The International Journal of Engineering and Science (IJES)
|| Volume || 6 || Issue || 9 || Pages || PP 22-29|| 2017 ||
ISSN (e): 2319 – 1813 ISSN (p): 2319 – 1805

The Variation of Pitch Height of Tones in Mandarin Utterance

*Qi Gong
College of Foreign Studies, Jinan University, Guangzhou, China
Corresponding Author: Qi Gong

ABSTRACT
In this study, the lowering pattern in Mandarin tone production is analyzed, and it is found that downstep exists for high (H) tones. They always drop to a lower scale compared to the forgoing ones. The degree of the first downstep is always larger than the following ones, but there are no differences in the degrees of downsteps for the later HL sequences. The speaker does not only raise the initial H tone when there are more downsteps, but also depress the final H tone. As for the L tones, due to the effect of final lowering, the lowering degree of the last downstep is always larger than the preceding ones. The utterance final L tones are not affected by the number of downsteps.

Keywords: *Tone, pitch, downstep*

Date of Submission: 30 AUG 2017 Date of Accepted: 17 SEP 2017

I. INTRODUCTION

Downstep refers to the stepwise lowering of High (H) tones in certain contexts in tone languages. In automatic downstep, H tones are lowered in sequences of alternating H and Low (L) tones. In nonautomatic downstep, often noted as H!H, there is no overt conditioning L tone between the two H tones [1]. In downstep, each successive H tone in longer downstepping sequences is lower than the preceding one, creating a cumulative 'staircase' pattern. The concept of downstep has been extended to studies of many nontonal languages, and has been incorporated in quantitative models of intonation in languages such as English, Swedish, and Japanese [2].

In tone languages, downstep is a common phenomenon, so speakers may employ 'foresight' in producing long downstepping sequences. Stewart [3] claimed that the pitch of H tones in downstepping sequences in Akan is sensitive to the number of following downsteps. He stated that the pitch of any particular high tone is raised by as many levels as there are downsteps in the subsequent part of the phrase, while the last H tone in the sequence tends to be realized at a constant level, its basic pitch. On the contrary, Schachter [4] maintained that the pitch of the first H tone in Akan is normally phonetically the same regardless of the number of the following downsteps, while later H tones descend to lower and lower values as the number of downsteps increases.

In Mandarin, there have been a number of studies related to tonal downstep, Xu [5] argued that anticipatory and carry-over tonal influences co-exist in Mandarin, and they differ both in magnitude and in nature. Carry-over effects are mostly assimilatory: the starting F0 of a tone is assimilated to the offset value of a previous tone. Anticipatory effects, on the other hand, are mostly dissimilatory: a low onset value of a tone raises the maximum F0 value of a preceding tone. Shih [6] pointed out that the F0 contour of a Mandarin utterance is affected by a number of factors, such as declination, downstep and final lowering, etc. Huang et al. [7] reported that in downstep in Mandarin, the low tone will compress the pitch range of the following syllables, and the main effect of downstep is on the topline.

The experiment reported here will investigate the pitch values of utterances with 2 to 5 HL sequences, i.e. utterances with 4 to 10 syllables, and the aim is to find out whether Mandarin utterances display downstep effects across H and L tones. The following questions are addressed:

(a) Do H and L tone sequences show lowering effects?
(b) What about the lowering degree at different position of the utterance?
(c) Is the initial H and L tone scaled higher as the number of downsteps increases?
(d) Does the final H and L tone drop to lower values in utterances with more downsteps?

II. METHOD

2.1 Stimuli

In Mandarin, there are four tones. Tone 1 is high, Tone 2 rising, Tone 3 low falling, and Tone 4 is a falling one. In order to address the questions of the present experiment, only Tone 1 (H) and Tone 3 (L) sequences are used. In the utterances designed, H and L tones alternate on successive syllables, that is, in the pattern of HLHL, HLHLHL, etc. In the corpus, each set contained 4 utterances, with 2 to 5 HL sequences, i.e. with 4 to 10 syllables in length. The following is one set of the utterances,

(1) Bianxie chugao. (2 HL sequences)
    To compile the draft.
(2) Bianxie gepu chugao. (3 HL sequences)
    To compile the music draft.
(3) Qinshou bianxie gepu chugao. (4 HL sequences)
    To compile the music draft himself.
(4) Kaishi qinshou bianxie gepu chugao. (5 HL sequences)
    To begin to compile the music draft himself.

In the corpus designed, there are four such sets, which make a total of 16 utterances.

2.2 Subjects and Recording

The utterances used in this experiment are recorded by eight native speakers of standard Mandarin, four males and four females. The test utterances for the experiment were recorded in a sound-treated room, with a short practice session before the actual recording. The utterances were presented in random order and were read 3 times by the subject, with the order of each repetition randomized separately. In the recording, the subjects were instructed to read in normal speed, in a natural style, without narrow focus. By this means, the subjects are expected to read each utterance as broad-focused. The total utterances used in this study are 384 (16 utterances × 3 repetitions × 8 speakers).

2.3 Measurements

Acoustic data are segmented and labeled, and F0 is extracted using Praat [8]. The extracted F0 is manually verified with reference to the cycle in the waveform. In this study, for the purpose of normalizing the F0 difference among the speakers, semitone is used as the unit of pitch, instead of Hertz, and the conversion is done by the following formula,

$$St = 12 \times \log_2 \left(\frac{F_0}{F_{0\min}} \right) \qquad (1)$$

In (1), F0 is the pitch value in Hertz, F0min as the low bound of pitch range of the speaker, and St is the semitone value.

1) Average pitch value of H tone

As Tone 1 in Mandarin is a level tone, average pitch value is calculated, instead of using the high point value. Average pitch value is the mean of the pitch values of a tone. For example, if the duration of the tone is 200 ms, the extraction will get 20 pitch values within it, and the average pitch value is the mean of the 20 values.

2) Low point value of L tone

Regarding Tone 3, which is a low falling tone, the low point value is used in this study. The low point value is the minimum pitch value of the tone, which may best represent the feature of low tones.

3) Lowing degree

In order to inspect the extent of the lowering effect, lowering degree is calculated. It is the difference between successive H or L tones, which is computed by the following formula,

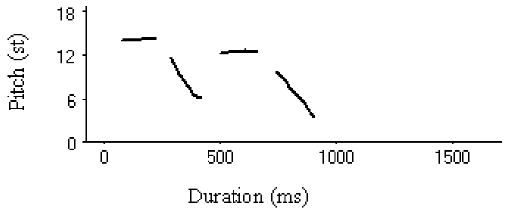
$$D_i = St_i - St_{i+1} \qquad (2)$$

In (2), Di stands for the value of lowering degree at position 'i', Sti for the pitch value of the tone at the same position, and Sti+1 for the pitch value of the following tone. Statistic analysis is done in SPSS.

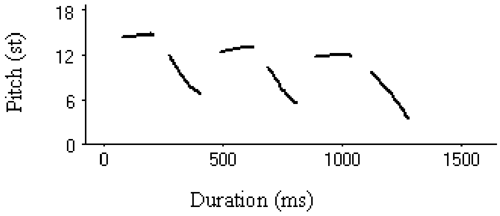
III. RESULT

Fig. 1 graphs the mean pitch values of utterances of various lengths for all the eight speakers, with (a) to (d) presenting values for utterances with 2 to 5 HL sequences respectively. In these graphs, the x-axis displays duration, and the y-axis displays pitch values in semitone. The line segments present the pitch contour, with each segment for one syllable, level one for H tone and low falling one for L tone.

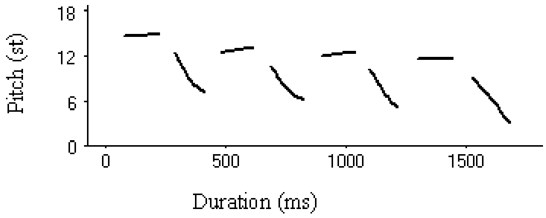
(a) Utterance with 2 HL sequences



(b) Utterance with 3 HL sequences



(c) Utterance with 4 HL sequences



(d) Utterance with 5 HL sequences

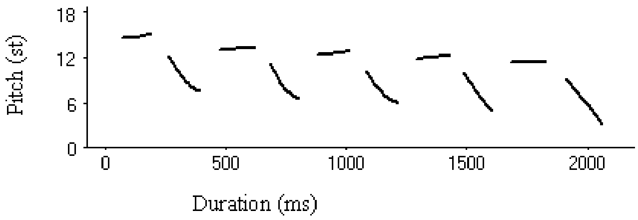


Figure 1. Pitch contour of utterances of various lengths[a]

a. The line segments present the pitch contour, with each segment for one syllable, level one for H tone and low f alling one for L tone.

3.1 H tones

3.1.1 Average pitch value

From Fig. 1 it can be seen that there is a prominent gradual lowering of H tones strongly resembling downstep throughout the utterances. Detailed analysis will be given as follow.

1) Utterance with 2 HL sequences

For utterances with 2 HL sequences, there is a downward trend for the H tones. It is shown from a repeated measures ANOVA result that there is significant difference between their pitch values: $F(1, 95) = 413.7$, $p < 0.001$, with the pitch of the first H tone (H1) higher than that of the second H tone (H2).

2) Utterance with 3 HL sequences

In regard to utterances with 3 HL sequences, downstep also exists for the H tones. Repeated measures ANOVA results show that H1 is higher than H2: $F(1, 95) = 492.7$, $p < 0.001$, and H2 higher than the third H (H3): $F(1, 95) = 173.7$, $p < 0.001$.

3) Utterance with 4 HL sequences

As for utterances with 4 HL sequences, there is still a gradual lowering of H tones. It is displayed from repeated measures ANOVA results that significant difference exists between pitch values of successive H tones, H1 vs. H2: $F(1, 95) = 388.5$, $p < 0.001$; H2 vs. H3: $F(1, 95) = 79.2$, $p < 0.001$; H3 vs. H4: $F(1, 95) = 88.8$, $p < 0.001$, with the preceding H tone higher than the following one.

4) Utterance with 5 HL sequences

When there are 5 HL sequences in an utterance, repeated measures ANOVA results show that the differences between the average pitch values of the successive H tones are still significant, H1 vs. H2: $F(1, 95) = 224.1$, $p < 0.001$; H2 vs. H3: $F(1, 95) = 83$, $p < 0.001$; H3 vs. H4: $F(1, 95) = 54.1$, $p < 0.001$; H4 vs. H5: $F(1, 95) = 98.4$, $p < 0.001$, with the foregoing H tone higher than the subsequent one.

3.1.2 Lowering degree

Table 1 displays the average lowering degrees (in st) of H tones for utterances of various lengths. In Table 1, Di stands for the lowering degree between the pitch values of Hi and Hi+1, for example, D1 stands for the lowering degree between those of H1 and H2, and D2 for that between H2 and H3. For utterances with 2 HL sequences, the lowering degree between the two H tones is 1.73 semitones.

Table 1. The average lowering degree (in st) of H tones for utterances of various lengths[a]

N. of HLs.	Lowering degree			
	D_1	D_2	D_3	D_4
2	1.73			
3	1.83	0.88		
4	2.02	0.67	0.74	
5	1.59	0.61	0.49	0.64

a. D_i stands for the lowering degree between the pitch values of H_i and H_{i+1}.

As for utterances with 3 HL sequences, it is shown from repeated measures ANOVA results that the lowering degrees for successive H tones are significantly different, with D1 greater than D2: $F(1, 95) = 72.6$, $p < 0.001$.

In regard to utterances with 4 HL sequences, repeated measures ANOVA results show that there is significant difference among the lowering degrees: $F(2, 190) = 87.2$, $p < 0.001$. Further analysis shows that D1 is greater than D2 and D3, D1 vs. D2: $F(1, 95) = 101.6$, $p < 0.001$; D1 vs. D3: $F(1, 95) = 142.9$, $p < 0.001$, but there is no significant difference between D2 and D3: $F(1, 95) = 0.39$, $p = 0.529$.

When there are 5 HL sequences in an utterance, it is indicated from repeated measures ANOVA results that the difference among the lowering degrees is also significant: $F(3, 285) = 37.2$, $p < 0.001$. Detailed analysis shows that D1 is larger than at any of the other positions, D1 vs. D2: $F(1, 95) = 40.2$, $p < 0.001$; D1 vs. D3: $F(1, 95) = 145.6$, $p < 0.001$; D1 vs. D4: $F(1, 95) = 46.9$, $p < 0.001$. However, there is no significant difference among D2, D3 and D4: $F(2, 190) = 1.19$, $p = 0.305$.

3.1.3 The initial H tone

In this subsection, the pitch values of the utterance initial H tones are analyzed. If initial H tones are scaled higher as the number of downsteps increases, the pitch values should rise for longer utterances. Table 2 displays the average values of the initial H tones and the ANOVA result, from which it can be seen that this is the case. Result from a repeated measures ANOVA shows that there is significant difference among them. Further analysis shows that, compared to that of utterance with 2 HL sequences, the initial H tones of longer utterances are higher, 2-HL vs. 3-HL: $F(1, 95) = 38.9$, $p < 0.001$; 2-HL vs. 4-HL: $F(1, 95) = 41.5$, $p < 0.001$; 2-HL vs. 5-HL: $F(1, 95) = 36.7$, $p < 0.001$. However, for utterences with three, four and five HL sequences, there is no significant difference among the pitch values of the initial H tones: $F(2, 190) = 0.83$, $p = 0.436$.

Table 2 Average pitch values (in st) of the initial H tones and the ANOVA result

N. of HLs.	2	3	4	5
Pitch value	14.12	14.61	14.72	14.73
ANOVA result	F(3, 285) = 16.7, p < 0.001			

3.1.4 The final H tone

With the effect of downstep, it is expected that as the number of downsteps increases, the final H tone will drop to lower and lower values. Table 3 presents the average pitch values of the final H tones and the ANOVA result, which indicates that there is indeed a continual lowering for the final H tones. It is shown from a repeated measures ANOVA result that there is also a significant difference among them. Further analysis shows that the final H tones of longer utterances are lower than those of shorter ones, regardless of the length of the utterance. To be specific, that of three HL utterance is lower than that of two HL utterance: $F(1, 95) = 70.9$, $p < 0.001$; that of four HL utterance lower than that of three HL utterance: $F(1, 95) = 12.6$, $p = 0.001$, and in turn that of five HL utterance is lower than that of four HL utterence: $F(1, 95) = 8.04$, $p = 0.006$. The final H tones do drop downwards as the number of downsteps increases.

Table 3 Average pitch values (in st) of the final H tones and the ANOVA result

N. of HLs.	2	3	4	5
Pitch value	12.39	11.90	11.62	11.40
ANOVA result	F(3, 285) = 76.7, p < 0.001			

3.2 Low tones

3.2.1 The low point value

As for low tones, Fig. 1 shows that there is also a gradual lowering of them throughout the utterances.

1) Utterance with 2 HL sequences

Similar to the H tones, for utterances with 2 HL sequences, there is also a downward trend for the L tones. Repeated measures ANOVA result shows that there is significant difference between their pitch values: $F(1, 95) = 167.9$, $p < 0.001$, with the pitch of the second L tone (L2) lower than that of the first L tone (L1).

2) Utterance with 3 HL sequences

For utterances with 3 HL sequences, lowering effect also exists for the L tones. It is shown from a repeated measures ANOVA result that L2 is lower than L1: $F(1, 95) = 83.9$, $p < 0.001$, and the third L (L3) lower than L2: $F(1, 95) = 138.7$, $p < 0.001$.

3) Utterance with 4 HL sequences

Regarding utterances with 4 HL sequences, there still exists a gradual lowering of L tones. ANOVA results display that significant difference exists between pitch values of successive L tones, L1 vs. L2: $F(1, 95) = 129.0$, $p < 0.001$; L2 vs. L3: $F(1, 95) = 94.0$, $p < 0.001$; L3 vs. L4: $F(1, 95) = 114.9$, $p < 0.001$, with the subsequent L tone lower than the preceding one.

4) Utterance with 5 HL sequences

When there are 5 HL sequences in an utterance, it is indicated from ANOVA results that the differences between the low point values of the successive L tones are still significant, L1 vs. L2: $F(1, 95) = 174.3$, $p < 0.001$; L2 vs. L3: $F(1, 95) = 94.8$, $p < 0.001$; L3 vs. L4: $F(1, 95) = 103.7$, $p < 0.001$; L4 vs. L5: $F(1, 95) = 122.7$, $p < 0.001$, with the subsequent L tone lower than the foregoing one.

3.2.2 Lowing degree

Table 4 shows the average lowering degrees (in st) of L tones for utterances of various lengths. For utterances with 2 HL sequences, the lowering degree between the two L tones is 2.33 semitones.

Table 4. The average lowering degree (in st) of L tones for utterances of various lengths

N. of HLs.	Lowering degree			
	D_1	D_2	D_3	D_4
2	2.33			
3	1.36	1.85		
4	1.20	1.20	1.96	
5	1.04	0.59	0.89	1.82

For utterances with 3 HL sequences, Repeated measures ANOVA results show that the lowering degrees for successive L tones are significantly different, with D2 greater than D1: $F(1, 95) = 5.88$, $p = 0.017$.

Regarding utterances with 4 HL sequences, it is shown from ANOVA results that there is also significant difference among the lowering degrees: $F(2, 190) = 11.6$, $p < 0.001$. Further analysis shows that D3 is greater than D1 and D2, D1 vs. D3: $F(1, 95) = 15.5$, $p < 0.001$; D2 vs. D3: $F(1, 95) = 14.1$, $p < 0.001$, but there is no significant difference between D1 and D2: $F(1, 95) = 0.000$, $p = 0.999$.

When there are 5 HL sequences in an utterance, it is displayed from repeated measures ANOVA results that the difference among the lowering degrees is still significant: $F(3, 285) = 24.3$, $p < 0.001$. Detailed analysis shows that D4 is larger than at any of the other positions, D1 vs. D4: $F(1, 95) = 18.0$, $p < 0.001$; D2 vs. D4: $F(1, 95) = 51.2$, $p < 0.001$; D3 vs. D4: $F(1, 95) = 22.9$, $p < 0.001$. In addition, D2 is smaller than D1 and D3, D2 vs. D1: $F(1, 95) = 21.3$, $p < 0.001$; D2 vs. D3: $F(1, 95) = 7.49$, $p = 0.007$. However, there is no significant difference between D1 and D3: $F(1, 95) = 1.72$, $p = 0.193$.

3.2.3 The initial L tone

Analysis on the initial H tones shows that they are scaled higher as the number of downsteps increases. In this subsection, the initial L tone will be investigated. Table 5 displays the low point values of the initial L tones and the ANOVA result, from which it can be seen that it is similar to the initial H tones. Result from a repeated measures ANOVA shows that there is significant difference among the initial L tones. Further analysis shows that, compared to that of shorter utterance, the initial L tones of longer utterances are higher, 2-HL vs. 3-HL: $F(1, 95) = 22.9$, $p < 0.001$; 3-HL vs. 4-HL: $F(1, 95) = 11.1$, $p < 0.001$; 4-HL vs. 5-HL: $F(1, 95) = 5.71$, $p = 0.019$. The longer the utterances are, the higher the low point values of the initial L tones are.

Table 5 Average low point values (in st) of the initial L tones and the ANOVA result

N. of HLs.	2	3	4	5
Pitch value	5.79	6.64	7.06	7.37
ANOVA result	$F(3, 285) = 41.5$, $p < 0.001$			

3.2.4 The final L tone

Table 6 displays the average low point values of the final L tones and the ANOVA result, which shows that, from utterance with two HL sequences to that with five HL sequences, there is no significant difference among their low point values. That is to say, the end points of the utterances tend to be constant regardless of the length of the utterances.

Table 6 Average low point values (in st) of the final L tones and the ANOVA result

N. of HLs.	2	3	4	5
Pitch value	3.46	3.43	3.08	3.04
ANOVA result	$F(3, 285) = 1.77$, $p = 0.153$			

IV. DISCUSSION

4.1 H tones

Results from this experiment show that, first of all, downstep exists for H tones, and this is true for utterances of various lengths, whether those with two or three HL sequences, or those with four or five HL sequences. This is also true for H tones at various positions, whether those at the earlier part of an utterance, or those at the later part. H tones always drop to a lower scale compared to the foregoing ones. When the lowering degree is analyzed, it is found that the degree of the first downstep is always larger than the rest. No matter how long the utterance is, this is always the case. As for the rest of the downsteps, that is, the second, the third and the fourth one, there are no significant differences among them.

Liberman and Pierrehumbert [9] put forward the Gradient model of downstep, which defines downstepping patterns as a gradual decay toward an abstract reference line, or asymptote. Their method of pitch assignment describes an exponentially decaying curve in which each step down is proportionally identical to the preceding one in terms of its distance from the reference line: Later downstep intervals are progressively smaller than earlier ones, and tend to become vanishingly small as the reference line is approached. This approach could be called a 'soft-landing' model of downstep implementation as it describes a curve similar to that of an aircraft gliding smoothly down to a landing strip.

As far as H tones are concerned, the downstepping pattern observed from this experiment resembles the soft-landing model to some extent, that is, the degree of the first downstep is the greatest. However, in this study, it is found that there are no differences in the degrees of downstep for the later HL sequences. The gradient for the later HL sequences keeps constant, rather than proportionally getting smaller.

In regard to the pitch values of the utterance initial H tones, it is shown that they shift upwards as the number of downsteps increases. This pattern can be attributed to 'foresight' in tone production, that is, the pitch of H tones in downstepping sequences is sensitive to the number of the subsequent downsteps. When there are more downsteps in an utterance, the speaker foresees this and will specify the initial H tone at a higher level. However, detailed analysis shows that this 'foreseeing' mechanism ceases to be in effect when there are three or more than three HL sequences in an utterance. That is to say, the effect of 'foresight' is limited, and the reason for this is that, generally speaking, there is a limit for a speaker's maximum pitch value.
Results from the previous section also show that the speaker does not only raise the initial H tone when there are more downsteps, but also depress the final H one. So there are two measures for the implementation of dowstep in Mandarin, to raise the initial H tone and to depress the final H one. Further analysis shows that the measure of the depressing of the final H tone is in effect even for utterance with four or five HL sequences. As is mentioned above, due to the constraint of the speaker's pitch range, the 'foresight' effect is limited. However, there is no limitation for the mechanism for the final H tone depressing. Compared to lowering the final H tone, raising the initial H tone is more energy consuming. Therefore, people tend to take the energy saving measure to realize the downstepping effect.

4.2 L tones

As for the L tones, it is shown that, similar to the H tones, there is also a downward trend for them, and this is true for utterances of various lengths, as well as L tones at different positions. In the context of downstep, the pitch of the L tones is also affected, with the later L tones lower than the earlier ones. The pitches of tones are relative. For example, the pitch of a male speaker is much lower than that of a female speaker, but when male and female speakers articulate a Tone 1 syllable, both will be correctly perceived. In this study, there are alternative H and L tones in the utterances, which will trigger downstep. In the context of downstep, the subsequent H tone will be lower than the forgoing ones, in which case the later tones will be reset in a lower scale. The L tones are basically lower than the H tones, so the subsequent L tones will be lower than the forgoing ones.
In regard to the lowering degree of the L tones, it is found that the degree of the last downstep is always larger than the forgoing ones, which is different from that of the H tones. No matter how long the utterance is, this is always the case. For utterance with 5 HL sequences, the lowering degree of the second downstep is the smallest. This lowering pattern is due to the 'final lowering' effect. Final lowering, the lowering of pitch at the end of an utterance, has been observed in many languages, like Spanish and Yoruba [10].
With respect to the the utterance initial L tones, it is indicated that they also shift upwards as the number of downsteps increases. When there are more downsteps, with the raising of the initial H tone, the entire pitch will be reset in a higher level, and the initial L tone is also raised to a higher level accordingly. Analysis shows that the low point values of longer utterances are always higher than those of short ones, which indicates that the initial raising effect for L tones is great in Mandarin.
As for the utterance final L tones, it is shown that the increase of the number of downsteps has no effect on their low point values. No matter how long the utterance is, the low point value tends to be constant. Researches show that, as far as statement is concerned, the utterance final pitch approaches a constant value, and this is true regardless of the initial pitch height of the utterance, or the variation of pitch within the utterance. Studies on both tone and non-tone languages, with both read corpus and natural speech, get similar result: the utterance final pitch tends to be constant [11].

V. CONCLUSION

In this study, the lowering pattern in Mandarin tone production is analyzed, and it is found that downstep exists for H tones, regardless of the length of the utterance, or the position of the tone. High tones always drop to a lower scale compared to the foregoing ones. The degree of the first downstep is always larger than the following ones, but there are no differences in the degrees of downsteps for the later HL sequences. The utterance initial H tones will shift upwards as the number of downsteps increases, but as there is a limit for a speaker's maximum pitch value, it ceases to be in effect for longer utterances. The speaker will depress the final H tone when there are more downsteps, and this is true for utterances of any lengths. Speakers tend to take the energy saving measure to realize the downstepping effect.
In regard to the L tones, it is found that lowering effect also exists for them, and due to the effect of final lowering, the lowering degree of the last downstep is always larger than the preceding ones. The utterance initial L tones also shift upwards as the number of downsteps increases, and this effect of great, with those of longer utterances always higher than those of shorter ones. However, the utterance final L tones are not affected by the number of downsteps, that is, they tend to be constant. The results from this study will be helpful for the pitch modeling in speech synthesis.

REFERENCES

[1] Laniran, Y. O. and Clements, G. N. Downstep and high raising: interacting factors in Yoruba tone produntion, Journal of Phonetics, 2003, (31), pp. 203–250.
[2] Ladd, D. R. Intonational phonology, Cambridge: Cambridge University Press,1996.
[3] Stewart, J. M. The typology of the Twi tone system, Bulletin of the Institute of African Studies, Vol.1, Legon, Ghana, 1965.
[4] Schachter, P. Some comments on J. M. Stewart's The typology of the Twi tone system, Bulletin of the Institute of African Studies, Vol.1, Legon, Ghana, 1965.
[5] Xu, Y. Contextual tonal variations in Mandarin, Journal of Phonetics, 1997, (25), pp. 61–83.
[6] Shih, C. Declination in Mandarin. Intonation: theory, models and applications, Proceedings of an ESCA work shop, Athens, Greece, 1997, pp. 293–296.
[7] Huang, X. Yang Y. and Lu S., Experimental studies on downstep in Chinese intonation, Acta Acustica, 2007, (32), pp. 57–61.
[8] Boersma, P. Praat, a syntem doing phonetics by computer, Glot International, 2001, 5:9/10, pp. 341–345.
[9] Liberman M. and Pierrehumbert, J. Intonational invariance under changes in pitch range and length, in Language and sound structure, M. Aronoff and R. T. Oehrle, Eds. Cambridge, MA: MIT Press, 1984, pp. 157–233.
[10] Truckenbrodt, H. Final lowering in non-final position, Journal of Phonetics, 2004; 32, pp. 313–348.

Qi Gong. "□ The Variation of Pitch Height of Tones in Mandarin Utterance." The International Journal of Engineering and Science (IJES), vol. 6, no. 9, 2017, pp. 22–29.