Influences of particles on Vietnamese tonal Co-articulation

Thị Lan NGUYỄN[1] Đỗ Đạt TRẦN[1]
(1) International Research MICA, 1 Dai Co Viet, Hanoi, VIETNAM
`thi-lan.nguyen@mica.edu.vn, Do-Dat.Tran@mica.edu.vn`

ABSTRACT

In continuous speech, the pitch contour exhibits variable patterns and it is strongly influenced by its tone context. Although several effective models have been proposed to improve the accuracy for tonal syllables, the quality of Vietnamese synthesis system is poor by lack of lexical parameters corresponding to each syllable in modelling of fundamental frequency. This problem will be clarified by our experiment in this study.

This paper presents our study on tonal co-articulation of particles which are frequently used in Vietnamese language. The obtained results show that tonal co-articulation phenomenon always takes place at the transition between two adjacent syllables, the progressive co-articulation is the basic tonal co-articulation and there is an influence of the function of particles on form of F0 contour of Vietnamese tones.

KEYWORDS : tonal co-articulation; Vietnamese tones; F0 generation ; lexical category; particles

163

1   Introduction

Vietnamese is a monosyllabic and tonal language. Each Vietnamese character corresponds to a syllable which is associated with a lexical tone. Syllables with different tone have different meanings, so the tone plays an important role in distinguishing lexical meaning for a family of syllables. As is well known, for a tonal language, like Vietnamese, Chinese (Mandarin or Cantonese) or Thai, fundamental frequency (F0) contours of utterances are always composed of tonal local features (tones and the co-articulation between adjacent tones) and the sentential intonation (corresponding to higher-level structures). This makes F0 variations of sentence more complicated than non-tonal languages such as English or French.

Recently, Vietnamese has been the subject of much linguistic research. Most of studies have concentrated on analyzing the characteristics of isolated word [2][3] or some on tonal features [1][4]. Researches on tonal co-articulation in Vietnamese were also presented in studies [1][4][5].

Tonal co-articulation phenomena are specific co-articulation phenomena for languages using tones. In Thai language [6], there are two types of tonal co-articulation: anticipatory and carryover co-articulation. The F0 contour shape of a syllable that is influenced by the succeeding syllable is called "anticipatory co-articulation". The carryover co-articulation occurs when the preceding syllable influences the succeeding syllable. Gandour [6] reported that the co-articulation effect of Thai tones is asymmetric. Thai tones were more influenced by carryover than by anticipatory co-articulation. The tonal co-articulation in Mandarin is also classified into both types: anticipatory and carryover co-articulation [7][8]. In Shen's study [7], nonsense strings of syllables were used as reading list, whereas Shir and Sproat [8] used real words and phrases. In Shir's study speakers were instructed to produce the utterances as naturally as possible. In order to study phenomena involved in co-articulation of tones in Mandarin, Xu Y [9] provided 4 theoretically possible transitional patterns between two tones. Similar to the results of [8], two in four possible tonal transitions (Exclusive Assimilation and Exclusive Carryover) were found for Mandarin language.

In Vietnamese, the studies [1][4] and [5] show that the progressive tonal co-articulation is stronger than the regressive tonal co-articulation. In these studies, authors focused on analysis of tonal co-articulation between two adjacent syllables, they have not take into account information about lexical category (or part of speech) of considered words.

In order to study the influence of lexical category in phenomena of tonal co-articulation, this paper presents our study on tonal co-articulation of particles (such as 'ạ', 'dạ', 'chứ', 'đi' …) which are frequently used in Vietnamese language. The paper is organized as follows. Section II introduces the speech corpus which is used in the analysis process. Section III gives the analysis results on the tonal co-articulation phenomenona which occur at the considered particles. Section IV gives some conclusions and perspectives.

2   Bi-tone speech corpus

A speech corpus of bi-tone pairs of two adjacent syllables was prepared in our experiment. In addition, each sentence in the text corpus is constructed of an affirmative sentence and a particle at the last position of the sentence, for instance: "*Lâu lâu mới gặp lại, chị vào nhà tôi chơi / nhé!*"(We haven't met each other for a long time, let's come to my house). In the sentence the word "*nhé*" is a particle and "*Lâu lâu mới gặp lại, chị vào nhà tôi chơi*" is an affirmative sentence. The bi-tone pairs in these sentences are the last two syllables. For example, in the

sentence *"Trước mẹ con vẫn gắp cho con kia!"* (Figure 1) the last two syllables *"con kia"* are analyzed. The corpus of these bi-tone pairs help to clarify the effects which the tonal context or the lexical context causes.
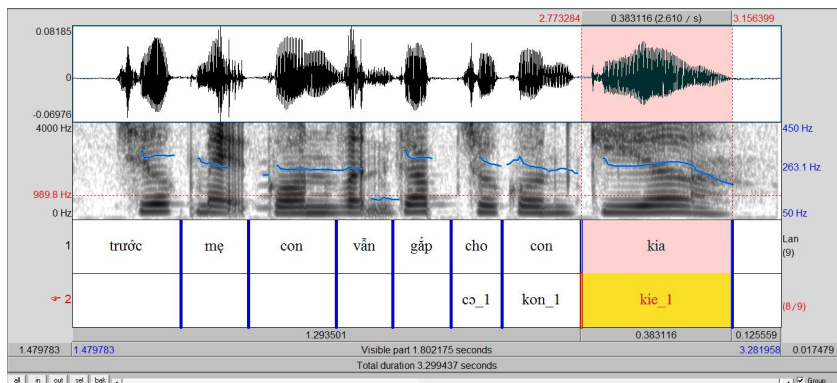

Figure 1: F0 contour of the sentence *"Trước mẹ con vẫn gắp cho con kia!"*

Therefore, a text corpus including 144 sentences in dialogs which were extracted from 6 Vietnamese famous short stories *"Bóng mây chiều"*, *"Dế mèn"*, *"Tuyển tập truyện ngắn Hồ Dzếnh"*, *"Kính vạn hoa"*, *"Tuyển tập truyện ngắn Thạch Lam"* và *"Tuổi 20 yêu dấu"* was collected.

These 144 sentences are divided into 6 groups as follows:
- *Group 1*: 24 sentences contain particles carrying the level tone (Tone 1).
- *Group 2*: 24 sentences contain particles carrying the falling tone (Tone 2).
- *Group 3*: 24 sentences contain particles carrying the broken tone (Tone 3).
- *Group 4*: 24 sentences contain particles carrying curve tone (Tone 4).
- *Group 5*: 24 sentences contain particles carrying the rising tone (Tone 5).
- *Group 6*: 24 sentences contain particles carrying the drop tone (Tone 6).

Each group is divided into 6 subgroups based on the tone (6 tones) which is carried by the preceding syllable of the particle. Each subgroup contains 4 sentences. For example:

Group 1 is composed of 6 subgroups: S1-1, S1-2, S1-3, S1-4, S1-5 and S1-6. In the subgroup S1-1, the sentences have particles carrying level tone and the preceding syllable also carrying level tone.
- *"Trước mẹ con vẫn gắp cho **con / kia**!"(In meal my mother used to pick food for me)*
- *"Có một việc quan trọng phải nhờ cậu mới **xong / đây**!"(There is an important work that we must need your help)*

The text corpus is then used to build a speech corpus. The recording progress was taken place in a quite environment at a quiet studio. A male and female who have Hanoi voice were asked to read each sentence two times at a normal speaking rate. The speech corpus was then labeled manually using the PRAAT program.

3   Analysis

We carry out an analysis of the variation of fundamental frequency for every subgroup of sentences. F0 values of the sentences are extracted automatically by using Praat software.
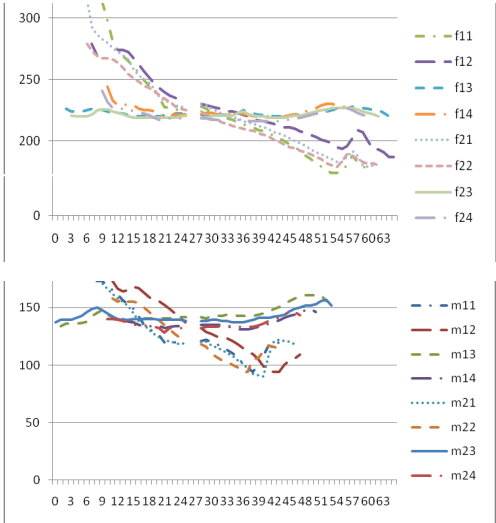
Figure 2 presents evolution of F0 contours in subgroup S2-2 (which has a pair of tones: falling – falling). In the figure *m* and *f* refer consequently "male" and "female"; m*ij* denotes the i*th* recorded sample of the j*th* sentence of the male speaker in the group; and f*ij* denotes the i*th* recorded sample of the j*th* sentence of the female speaker.

We can see in Figure 2, the F0 contours of the preceding falling tones fall slowly and converge at the ending position. However, values of F0 at the beginning positions of these syllables are quite different. By carrying a larger analysis, we found that the difference is caused by an effect of tonal co-articulation. The tones of the syllables that stand just before the analyzed syllables are different. For example, concerning f11 and f14 sentences, the three last syllables of these sentences are "có tiền à" (/kɔ_5 tien_2 a_2/) and "nhiều tiền mà" (/ɲiew_2 tien_2 ma_2/) consequently. It is easy to see that the last two syllables have the same pair of tones (falling – falling), but the tone of the 1st syllable in case of f14 is the falling tone (Tone 2), while the tone of the 1st syllable in case of f11 is the rising tone (Tone 5). Thus, F0 values of the beginning positions in sentence f11 are higher than that of sentence f14.

Concerning the falling tone carried by the particle words, in Figure 2 we found that, the shape of F0 contours of this tone has a little difference in comparison to that in the static mode. The F0 contours of these syllables fall slowly at the start positions and then rise to the ending positions. This can be explained as follows: a particle is also the last syllable of an imperative or a question, so its F0 contour usually rises up, like the results of studies [10][11].

To observe the tonal co-articulation more clearly, similar to the research of Brunelle [1], we carried out comparing an effect of the preceding syllable's tone on the succeeding syllable's tone. In order to facilitate the analysis, the values of the extracted F0 are normalized into a sequence of six points. Each syllable is divided into 6 parts and the average value of F0 is calculated for each part. Therefore, for a syllable pair, points from 0 to 5 are belong to the first syllable (the last syllable of the affirmative sentence), and points from 6 to 11 are belong to the second syllable (a particle).

4   The results

The normalized F0 contours of the subgroups of sentences are presented in figures 3, 4, 5, 6, 7, 8. In these figures:

- t1 presents the average normalized F0 contour of the set of sentences that have the preceding syllable (the syllable stands before the particle) porting the level tone (Tone 1).
- Similar to t1, t2 – t6 present the average normalized F0 contours of sets of sentences that have the preceding syllable carrying tones 2, 3, 4, 5 and 6 consequently.

From these figures, the first result can be easily seen that: tonal co-articulation phenomenon always happens at the transition between two adjacent syllables, and the progressive co-articulation seems the basic tonal co-articulation, like the results presented in [1][4][5].

We pay more attention to representations of Tone 1 and Tone 2 which are presented in the figures 3 and 4.
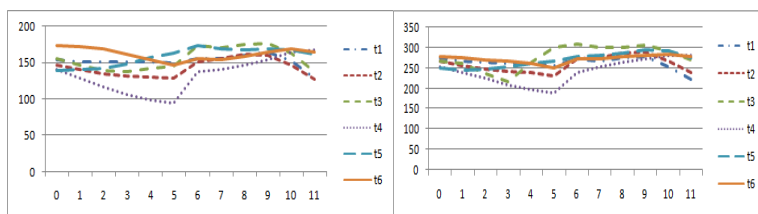

Figure 3: F0 contours of tone pairs in Group 1 (particle carries level tone) with male (left) and female (right) speakers
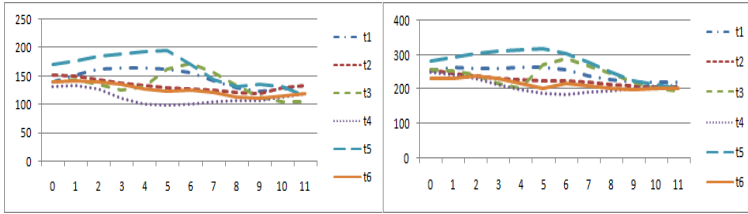
Figure 4: F0 contours of tone pairs in Group 2 (particle carries falling tone) with male (left) and female (right) speakers

In Figure 3, after ending the preceding syllable, F0 contours of Tone 1 (level tone) tend to rise to the end of succeeding syllables and it only fall lightly at the last point. Whereas, in studies [1][11], the F0 contour of Tone 1 is quite stable and falls lightly down to the end of the succeeding syllable (same as Tone 1 in isolated syllable).

Like the particles carrying Tone 1, we can see clearly the same result in case of the particles porting Tone 2 (falling tone). In Figure 4, F0 contour of this tone can start at either low or high register. At the last half of syllable F0 contour either is stable or rises slightly. Whereas, in static mode, Tone 2 starts lower than the level tone then falls slowly and its slope is never goes up. In the studies of [1][11], in dynamic mode (continuous speech), F0 contours of Tone 2 falls slowly. This results shows that there is an affect of the function of particles on evolution of F0. The result seems logic, because the particles in our corpus stand at the last position of sentence. The function of these words is to make the sentence become either an interrogative or an imperative sentence. According to studies [10][11], in Vietnamese interrogative and imperative sentences, the contour of the last syllable or of its second half tends to increase.
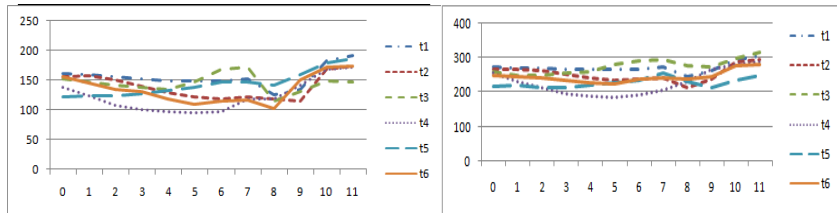

Figure 5: F0 contours of tone pairs in Group 3 (particle carries broken tone) with male (left) and female (right) speakers
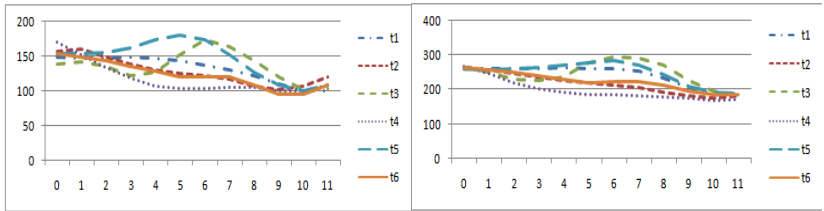
Figure 6: F0 contours of tone pairs in Group 4 (particle carries curve tone) with male (left) and female (right) speakers
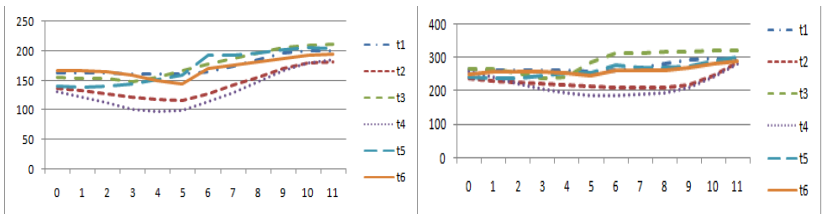

Figure 7: F0 contours of tone pairs in Group 5 (particle carries rising tone) with male (left) and female (right) speakers
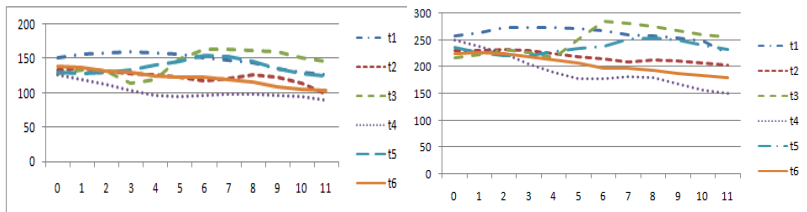

Figure 8: F0 contours of tone pairs in Group 6 (particle carries drop tone) with male (left) and female (right) speakers

In remaining tones, the phenomenon in which F0 contour of the last half of particle rises up also takes place. However, phenomenon is caused not only by the function of particles but also by the evolution of F0 contour of the tone that syllables carry. These tones (Tone 3, Tone 4 and Tone 5) have the F0 contour rising at the last half of syllable.

The standard deviation of 6 tones at 6 points is presented in following tables:

Table 1: Standard deviation (%) of tones at 6 points (male)

Point	P1	P2	P3	P4	P5	P6
Tone 1	8.1	6.6	5.6	4.5	5.5	11.5

169

Point	P1	P2	P3	P4	P5	P6
Tone 2	18.0	12.9	8.1	7.6	7.8	8.5
Tone 3	18.0	14.9	9.6	10.4	6.2	7.8
Tone 4	19.1	15.4	12.1	7.6	3.7	5.2
Tone 5	18.0	14.1	10.5	7.7	6.0	5.3
Tone 6	17.7	17.3	16.2	16.2	15.3	17.2

Table 2 : Standard deviation (%) of tones at 6 points (female)

Point	P1	P2	P3	P4	P5	P6
Tone 1	7.6	5.3	4.3	3.85	5.26	8.73
Tone 2	17	13	9.1	4.66	3.22	4.09
Tone 3	12	11	8.2	8.04	7.03	7.12
Tone 4	16	16	14	9.12	5.08	2.56
Tone 5	17	16	16	14.6	10	4.55
Tone 6	17	16	16	16.5	17.1	16.7

We can see that the standard deviation values at initial points (P1, P2 and P3) which are close to the preceding syllable are almost higher than that's of the final points (P4, P5 and P6). This is entirely consistent with above results.

The variation of F0 of Tone 6 is quite different in comparison with other tones. Because of the glottalization phenomenon at ending position, F0 contour at the second last syllable of syllables porting this tone changes in large range and this can make a non precise measurement of F0 by Praat. Therefore, value of standard deviations is quite high.

Conclusion and perspectives

The paper presents our research on phenomena of tonal co-articulation on particles in Vietnamese. The obtained results show that tonal co-articulation phenomena always take place at the transition between two adjacent syllables, and the progressive co-articulation is the basic tonal co-articulation. And *there is an influence of the function of particles on form of F0 contour of Vietnamese tones.* Therefore, it has to take into account this information when doing an analysis variation of fundamental frequency, especially on modeling Vietnamese intonation for Vietnamese text to speech synthesis.

In our study, the corpus is still limited. There are only two speakers and the particles always stand at the last position of sentences. These limitations will be studied more deeply in next researches with more speakers and other specific words at different positions.

Acknowledgments

References

[1] Brunelle M. "Co-articulation effects in Northern Vietnamese tones", Proc. ICPhS, Barcelona, pp. 2673 – 2676, 2003.

[2] Nguyen, Q.C, "Reconnaissance de la parole en langue Vietnamienne", PhD. thesis INP-Grenoble, France, June 2002.

[3] Tran D.D, Castelli E., Serignat JF., Le X.H., Trinh V.L., "Influence of F0 on Vietnamese syllable perception", Proc. of Interspeech2005, Lisbon, pp. 1697-1700, 2005.

[4] Han M.S. and Kim K., "Phonetic variation of Vietnamese tones in disyllabic utterances", Journal of Phonetics April 1974, pp.223-232, 1974.

[5] Tran D.D, Castelli E., "Generation of F0 contours for Vietnamese speech synthesis", Third International Conference on Communications and Electronics (ICCE), Nha Trang, 2010.

[6] Gandour, J., Potisuk, S., & Dechangkit, S. (1994). "Tone co-articulation in Thai,Journal of Phonetics", 22, 477-492.

[7] Shen, X. S. (1990). "Tonal co-articulation in Mandarin. Journal of Phonetics", 18,281-295.

[8] Shih, C., & Sproat, R (1992). "Variations of the Mandarin rising tone". In Proceedings of the IRCS Workshop on Prosody in Natural Speech No. 92-37, (pp. 193-200). Philadelphia: The Institute for Research in Cognitive Science, University of Pennsylvania.

[9] Xu Y., "Contextual tonal Variations in Mandarin", Journal of Phonetics 25, 61-83, 1997.

[10] Nguyen, T. T. H. 2004, Contribution à l'étude de la prosodie du vietnamien. Variations de l'intonation dans les modalités: assertive, interrogative et impérative, PhD. thèses, Doctorat de Linguistique Théorique, Formelle et Automatique, Paris.

[11] Vu M.Q., Tran D. D., Castelli E. 2006, Prosody of Interrogative and Affirmative Sentences in Vietnamese Language: Analysis and Perceptive Results, The Ninth International Conference on Spoken Language Processing – INTERSPEECH 2006 - ICSLP, Pittsburgh, Pennsylvania, USA, September 2006.

[12] Tran D.D. Castelli E., Serignat J.F., Trinh V.L. & Le X.H., "Linear F0 Contour Model for Vietnamese Tones and Vietnamese Syllable Synthesis with TD-PSOLA", Proc. TAL2006, La Rochelle, April 2006.