based on data illustrated in figure 3, we may conclude that learners' perceptual ability in identifying focus location was very low regardless their L1 background.

⁶Average IA in this statement refer to the Mean calculated from the IA of a particular condition (i.e. T4SFF or T2SFF) by both types of listeners.

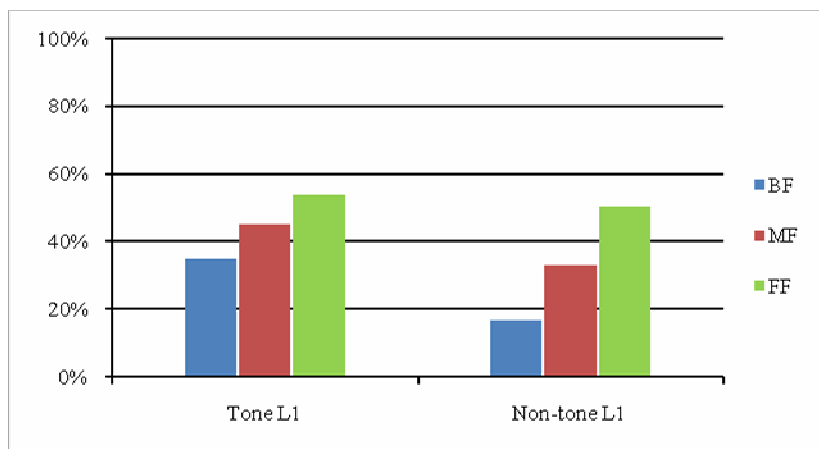


Figure 3. IAs of sentence focus-position by both types of listeners

Although, narrow focus and broad focus in Mandarin have a very distinctive acoustic pattern (see Chen & Gussenhoven, 2008), listeners' average IA for focus-identification was only 39%, lower than the average IAs of tone-type identification (47.71%). The low rate is assumed due to lack of training in the classroom since Mandarin prosody learning mostly concern with recognizing the canonical form of lexical tones and slightly touches upon the post-lexical use of Mandarin prosodic features.

Even though the average IA was very low, both types of listeners had at least reached 50% of accuracy for identifying focus at the end of an utterance. IAs for tone language L1 listeners and non-tone language L1 listeners respectively were 53.75% and 50.40%. On the other hand, among the three contexts given, namely broad focus (BF), medial focus (MF), and final focus (FF), the least easy to be perceived was utterances with broad focus, especially for listeners with non-tone L1 language background.

Both types of listeners shared similar IA rate for recognizing focus produced in question intonation, but both groups showed a different accuracy trend within the three focus types. Listeners with tone language L1 had gradual trend for each IA rate; on the other hand, IAs found in listeners with non-tone language L1 were very contrast for each focus types. For example, listeners with non-tone language L1 had very low IA rate (16.67%) for broad focus, meanwhile IA rate of medial focus was twice higher than the IA rate of broad focus, and IA rate of final focus was nearly twice higher than IA rate of medial focus. In other words, listeners with non-tone language L1 had preference on which focus types that was somewhat less difficult to be perceived.

Figure 4 illustrates listeners' accuracy percentage for identifying the presence and location of focus in varying intonation contexts. The horizontal axis stands for the combined contexts of the target utterances heard by the listeners. The utterance condition is a combination of intonation types (S, Q) and focus types (B, M, F). For example, SBF refers to statement in broad focus, QMF refers to question with medial focus, and so on.

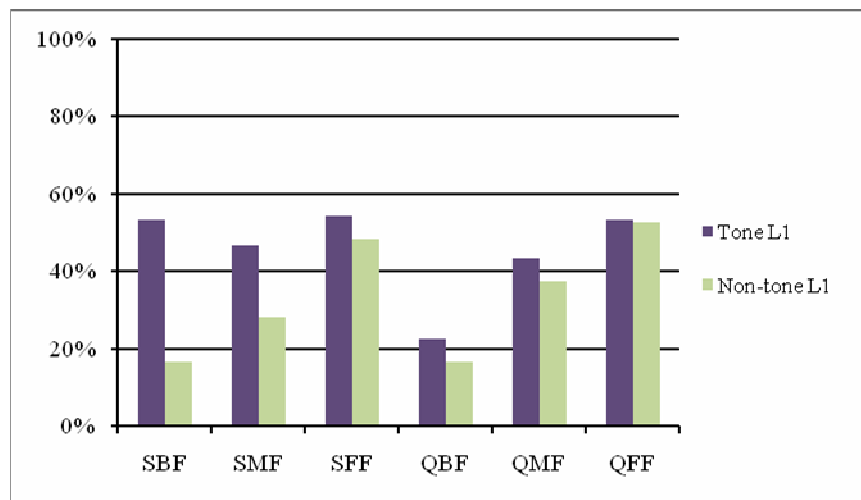


Figure 4. Identification Accuracy of three focus types in varying intonation types

The bar chart reveals that low IA found in both groups were not dependable to any intonation types, especially for tone language L1 listeners since their average IAs of each context was in between 43% to 53%. They only occasionally failed to reach 40% when identifying broad focus in question intonation. In addition, listeners with non-tone language L1 also had low IA rate for QBF. Average IA on QBF for both groups was only 19.6%. I can temporarily conclude with caution that for both types of listeners identifying broad focus in question intonation were much more difficult than any other focus context. In addition, from the bar chart presented in figure 4, it appears that for listeners with tone language L1 background, identifying focus in statement were somewhat less difficult; the average IA was 11% higher than the average IA in question. On the other hand, listeners with non-tone language L1 background had no preference on which intonation type that would be beneficial for them in identifying focus-position; their average IAs for focus-position identification in statement and question respectively were 31.11% and 35.55%.

From listeners' answers on focus type identification test, it was observable that both groups were most likely confused with medial and final focus. They recognized medial focus as final focus and vice versa. This particular finding is not supported by previous literature which have shown learners' confusion between initial and medial focus since both initial and medial focus were cued acoustically by the compressed and low pitch range of the post-focused words (Liu & Xu, 2005:82). From the present study, we may conclude that listeners, regardless their L1 background, have a very minor knowledge regarding the acoustic cues for sentence focus.

As mentioned earlier, the realization of focus in QBF was the least easy to recognize. Listeners mostly perceived broad focus in this condition as final focus. This confusion was related to tone types. If a question with broad focus was ended in T4, listeners would identify this utterance as statement with final focus; and if a question with broad focus were ended in T2, listeners would identify this utterance as question with final focus. This is assumed due to the similarity in F_0 between the upward trend in question intonation and the final raise of T2. Listeners misperceived this acoustic cue as

pitch range expansion of lexical tone which does commonly occur due to the acoustic effect on a focused word.

3.3. Tone language L1 Learners vs. Non-tone language L1 Learners

Despite the distinct L1 backgrounds of the listeners, both groups apparently had low accuracy rate for either recognizing tone type or identifying the presence and location of focus. As shown in figure 5, the average IAs of both groups was lower than 50%. The average IAs for various tasks by listeners with tone language L1 background was 47.5% and for listeners with non-tone language L1 background was 39.2%.



Figure 5. Average IAs between two types of listeners.

A lower than 50% IA indicates that to identify lexical tones and determine focus location that were simultaneously produced in sentential level were difficult to perceive for both types of listeners. In addition, a less than 10% difference on the perception ability of both groups implies that both types of learners may have equal opportunity in their development in learning Mandarin prosody. From the data collected, I found one good achiever listener from the non-tone language L1's group that reached 52% accuracy for identifying tone types. This IA rate was better than some of the tone language L1 listeners who only reached 40%.

A closer investigation towards the perception results revealed that both groups shared similar responses for the perception tasks, especially for tone-type identification task. Figures shown in Table 1 present the average rate of each type of listeners' responses for identifying tone-type identification, while Table 2 shows the average rate of each types of listeners' responses for determining focus location.

Table 1. Listeners' responses on tone-type identification

Target	Responses									
	T1		T2		T3		T4		T0	
	Tone L1	Non-Tone L1	Tone L1	Non-Tone L1	Tone L1	Non-Tone L1	Tone L1	Non-Tone L1	Tone L1	Non-Tone L1
T2										
S	1.67%	1.67%	25%	21.67%	73.33%	76.67%	0%	0%	0%	0%
Q	1.67%	3.33%	43.33%	53.33%	55%	4.33%	0%	0%	0%	0%
T4										
S	0%	5%	0%	3.33%	1.67%	1.67%	98.33%	90%	0%	0%
Q	35%	55%	15%	11.67%	10%	11.67%	40%	18.33%	0%	3.33%

Table 2. Listeners' responses on focus-position identification

Target	Responses					
	BF	BF		MF		FF
	Tone L1	Non-Tone L1	Tone L1	Non-Tone L1	Tone L1	Non-Tone L1
S	54.17%	19.17%	13.3%	23.33%	32.5%	57.5%
Q	23.33%	14.17%	23.33%	27.5%	53.33%	58.33%
MF						
S	25%	17.5%	42.5%	30.83%	32.5%	51.67%
Q	19.17%	5%	45%	36.67%	35.83%	58.33%
FF						
S	10.83%	14.17%	31.67%	34.17%	57.5%	51.67%
Q	11.54%	15%	30.83%	28.33%	56.67%	56.67%

As mentioned in earlier section, the average IA for identifying T2 under narrow focus in statements was very low (21.6%), moreover from the percentage displayed in table 1, it is observable that both groups of listeners often misperceived T2 in this condition as T3; shown by an approximately 70% responses for each group. The F_0 range expansion due to the realization of focus and final downward in statements misled them to determine what they heard as T3. Pitch movements signalling lexical tone and the ones signalling sentence boundary were also incongruent in T4QFF. The fall of T4 under narrow focus in question intonation was somewhat less steep than its canonical form (51). Most of the listeners misperceived T4 in this condition as T1. However, the percentage for the two types of listeners was not identical. The number of listeners with non-tone L1 background that misperceived T4 as T1 was bigger than the counterpart, their percentage respectively are 55% and 35%.

In general, listeners' responses were affected by the rise and fall of boundary tone. However, a comparison between their responses has shown a strong tendency for the non-tone language L1 listeners to determine their responses by merely relating the pitch movement of the last syllable to the canonical contour of lexical tone. They could perceive the falling and rising tone at the end of an utterance but failed to involve the knowledge that pitch movements over stretches of utterance in a tone language have lexical and sentential use. Hence, their strategy was only beneficial if the contour of boundary tone was congruent with the contour of lexical tone, such as identifying T4SFF and T2QFF. Their identification accuracy for T2QFF was slightly higher (53.33%) than the tone-language L1 listeners' (43.33%). Therefore, it presumably was

also due to this strategy that made most of the non-tone language L1 listeners misperceived T4 as T1 in T4QFF. On the other hand, the tone-language L1 listeners' responses were somewhat affected by the excursion of the pitch range. This is observable for identifying T2QFF and T4SFF.

Slightly different from the perception of tone types, in Table 2 it can be seen that the two groups did not always demonstrate similar mistakes, for example, in locating the presence of focus. Listeners with a tone language L1 were 50% accurate in identifying utterance with broad focus. On the other hand, nearly all listeners with a non-tone language L1 failed to recognize broad focus regardless the type of utterance. Most of them misperceived broad focus as final focus, moreover, most of the utterance they heard were considered as utterances with final focus. Among the three options that they could choose, BF, MF, or FF, an average of 55% responses were FF. This tendency presumably was due to L1 interference since in Indonesian a focalized constituent appears at the final word of an intonation phrase. In addition, speakers of Indonesian with a non-stress substrate language, in most cases, will put accent for a focused word somewhere in between the penultimate and the final syllable (Van Zanten & Goedemans 2009:215).

4. CONCLUSION

As shown by the perceptual experiment, both types of listeners, regardless their L1 background, did not perform well during tone-type identification and focus-position identification. Their average IA for both tasks was only 43.35%. This result shows that certain acoustic interaction due to the simultaneous uses of pitch in Mandarin, namely to indicate lexical meaning, to mark prominence, and to qualify the information, contributed confusion on learner's perception. The identity of lexical tone is hindered by the use of F0 at a post-lexical level, which contributes to learners' confusion when recognizing the lexical tones. As expected, learners' confusion was related with the incongruence of pitch movements. One particular context that both listeners found as the most difficult one was T2SFF. They confused with the identity of tone 2 under narrow focus produced in statement intonation. On the other hand, during the focus-position investigation, the most challenging context was QBF. The identification of broad focus was the least easy one compared to medial and final focus, especially when it appeared in question. Therefore, contexts which were considered difficult by listeners with a tone language background were also hard to handle by their counterpart. However, the two types of listeners did not always have similar responses, particularly for T2QFF, T4QFF, and SBF.

In summary, the identification accuracy rates between the two types of listeners were not fundamentally contrastive. Listeners with a tone language L1 background do not significantly have high sensitivity to Mandarin lexical tones. The IA rates shown in the previous section reveal the fact that for a small group of Indonesian learners, L1 background has a minor effect on learners' perception. From this finding, could be drawn a preliminary conclusion that in addition to L1 background, another factor that could also affect listeners' sensitivity is the exposure to the target language. This factor is conceivably the reason why the result of this research does not conform to the result from Jiang and Heuven's research (Jiang & Van Heuven 2007) that revealed L2 listeners' sensitivity on Mandarin tone and intonation is in line with their L1 background.

Since this study focused on describing the similarities and differences between two types of listeners, only a simple calculation was used in comparing results from the two groups of learners. Thus, it could not reveal whether differences found were significant or not. Future study needs to involve statistic measurement with a controlled variable to analyse the variance between two distinct groups. In addition to the data collection, future studies should use realistic context or dialogue to measure other aspects such as pragmatic context role in helping listeners' comprehension which may be beneficial for them when performing identification task. Moreover, future research should be conducted to a larger number of subjects to generalize the result. In addition, questionnaires or interviews may be utilized in giving prior knowledge for the experimenter regarding the listeners' experience in using Chinese dialect for daily communication—whether the Chinese dialect being spoken is standard (especially for the tones).

REFERENCES

- Chao, Y.R. (1968). *A Grammar of Spoken Chinese*. Berkeley, CA: University of California Press.
- Chen, A. & Kager, R. (2011). *The perception of lexical tones and tone sandhi in L2: Success or failure*. Proceedings of the 17th International Congress of Phonetic Sciences, August 17-21, Hong Kong.
- Chen, Y. & Gussenhoven, C. (2008). Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics*. Vol. 36 Issue 4, pp. 724-746. DOI: 10.1016/j.wocn.2008.06.003.
- Ding, H., Jokisch, O., & Hoffmann, R. (2010). *Perception and production of Mandarin tones by German speakers*. Proceedings of Speech Prosody 2010, Fifth International Conference, May 10-14, Chicago, Illinois.
- Gao, Man. (2016). *Perception of lexical tones by Swedish learners of Mandarin*. Proceedings of the joint workshop on NLP for Computer Assisted Language Learning and NLP for Language Acquisition, November 16, SLTC, Umeå, Sweden.
- Guo, L. & Liang Tao, L. (2008). *Tone production in Mandarin Chinese by American students: A case study*. Proceedings of the 20th North American Conference on Chinese Linguistics April 25-27, Columbus, OH.
- Hasanah, H.N. (2011). *Analisis kesalahan tonal pembelajar Bahasa Mandarin: Studi kasus di sebuah universitas negeri di Jakarta*. Unpublished Master's Thesis. Depok: Faculty of Humanities, Universitas Indonesia.
- Lehiste, I. (1970). *Suprasegmentals*. Cambridge, Mass.: MIT Press.
- Liang, J. & Van Heuven, V.J. (2007). Chinese tone and intonation perceived by L1 and L2 listeners. In C. Gussenhoven and T. Riad (Eds.). *Tones and tunes Volume 2: Experimental studies in word and sentence prosody*. Berlin: Mouton de Gruyter.
- Lin, M. & Li Z. (2011). *Focus and boundary in Chinese intonation*. Proceedings of the 17th International Congress of Phonetic Sciences, August 17-21, Hong Kong.
- Liu, F. & Xu, Y. (2005). Parallel encoding in focus and interrogative meaning in Mandarin intonation. *Phonetica*, 62, 70-87. <https://doi.org/10.1159/000090090>.

Hasanah, Nurul Hana. 2019. Perception of Tone and Focus in Mandarin by Indonesian Learners: A Case study. *IJOTL-TL* (2019), 4(1): 11-26. DOI:10.30957/ijotl-tl.v4i1.553.

- Liu, M., Chen, Y. & Schiller, N.O.(2016). *Context effects on tone and intonation processing in Mandarin*. Proceedings of Speech Prosody, Eighth International Conference, 31 May - 3 Jun, Boston, USA.
- Norman, J. (1988). *Chinese*. Cambridge: Cambridge University Press.
- Van Heuven, V.J.(1994). Introducing prosodic phonetics. In C. Odé& V.J. van Heuven (Eds.).*Experimental studies of Indonesian prosody* [Semaian 9] (pp. 1-26).Leiden: VakgroepTalen en Culturen van Zuidoost-Azië en Oceanië, Leiden University.
- Van Zanten, E. &Goedemans, R. (2009). Prominence in Indonesian: Stress, phrases, and boundaries. *Wacana, Jurnal Ilmu Pengetahuan Budaya* Vol. 11 No. 2, pp. 197-225.
- Xu, B.R.&Mok, P.P. (2012). *Cross-linguistic perception of intonation by Mandarin and Cantonese listeners*. Proceedings of Speech Prosody, Sixth International Conference, May 22-25, Shanghai, China.
- Yang, C. (2016). Tone errors in scripted conversations of L2 Mandarin Chinese. *Chinese as a Second Language Research* Vol. 5 No. 1, pp.63-86. DOI: 10.1515/caslar-2016-0003
- Yuan, J. (2011). Perception of intonation in Mandarin Chinese.*J. Acoust. Soc. Am.*Vol. 130 No. 6, pp.4063–4069. DOI: 10.1121/1.3651818.
- Yuan, J.,&Shih, C. (2004). *Confusability of Chinese intonation*.Proceedings of Speech Prosody, Fourth International Conference, March 23-26, Nara, Japan.

IJOTL-TL, Vol. 4, No. 1, January 2019

p-ISSN: 2502-2326; e-ISSN: 2502-8278

<https://soloccls.org> ; Email: ijotl@gmail.com

Center of Language and Cultural Studies, Surakarta, Indonesia

Hasanah, Nurul Hana. 2019. Perception of Tone and Focus in Mandarin by Indonesian Learners:
A Case study. *IJOTL-TL* (2019), 4(1): 11-26. DOI:10.30957/ijotl-tl.v4i1.553.
