

AN INITIAL ANALYSIS ON THE INTERACTIONS OF VIETNAMESE LINGUISTIC TONES & VIETNAMESE FOLK MUSIC¹

Phan Gia Anh Thu²

Columbia University Teachers College & Temple University, agp2132@tc.columbia.edu
&

Ngô Thanh Nhân³

New York University, Temple University & Folk Arts–Cultural Treasures Charter School

Abstract

This paper describes the interaction between the six Vietnamese *thanh* “tones” in the lyrics and melodies of popular Vietnamese folk songs. Based on a universal platform—strings of sounds, of language and music entwined in folk songs, a theoretical framework is found where Western music theories, Vietnamese music theories, phonology, and Vietnamese linguistics can interoperate. The investigation begins with an automated *notated pitch/intensity contour* (nPIC) graph of a sung folk song, from a peak frequency spectrogram, on which, onsets of syllables (i.e. words) are identified. This is called a “*syllamelis*” (plural, *syllameles*). On this nPIC frame a music staff is overlaid. In order to exhaust the interactions of the linguistic tones and the melody, four general concepts are advanced: a *syllamelis*, a *toneume*, adjacency, and congruence (i.e. how well lyric tones are realized in the melody). A “*toneume*” (tone+neume) is a unit of tone within a *syllamelis*, established by the distinctive linguistic features of tone pitches. The inter-toneume comes into play with the need for pitch behavior between two adjacent speech tones. The complex congruence relationships are found between (1) speech tone of different dialects and the melodic toneumes, (2) two adjacent toneumes and their corresponding speech inter-tone behavior, (3) general speech tone pitches and singer’s idiolectal toneumes in the song, and (4) performed *rung* “*vibrato*” and tones, and the existing ambiguous definitions of Vietnamese musical modes / airs presented in scholar literature. The analysis is based on 412 *syllamelic* nPICs generated from 7 recordings of singing and 6

¹ The authors thank the *Mekong Traditional Vietnamese Instrumental Ensemble Troupe* (MTVIET) for the ideas coming from an MTVIET discussion on March 4, 2016.

² *Music Education*, Teachers College - Columbia University, and *Fellow* of Temple University Center for Vietnamese Philosophy, Culture & Society. Email: agp2132@tc.columbia.edu.

³ New York University, *Linguistic String Project*, Temple University *Center for Vietnamese Philosophy, Culture & Society*, and Folk Arts—Cultural Treasures Charter School, *Đàn Tranh Ensemble*. Email nhan@cs.nyu.edu & nhan@temple.edu.

readings of the same southern Vietnamese folk song, *Ru Con miền Nam* “Lullaby”. The results reveal that toneumes do manifest themselves distinctly in their syllables, leaving space for inter-toneume expression and the artist’s idiosyncratic interpretation. For a tonal language, the lyric vs toneume/syllable congruence is most favored whereas incongruences may serve to enhance the idiosyncratic creativity of the artists. The method proposed in this paper: the nPIC with syllable marking and the behavior of tones and vibratos in the melody, helps teachers to approach widely available music and songs online and prepare lessons quickly and accurately with more challenging exercises. It also helps students everywhere to understand the inner structures of language and music.

Key words: notated pitch/intensity contour, key pitches in context, tone, congruence, inter-tone, syllable, toneume, *vibrato*.

Paper’s ID: VS2.032P1

A. Introduction

Teaching Vietnamese traditional and folk music to children is a challenge to any teacher, new or experienced, the authors included. Teachers, who are in-demand, have to improvise teaching tools to compensate for the shortage of basic research in Vietnamese traditional and folk music.

The authors are volunteers teaching music to members of the Cambodian and Vietnamese communities in the Bronx. The communities requested that a number of popular songs that are appreciated and listened to be taught. This request motivates the authors to find ways to address the need for new pedagogical methods.

In a previous study, Ngô & Phan (2016), the authors developed an evidence-based approach by introducing the concept of *notated pitch/intensity contour vs time graph* (nPIC), a direct visual peak frequency spectrogram of pitch/intensity contour over time representation in one folk song recording, to expose characteristic pitches of Vietnamese music on a western staff. We further introduced the concept of *frequencies of occurrence* of pitch onsets of the song, as well as the *key pitches in context* of adjacent pitches. This simple process allows the lesson to be prepared and taught quickly, and surprisingly confirms the intuitive characterizations by grandmasters of Vietnamese music. This paper addresses a frequently asked question: the interactions of speech tones and melody in Vietnamese folk songs.

B. Theoretical Framework

Tones in a song lyric, manifested in pitch, and the pitch contour of the song melody are sung by one voice. This observation lays the basis for an interdisciplinary approach: combining Western and Vietnamese music theories with Western and Vietnamese linguistic theories.

Introduction to Vietnamese linguistic tones

Vietnamese is a tonal, monosyllabic language, where each syllable, *tiếng*, is independent. No morpheme or word is smaller than a syllable. Each syllable consists of one tone, *thanh*, and a segmental syllable consisting of an onset and a rhyme, *vần*.

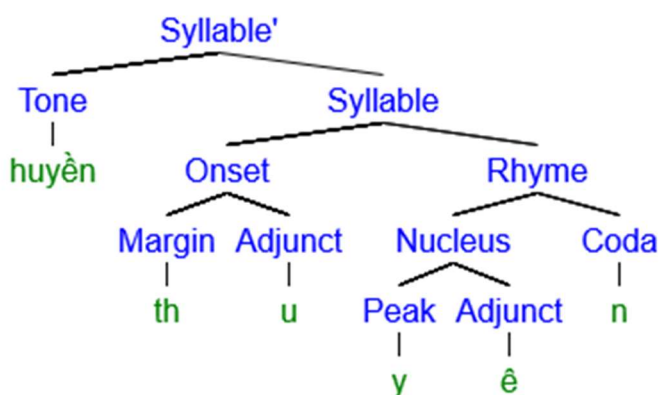


Chart 1: The Vietnamese syllable structure

The onset margin is obligatory, having one of 23 consonants (glottal stop /ʔ/ included), and an optional onset adjunct of a semivowel /w/.⁴ The rhyme has a nucleus of 13 vowels and 3 diphthongs. The coda can be one of the 13, either open or one of the 2 semivowels /w/ or /y/, or one of 3 stops and 3 nasal stops. The Vietnamese syllable structure rules, associated with Chart 1, generate only

over 14,000 well-formed syllables that sound perfectly Vietnamese. However, Vietnamese only use 6,979 of them. For example, **đạ* sounds Vietnamese but is not found in the dictionary.⁵

Vietnamese has 6 linguistic tones, *thanh*, which are briefly described with their traditional ordering and spelling frequencies of occurrence (fo.) in square brackets,⁶ the orthography, tone pitches and tone letters.

	<i>bằng</i> “level”	<i>trắc</i> “contour”	
<i>bổng</i> [high]	<i>ngang</i> “high level” Orthography: ◯ Tone order: 1 [fo. 1333] Tone pitch / ⁴⁴⁴ / Tone letter: 1	<i>hỏi</i> “low rising” Orthography: ◌̣ Tone order: 3 [fo. 830] Tone pitch : / ²¹⁴ / Tone letter: 3 4	<i>sắc</i> “high rising” Orthography: ◌̣̣ Tone order: 5 [fo. 1759] Tone pitch : / ³⁴⁵ / Tone letter: 1

⁴ Transcribed in this paper in International Phonetic Association’s convention, IPA for short.

⁵ cf. Ngô, 1984, Chapter 3, *Remarks on the phonological description of Vietnamese*.

⁶ cf. Ngô, 1984, p. 78.

<i>trầm</i> [low]	<i>huyền</i> “low level” Orthography: ̀ Tone order: 2 [fo. 1157] Tone pitch: / ²²² / Tone letter: ˉ	<i>ngã</i> “creaky rising” Orthography: ̣ Tone order: 4 [fo. 486] Tone pitch: / ⁴¹⁵ / Tone letter: ˋ ˊ	<i>nặng</i> “creaky falling” Orthography: ˘ Tone order: 6 [fo. 1414] Tone pitch: / ²¹¹ / Tone letter: ˋ
----------------------	---	--	---

Table 1. Vietnamese tones and their representations.

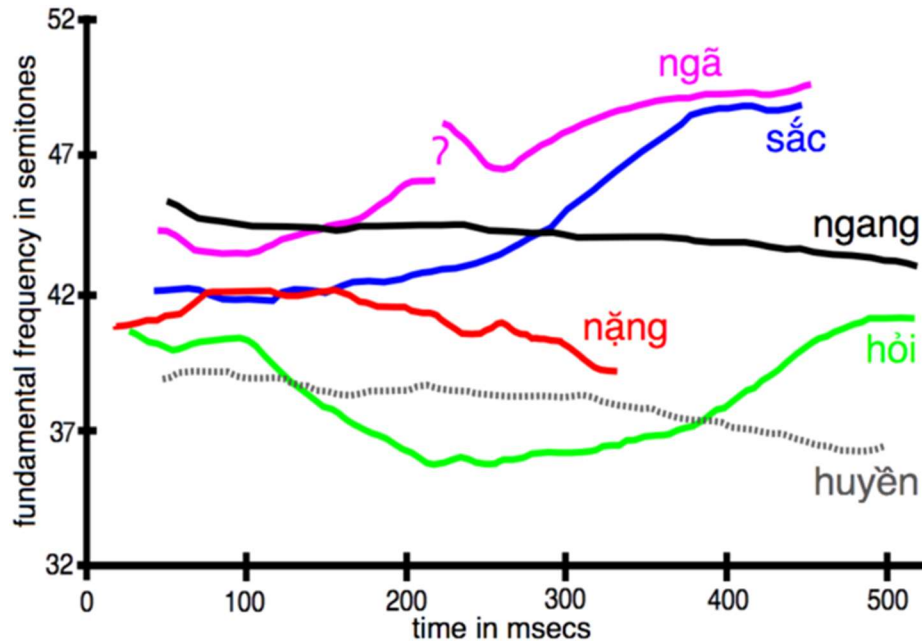


Figure 1. An example of Vietnamese tone in fundamental frequency vs. time.⁷

Note that tone *ngã* “creaky rising” is lost in the Central and Southern dialects of Vietnam. The tone *ngã* is pronounced like tone *hỏi*.

In this paper, we choose tone pitch representation for the lyric, for ease of exposition of the relationship between tone and melody. Phonetically, the Vietnamese tone begins 100 *ms.* into the syllable, after the onset at the rhyme, and is about 500 *ms.* long.⁸

There is no tone sandhi in Vietnamese. The monosyllabicity of Vietnamese can also be described as tones of two adjacent syllables that form two discrete and distinct pitch events. However, in normal speech, adjacent tones require a transition between two tone pitches: between one tone offset and the next tone onset, which can be called a phonetic

⁷ Pitch contours and duration of the six Northern Vietnamese tones as spoken by a male speaker (not from Hanoi). Fundamental frequency is plotted over time. From Nguyễn & Edmondson (1998).

⁸ According to Han & Kim (1974) and Tuệ and Minh (1976, pp. 72-73), the initial pitch of the tone picks up the level of pitch left off by the preceding tone. They further report that the characteristic pitch shape of each tone is only realized at about 100 *ms.* into the syllable, passing the onset and starting at the vocalic peak and terminating about 500 *ms.* later (the duration of the syllable is about 500 *ms.*). For this reason, they suggest that at the normal rate of speech, each tone is at least 100 *ms.* separate from the other, and is safe for tone sandhis of any kind.

inter-tone.⁹ This normally occurs at the coda of the first syllable and the onset of the second syllable where the pitch gap has to be accommodated. For example, in the folk saying, *chị ngã em nâng* “the older sister falls, the younger sibling helps her up,” the end of tone *nặng* /²¹¹/ in *chị* “older sister” is [1], and the beginning of tone *ngã* /⁴¹⁵/ in *ngã* “fall” is [4], the inter-tone is [14] (the gap is +3); the end of tone *ngã* /⁴¹⁵/ in *ngã* is [5], and the beginning of tone *ngang* /⁴⁴⁴/ in *em* “younger sibling” is [4], the inter-tone is [54] (the gap is -1), and so on. This implies that in speech, the inter-tone acts like a pitch glue that helps in smoothing the contour of the entire phrase.

C. Methodology

The methodology begins with the *notated pitch/intensity vs time contour* graph (nPIC), of the peak frequency spectrogram of a recorded song, on which a music staff is drawn, exposing the music notes, as presented in Ngô & Phan (2016). The nPIC is generated by *Sonic Visualiser*.¹⁰ This allows *key pitches in context* (kpic), i.e. how often 2 or 3 or 4 or more adjacent pitches occur in the melody, to bring out pitch patterns of a music piece.

In order to formalize the analysis in this paper, a set of mappings of the lyric syllables and their tones in the melody is introduced. That is, the concepts of a syllamelis and a toneume. Both are identified using the lyric.

C.1. A syllamelis (syllabic+neume+melisma,¹¹ plural, *syllameles*, Sanskrit unit symbol ष /sa/) is a lyric syllable sung in the melody. It is identified by an automatic onset detection on the melodic nPIC (in this paper, the *Aubio Onset Detector*¹² is used), narrowed down by a heuristic manual matching with the lyric syllable. It is marked by a point in time and a pitch for each syllable onset on nPIC. This process can be automated.

C.2. A toneume (tone+neume, Sanskrit unit symbol ष /ta/) is a lyric tone sung in the melody. It is identified by the distinctive linguistic pitch features of the tone, demarcated by a range of time points on the melodic nPIC.

C.3. An inter-toneume is identified in two adjacent syllameles by one toneume coda and the next toneume onset.

⁹ Rodger, M.W.M & Craig, C.M. “Moving with beats and loops,” p. 212.

¹⁰ As described in the About *Sonic Visualiser* page, a freeware program “for viewing and exploring audio data for semantic music analysis and annotation.”

¹¹ In music, a syllabic is a syllable sung in one note, a neume in 1-4 notes, and a melisma, in more than 1 note.

¹² *Aubio Onset Detector*, v.2.

These three concepts help to lay the groundwork to describe the interactions between linguistic tones in song melodies.

C.4. Congruence: This paper defines the concept of *congruence*, as to how well linguistic tones in lyrics are realized in the melody, i.e. between each tone pitch and its corresponding toneume, between lyric inter-tones and their melodic inter-toneumes, and the behavior of *vibratos*, spikes and dips on the toneumes and the inter-toneumes.

D. An example: A southern Vietnamese folk lullaby *Ru Con miền Nam*

For this paper, *Ru con miền Nam* “Lullaby from the South”¹³ was suggested by a community member, and is chosen to illustrate these procedures. The song is sung by Ms. Bích Tuyền, accompanied by the *đàn bầu* monochordist Hoàng Thịnh, (date unknown). A *dạo* “promenade” of *đàn bầu* takes up the first 30 seconds, and recital of the *ca dao* “folk poetry” of four six-eight syllable verses takes up the next 1:50 minutes before the main song continues for another 1:23:20 minutes. The main song was extracted in mp3 format. We call this extracted recording *Ru con*, for short.

There are different lyric versions for this lullaby,^{14,15} nevertheless, the folk version is still a masterpiece and is the one most favored by the community.

	Tentative translation:
“ <i>Gió mùa thu... mẹ ru mà con</i>	1. Autumn breeze helps Mother to lull her baby
<i>ngủ...</i>	to sleep.
<i>Năm ... canh dài...</i>	2. Five times the timekeeper had struck
<i>Năm ... canh dài,</i>	3. for ten hours straight...
<i>thức đủ vừa năm...</i>	4. I have been up all ten.
<i>Hỡi chàng... chàng ơi!</i>	5. Oh, lover... please, lover!
<i>Hỡi người... người ơi!</i>	6. Oh, man... please, man!
<i>Em nhớ tới chàng.</i>	7. I am thinking of you,
<i>Em nhớ tới chàng!</i>	8. I really miss you!
<i>Hãy nín... nín đi, con!</i>	9. Hush... don't cry, baby!
<i>Hãy ngủ... ngủ đi, con!</i>	10. Go to sleep... sleep well, baby!
<i>Con hời mà con hời!</i>	11. Oh child, please hush, oh baby!
<i>Con hời, con hời...</i>	12. Oh child, please hush, oh baby!
<i>Con hời, con hời, hời con!”</i>	13. Baby, please baby, oh please!”

¹³ Cf. <https://www.youtube.com/watch?v=xGHLTQzib3Y>.

¹⁴ 3 versions can be found at <http://dotchuoinon.com/2015/01/18/dan-ca-dan-nhac-vn-hat-ru-con-mien-nam/>.

¹⁵ 4 versions of four seasons can be found at <http://www.tuanpham.org/EnglishLyricsFull.htm>.

There are 59 tones in the lyric, with their corresponding 59 syllables in the nPIC. There are 23 tone *ngang* /⁴⁴⁴/ “high level”, 15 tone *huyền* /²²²/ “low level”, 12 tone *hỏi/ngã* /²¹⁴/ “concave tone”, 8 tone *sắc* /³⁴⁵/ “high rising”, and 1 tone *nặng* /²¹¹/ “creaky falling”. The IPA transcription of the song with tone pitches is shown below, in Table 2, where tone *ngã* “creaky rising,” tone pitch /⁴¹⁵/, merges with tone *hỏi* “low rising,” at tone pitch /²¹⁴/:

1.1 <i>gió</i> /yɔ ³⁴⁵ / [yɔ ⁴⁴⁴ ɔ ³⁴⁵]	1.2 <i>mùa</i> /muə ²²² / [muə ²²² ə ⁴⁴⁴]	1.3 <i>thu</i> /t ^h u ⁴⁴⁴ / [t ^h u ³⁴⁵ u ⁴⁴⁴]	1.4 <i>mẹ</i> /mɛ ²¹¹ / [mɛ ²¹¹]	1.5 <i>ru</i> /ru ⁴⁴⁴ / [ru ⁴⁴⁴]	1.6 <i>mà</i> /ma ²²² / [ma ²²²]	1.7 <i>con</i> /kɔn ⁴⁴⁴ / [kɔn ²²² ə ⁴⁴⁴]	1.8 <i>ngũ</i> /ŋu ²¹⁴ / [ŋu ²²² u ³⁴⁵ u ⁴⁴⁴]
2.1 <i>nấm</i> /næm ⁴⁴⁴ / [næm ⁴⁴⁴ ə ²²² ə ³⁴⁵]	2.2 <i>canh</i> /kɛŋ ⁴⁴⁴ / [kɛŋ ⁴⁴⁴]	3.3 <i>dài</i> /yay ²²² / [yay ²²²]					
3.1 <i>nấm</i> /næm ⁴⁴⁴ / [næm ⁴⁴⁴ ə ²²² ə ³⁴⁵]	3.2 <i>canh</i> /kɛŋ ⁴⁴⁴ / [kɛŋ ⁴⁴⁴]	3.3 <i>dài</i> /yay ²²² / [yay ²²²]					
4.1 <i>thức</i> /t ^h ʊk ³⁴⁵ / [t ^h ʊk ³⁴⁵]	4.2 <i>đủ</i> /du ²¹⁴ / [du ²²² u ³⁴⁵]	4.3 <i>vừa</i> /vʊə ²²² / [vʊə ²²² ə ⁴⁴⁴]	4.4 <i>nấm</i> /næm ⁴⁴⁴ / [næm ³⁴⁵ ə ⁴⁴⁴]				
5.1 <i>hỏi</i> /hɔy ²¹⁴ / [hɔ ²²² ɔy ³⁴⁵]	5.2 <i>chàng</i> /tʃaŋ ²²² / [tʃaŋ ⁴⁴⁴ ə ²²²]	5.3 <i>chàng</i> /tʃaŋ ²²² / [tʃaŋ ³³³ ə ⁴⁴⁴]	5.4 <i>ơi</i> /ʔɔy ⁴⁴⁴ / [ʔɔy ³⁴⁵ ə ⁴⁴⁴]				
6.1 <i>hỏi</i> /hɔy ²¹⁴ / [hɔ ²²² ɔy ³⁴⁵]	6.2 <i>người</i> /ŋuəy ²²² / [ŋuəy ⁴⁴⁴ ə ²²²]	6.3 <i>người</i> /ŋuəy ²²² / [ŋuəy ³³³ ə ⁴⁴⁴]	6.4 <i>ơi</i> /ʔɔy ⁴⁴⁴ / [ʔɔy ³⁴⁵ ə ⁴⁴⁴]				
7.1 <i>em</i> /ʔɛm ⁴⁴⁴ / [ʔɛ ⁴⁴⁴ əm ⁴³]	7.2 <i>nhớ</i> /ɲɔ ³⁴⁵ / [ɲɔ ²²² ə ³⁴⁵]	7.3 <i>tới</i> /tɔy ³⁴⁵ / [tɔ ²²² ɔy ³⁴⁵]	7.4 <i>chàng</i> /tʃaŋ ²²² / [tʃaŋ ²²²]				
8.1 <i>em</i> /ʔɛm ⁴⁴⁴ / [ʔɛ ⁴⁴⁴ əm ⁴³]	8.2 <i>nhớ</i> /ɲɔ ³⁴⁵ / [ɲɔ ²²² ə ³⁴⁵]	8.3 <i>tới</i> /tɔy ³⁴⁵ / [tɔ ²²² ɔy ³⁴⁵]	8.4 <i>chàng</i> /tʃaŋ ²²² / [tʃaŋ ²²²]				
9.1 <i>hãy</i> /hɛy ²¹⁴ / [hɛ ²²² y ³⁴⁵]	9.2 <i>nín</i> /nin ³⁴⁵ / [nin ⁵⁵⁵]	9.3 <i>nín</i> /nin ³⁴⁵ / [nin ⁵⁵⁵]	9.4 <i>đi</i> /di ⁴⁴⁴ / [di ⁴⁴⁴]	9.5 <i>con</i> /kɔn ⁴⁴⁴ / [kɔn ³⁴⁵ ə ⁴⁴⁴]			
10.1 <i>hãy</i> /hɛy ²¹⁴ / [hɛ ²²² y ³⁴⁵]	10.2 <i>ngũ</i> /ŋu ²¹⁴ / [ŋu ²²² u ³⁴⁵]	10.3 <i>ngũ</i> /ŋu ²¹⁴ / [ŋu ²²² u ³⁴⁵]	10.4 <i>đi</i> /di ⁴⁴⁴ / [di ⁴⁴⁴]	10.5 <i>con</i> /kɔn ⁴⁴⁴ / [kɔn ³⁴⁵ ə ⁴⁴⁴]			

11.1 <i>con</i> /kɔ̃n ⁴⁴⁴ / [kɔ̃n ⁴⁴⁴ ə ³³³]	11.2 <i>hòi</i> /hɔ̃y ²²² / [hɔ̃y ²²²]	11.3 <i>mà</i> /ma ²²² / [ma ²²²]	11.4 <i>con</i> /kɔ̃n ⁴⁴⁴ / [kɔ̃n ²²² ə ⁴⁴⁴]	11.5 <i>hòi</i> /hɔ̃y ²¹⁴ / [hɔ̃ ²²² y ³⁴⁵ ə ⁴⁴⁴]
12.1 <i>con</i> /kɔ̃n ⁴⁴⁴ / [kɔ̃n ⁴⁴⁴]	12.2 <i>hòi</i> /hɔ̃y ²¹⁴ / [hɔ̃ ⁴⁴⁴ y ³⁴⁵]	12.3 <i>con</i> /kɔ̃n ⁴⁴⁴ / [kɔ̃n ⁴⁴⁴ ə ³³³]	12.4 <i>hòi</i> /hɔ̃y ²²² / [hɔ̃y ²²²]	
13.1 <i>con</i> /kɔ̃n ⁴⁴⁴ / [kɔ̃n ⁴⁴⁴]	13.2 <i>hòi</i> /hɔ̃y ²¹⁴ / [hɔ̃ ²²² y ³⁴⁵]	13.3 <i>con</i> /kɔ̃n ⁴⁴⁴ / [kɔ̃n ⁴⁴⁴ ə ³³³]	13.4 <i>hòi</i> /hɔ̃y ²²² / [hɔ̃y ²²²]	13.5 <i>hòi</i> /hɔ̃y ²¹⁴ / [hɔ̃ ³³³ y ³⁴⁵]

Table 2: Syllables in *Ru Con miền Nam*, and their phonemic and phonetic descriptions.

D.1. Notated pitch/intensity contour graph (nPIC)

The spectrograms of the *Ru con* singing and reading inputs are graphed with values on a vertical axis showing frequencies in *Hertz* (Hz), pitches in *Ellis cents* (c) colored with intensity in *decibels* (dB), and values on a horizontal axis indicating time in *seconds* (sec). This is called the *pitch/intensity contour* (or PIC) of the song. Specifically, the following set of figures is generated by *Sonic Visualiser*. For example, at point 9.139 sec into the recording, the peak frequency spectrogram pane shows the following singing data, with the peak pitch ranging from $F\sharp^3-18c$ (18c under $F\sharp^3$) to $F\sharp^3+2c$ (2c over $F\sharp^3$):

at Time range: 9.139 — 9.233 sec

Peak Frequency: 183.2 — 185.187 Hz — Bin Frequency: 172.266 — 183.032 Hz

Peak Pitch: $F\sharp^3-18c$ — $F\sharp^3+2c$ — Bin Pitch: F^3-23c — $F\sharp^3-18c$

dB: -36 — -29 — Phase: -1.63752 — 2.1971

D.2. Syllable onset assignments—syllameles

Over the nPIC, *Aubio Onset Detector* is run “to detect onset times, the beginning of discrete sound events, in audio signals.” Each of the lyric syllable onsets were narrowed down to less than 5 peaks, and settled with a heuristic matching. The syllamelis is identified by its onset time and its pitch on the nPIC (drawn by a vertical purple line). Thus, there are 59 syllameles: 6 syllabics (one pitch), 19 neumes (one to four pitches), and 34 melismata (more than one pitches).

D.3. Music Staff

A G clef was superimposed on the PIC thanked to its *Ellis cent* measurement.

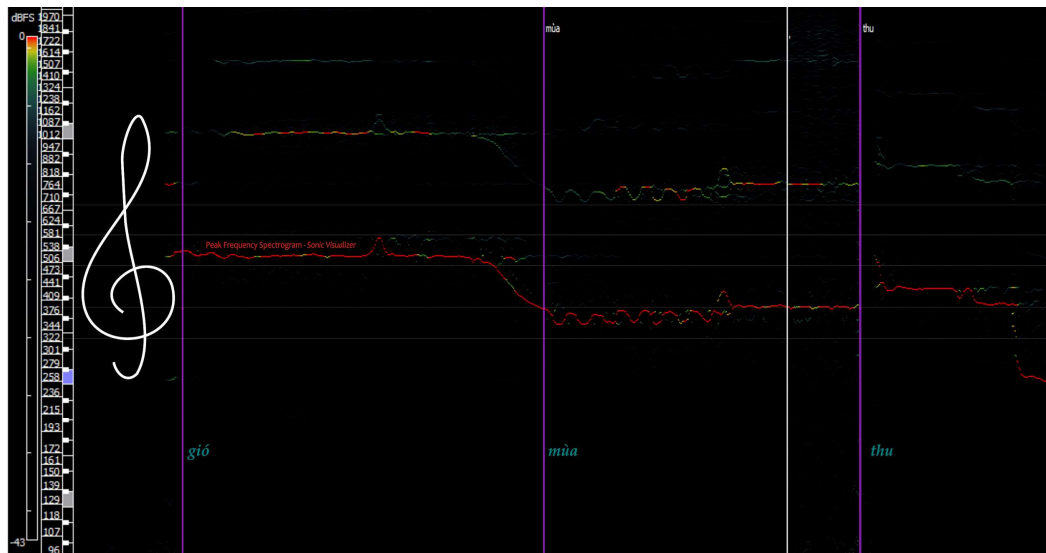


Figure 2. *Ru Con* nPIC with an overlaid G clef based on *cent* measurements.

D.4. Dialect tone readings of the *Ru Con* lyrics

Six recordings of the normal reading of the *Ru Con* lyrics were also made of males and females of the three major Vietnamese dialects: Hà Nội, Huế and Sài Gòn. These recordings done without any instructions to the readers show tones in regular speech. Their pairing is hoped to help bring out the effects of tones in singing.

Six nPICs were obtained from the six voice recordings of the *Ru Con* lyric (from now on we refer to them collectively as the “6 readings”). Aggregated syllamelis pairings of *đàn bầu*, the singing in southern female accents, and the six readings are presented online at http://vietcenter.cla.temple.edu/mpic/pw_rucon.php. (472 segmented nPIC syllameles).

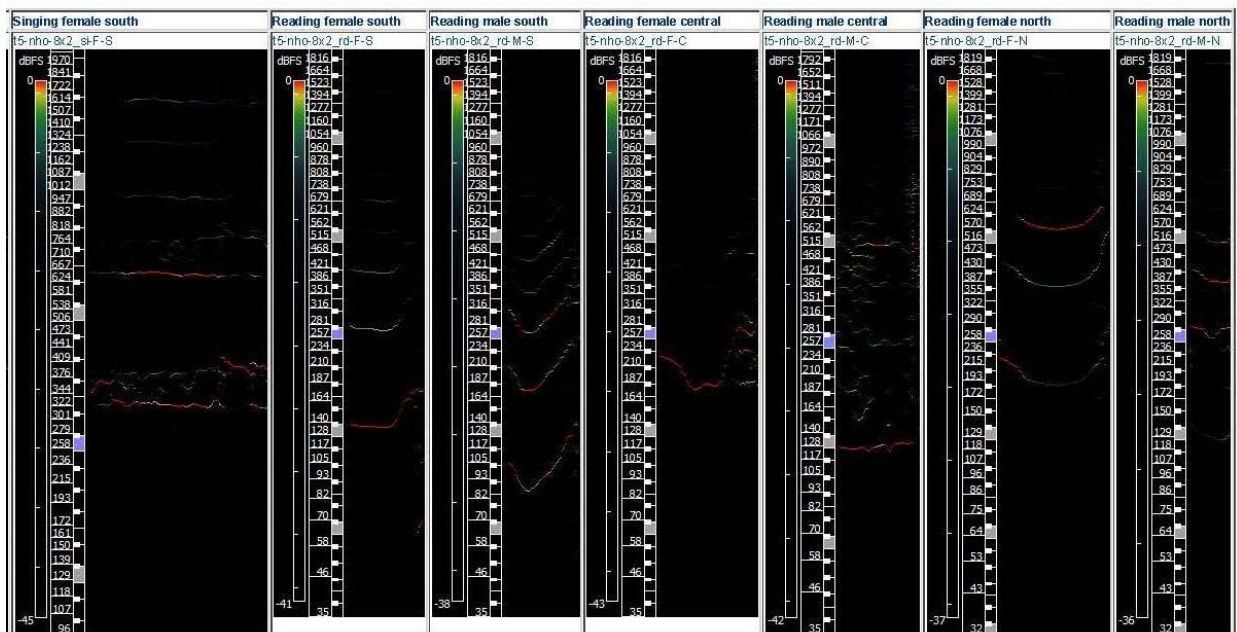
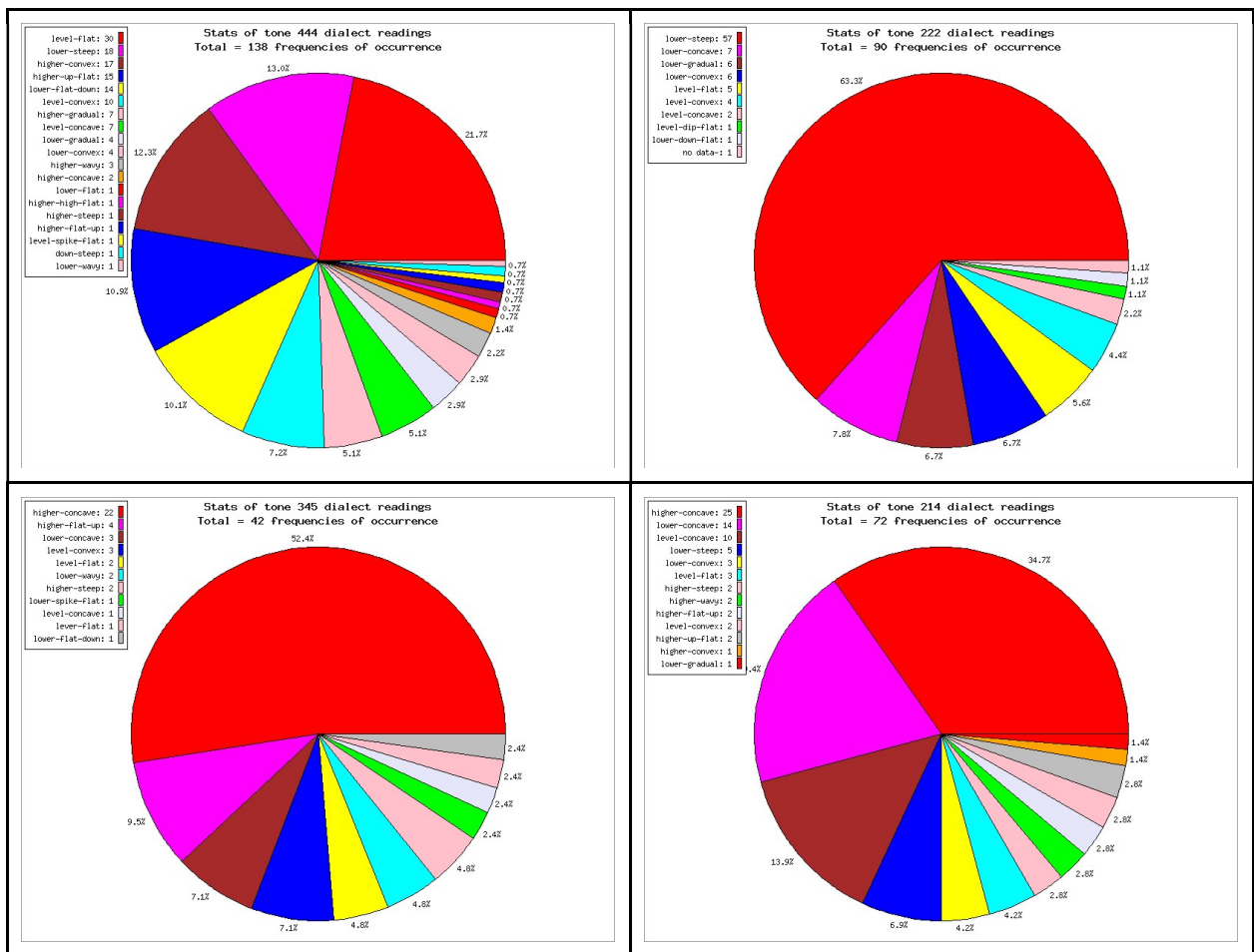


Figure 3: Pairings of 8 syllameles $\overline{\text{ng}}_{8,2}$ “*nho*³⁴⁵” for lyric syllable #8.2.

The patterns of pitches are shown in Figure 4. The followings are noted¹⁶:

1. Tone *ngang* /⁴⁴⁴/ “high level”, total 134 cases, shows 19 complex patterns.
2. Tone *huyền* /²²²/ “low level”, total 90 cases, the majority of the cases 81/90 shows lowering, only 12/90 readings show leveling.
3. Tone *sắc* /³⁴⁵/ “high rising”, total 42, the majority 28/42 shows rising.
4. Tone *hỏi/ngã* /²¹⁴/ “low rising”, total 72, the majority 42/72 shows concave pitch.
5. Tone *nặng* /²¹¹/ “creaky falling”, total 6, half shows steep falling.



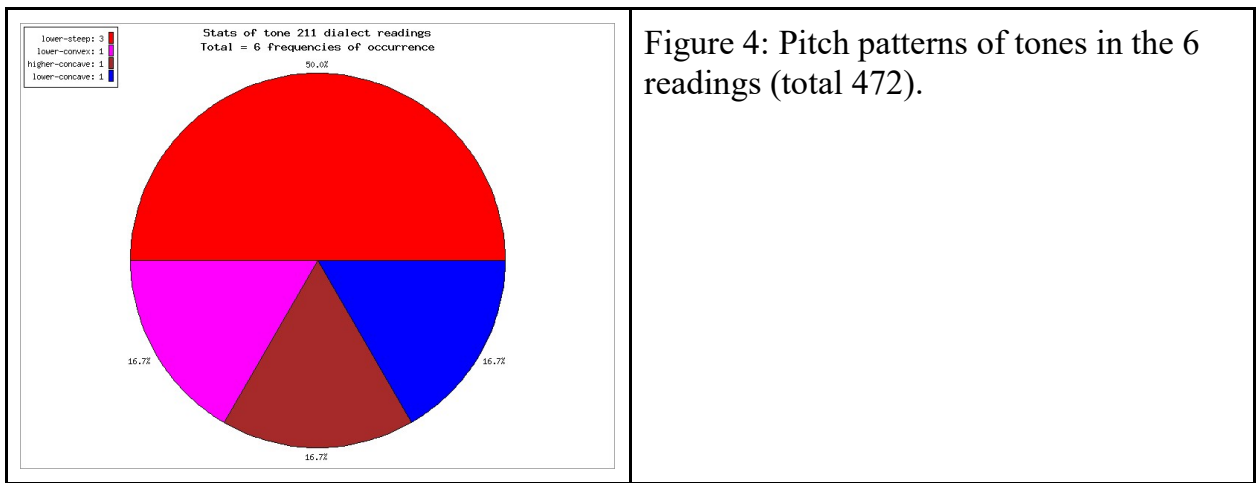


Figure 4: Pitch patterns of tones in the 6 readings (total 472).

Figure 4: Pitch patterns of tones in 6 readings.

D.5. Toneume: Matching phonemic and phonetic tone pitch in their syllamelis

The process of identifying the lyric tones in their corresponding syllameles requires several steps:

1. Lining up the phonemic and phonetic manifestations of tones onto the corresponding syllameles. For example, the tone *sác* “high rising” of word 1.1 *gió* “wind; breeze”, phonemic /yo³⁴⁵/ and phonetic [yo⁴⁴⁴ ɔ³⁴⁵] of the artist, in which the main tone is a spike, C⁵ to D⁵, surrounded by the long levelled high C⁵, helps to identify the tone [³⁴⁵] in Figures 5 and 6.

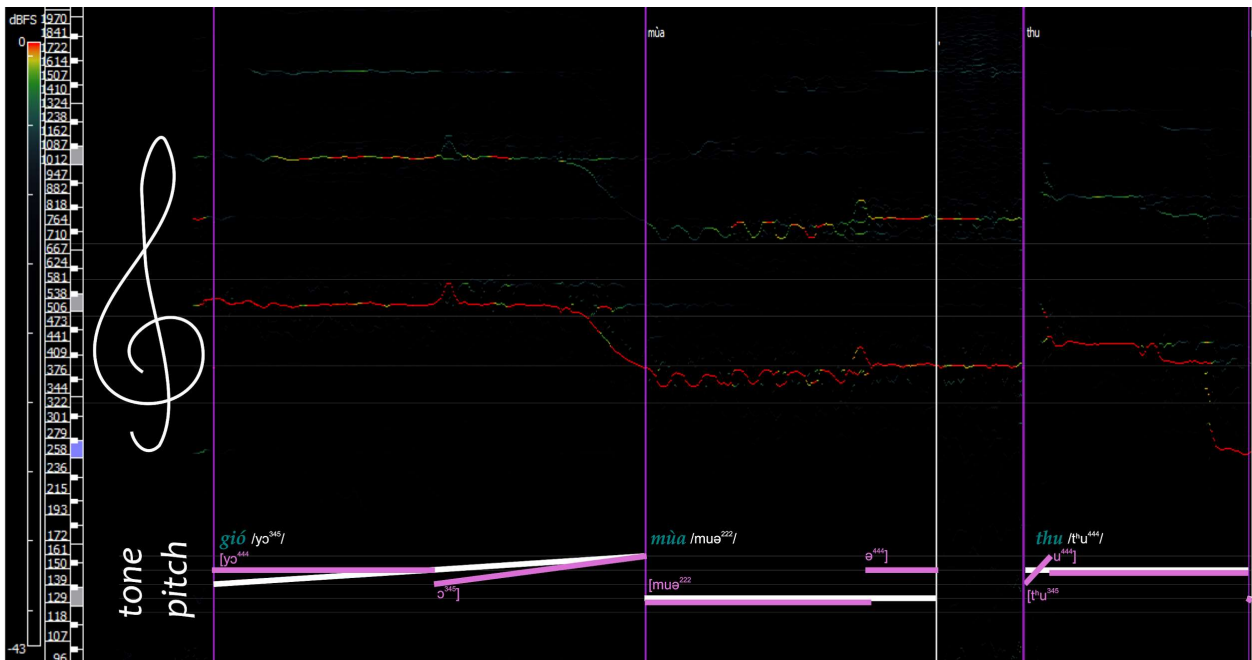


Figure 5: Linguistic tone pitch patterns pairing with their corresponding syllameles.

2. Finding the distinct patterns of tone pitches on the syllamelic pitch contours. This

is generally feasible with adjustments against *vibratos* (waving of pitch¹⁷), spikes (sudden sharp rise in pitch) and dips (sudden sharp drop in pitch).

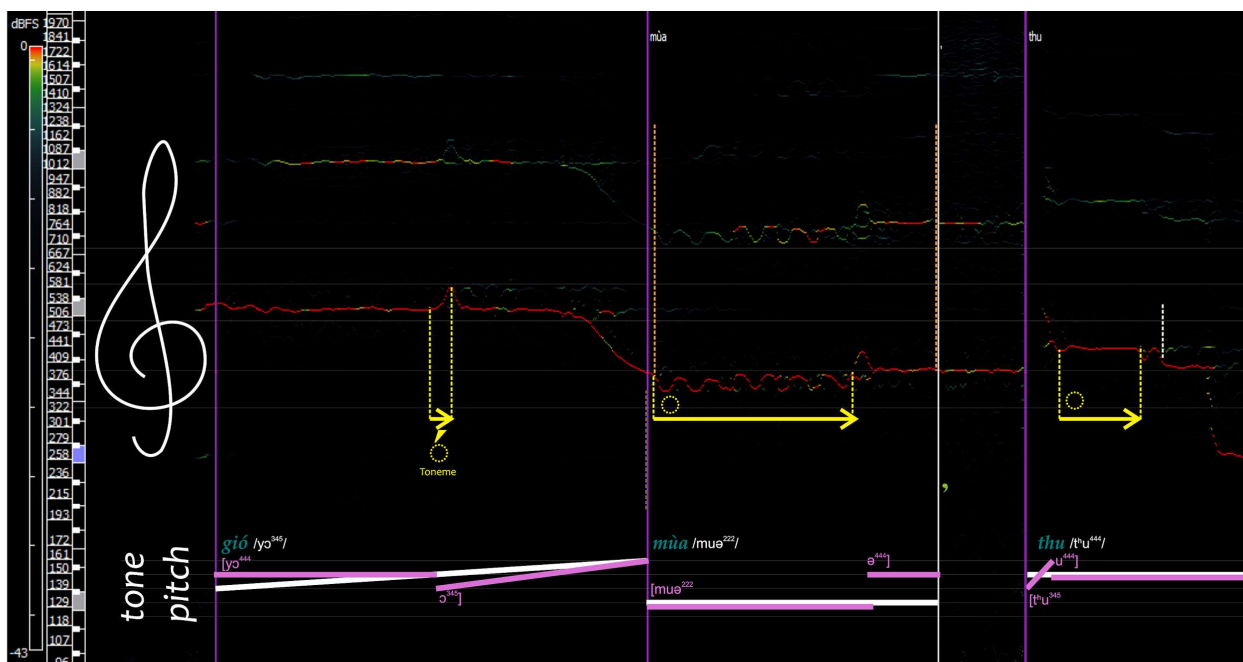


Figure 6: Toneumes (yellow arrows) in each syllable (between purple vertical bars). Ignoring the effects of 110 *vibratos*, 16 spikes and 4 dips, a toneume is identified in each syllable. They are fully demarcated on the nPIC of *Ru Con*, shown at http://vietcenter.cla.temple.edu/mpic/Ru_Con_nPIC_tones-in-syllables.jpg (henceforth, the full nPIC chart).

D.6. Congruence

Congruence, associated with adjacency, is a major concept which accounts for the interactions of Vietnamese linguistic tones (plus their inter-tones) and the folk song melody. It relies on the syllables and toneumes and adjacent inter-toneumes as the loci where congruence can be properly paired.

Rest (between phrases and syncopation) and breath catching are considered breaks in the continuity of the melody (marked by green commas [,] in the full nPIC chart), where congruence cannot be accounted for (for example, between syllables $\bar{\mathfrak{S}}_{2.2}$ and $\bar{\mathfrak{S}}_{2.3}$, or between toneumes $\bar{\mathfrak{T}}_{2.1}^{222}$ and $\bar{\mathfrak{T}}_{2.2}^{444}$ in Figure 6).

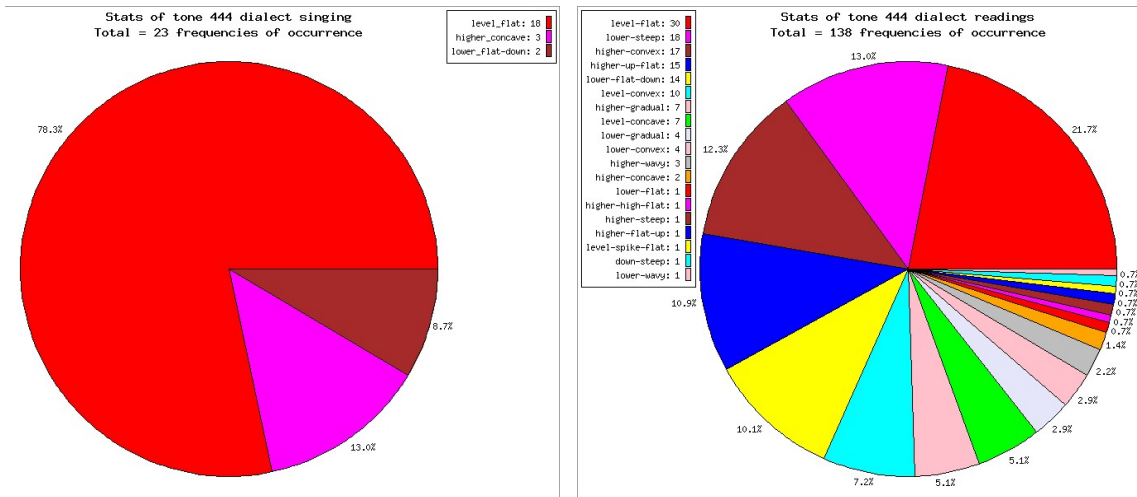
There are four types of congruences:

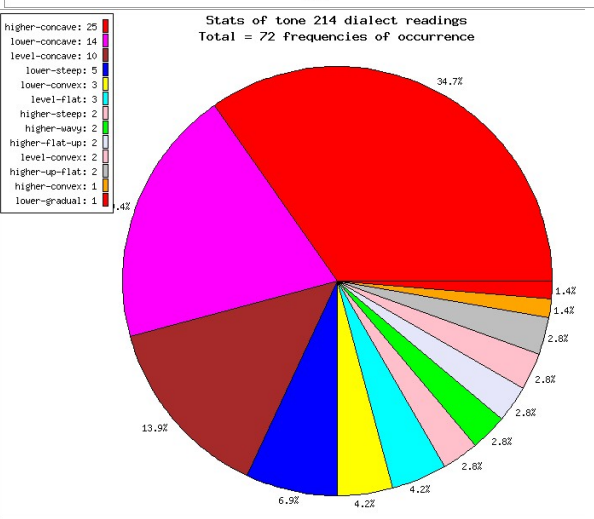
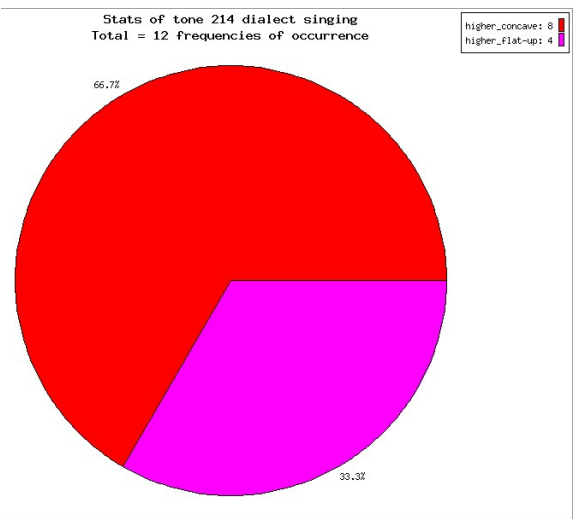
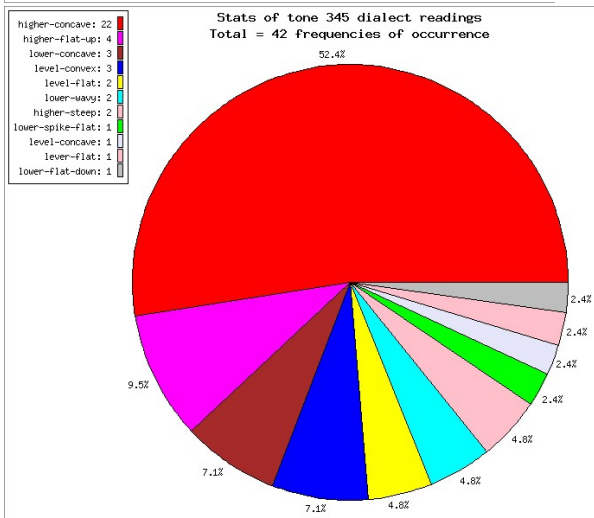
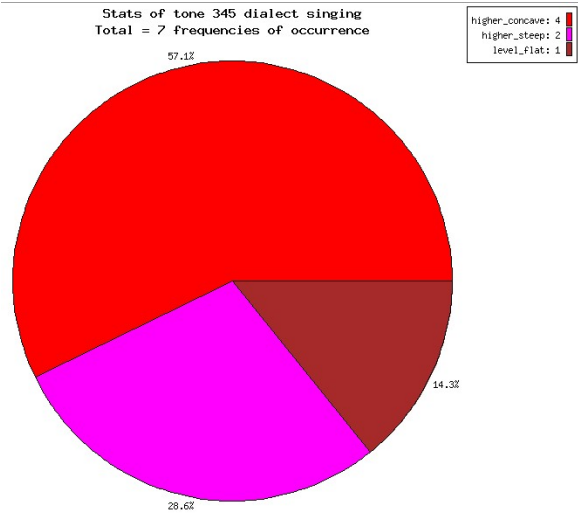
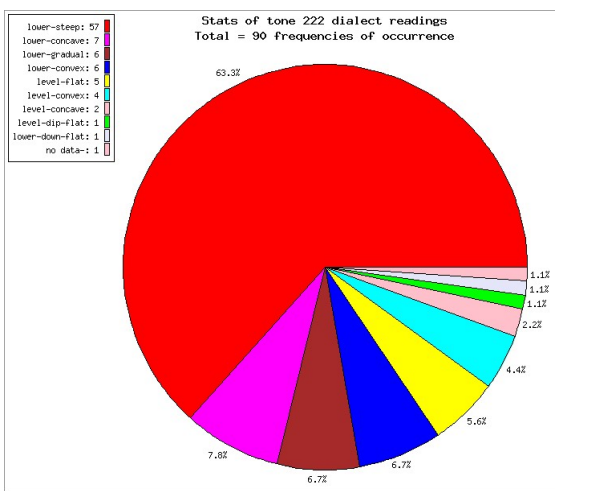
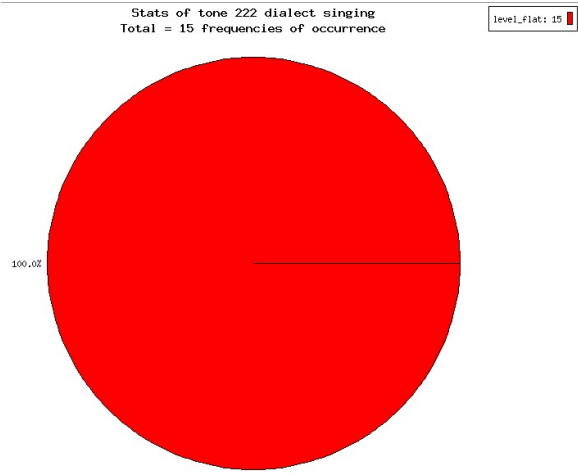
¹⁷ Oxford Music Online, Vibrato. Cf.

http://www.oxfordmusiconline.com.ezproxy.cul.columbia.edu/subscriber/article/opr/t114/e7129?q=vibrato&search=quick&pos=2&_start=1#firsthit

1. **Toneumes and their corresponding phonetic tone pitches.** The pairings of each toneume of the artist singing and of the 6 readings show the artist’s consistency, in Figure 7¹⁸.

- a. Tone *ngang* [444] “high level” is congruent with the dominant pattern of being flat at the high register.
- b. Tone *huyền* [222] “low level” is incongruent because the singing toneume stays flat at low register (tone pitch [222]), whereas the reading dialects show 63.3% steep falling, 7.8% concave falling, 6.7% gradual lowering, and 6.7% convex lowering. The lowering pattern is dominant (81/90).
- c. Tone *sắc* [345] “high rising” is congruent with the dominant pattern of concave rising. Both the singing and reading nPIC are different from the phonemic tone pitch of steep rising [345].
- d. Tone *hỏi/ngã* [214] “low rising” is congruent with the dominant pattern of concave fall from high ends (as suggested by the phonemic “low rising”).
- e. Tone *nặng* [211] “creaky falling” is incongruent as the singing toneume stays flat at a low register, while the reading dialects show 50% steep lowering in a low register (as suggested by its tone pitch [211]).





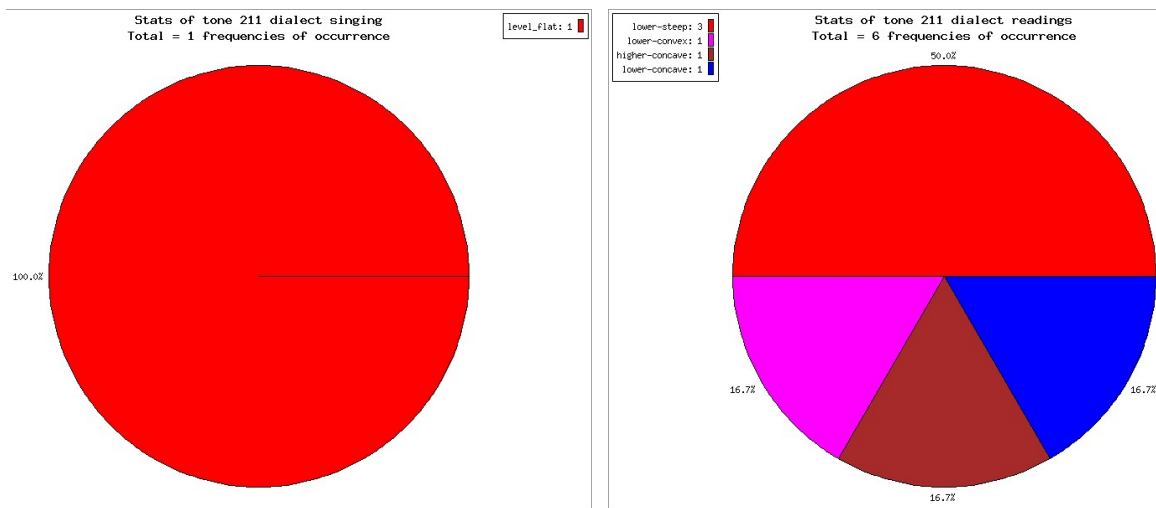


Figure 7: toneume vs tone pitch congruence

2. **Two adjacent inter-toneumes and their corresponding phonetic inter-tones**, marked by congruent \cong vs. incongruent $\not\cong$ on the full nPIC chart. The results show that the melody and the lyrics of *Ru Con* seem to flow well. The more congruent, the more natural the song sounds and is appreciated.

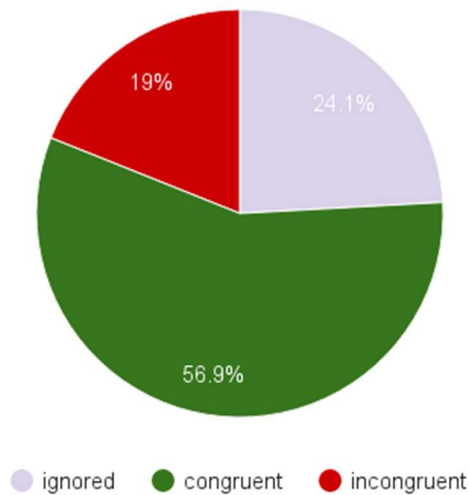


Figure 8: Inter-tone vs inter-toneume congruence.

Transitions in pitch between two adjacent toneumes are dominantly found in their proper syllameles (light green arrows on the full nPIC chart), showing the inter-toneumes in the direction of the next toneumes. The loci in the inter-toneumes are usually found after the first toneume and the onset of following toneumes (except for the spikes or dips due to the artist's style).

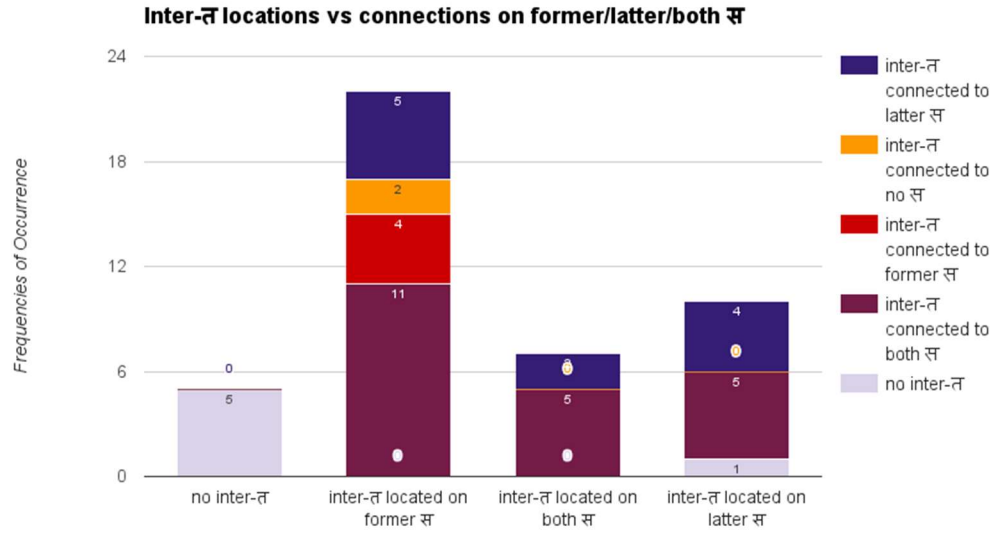


Figure 9: Detailed description of inter-toneume locations and connections

In *Ru Con*, 22 inter-toneumes are found on the first syllamelis, 10 on the second syllamelis, and 21 on both syllameles, ignoring 6 disrupted transitions due to rests and breath takings.

A toneume can move horizontally (time) and vertically (pitch) within its syllamelis without losing its distinct pitch characteristics. This behavior demonstrates the nature of linguistic tones. The tone pitches, in normal speech, vary from person to person and fluctuate due to different emotions. However, identical adjacent tone pitches, with large gaps between the end pitch of one tone and the onset pitch of the next, are potentially in conflict with the melody. For this reason, a formal *inter-toneume* is important in the formal description of the speech as well as the melody flows.

— In *Ru Con*, the two successive level toneumes, both *ngang* “high level” [444] and *huyền* “low level” [222], at times do not maintain the same pitch. For example, in phrase 2, *năm canh dài*, the three toneumes $\bar{t}_{2,1}^{444} \bar{t}_{2,2}^{444} \bar{t}_{2,3}^{222}$ show $\bar{t}_{2,1}^{444} \bar{t}_{2,2}^{333} \bar{t}_{2,3}^{222}$, where toneume $\bar{t}_{2,2}^{333}$ is lowered to accommodate the transition to toneume $\bar{t}_{2,3}^{222}$.

3. **Singers’ idiolectal toneumes** demonstrate changes in the contour of the tone pitches:

- toneumes *sắc* [345] “high rising” and *hỏi* [214] “low rising” have their first part flattened with light *vibratos*, followed by a sharp rise,
- toneumes *huyền* [222] “low level” are always falling rather than levelling, in congruence with the 6 normal readings.
- toneumes *hỏi* [214] “low rising” show a prolonged flat trough with light *vibratos*...

4. **Performed *rung* “vibrato”** does affect toneumes, and defining Vietnamese musical modes / airs presented in scholarly literature. There are 110 *vibratos* identified in *Ru con*. Their pitch data is obtained directly from *Sonic Visualizer*, showing highest, lowest and mean pitches of toneumes.¹⁹

When midpoints of *vibratos* of the toneumes are charted in ascending pitch order, 5 distinct groups become visible, sketching a Vietnamese modal/air system. See Figure 10 below, that matches scale *Ai-Oán* “mourning” mode of Figure 11.²⁰ This analysis independently confirms the conclusion in Ngô & Phan (8/2016), especially strongest regarding F#⁴ and G⁴.

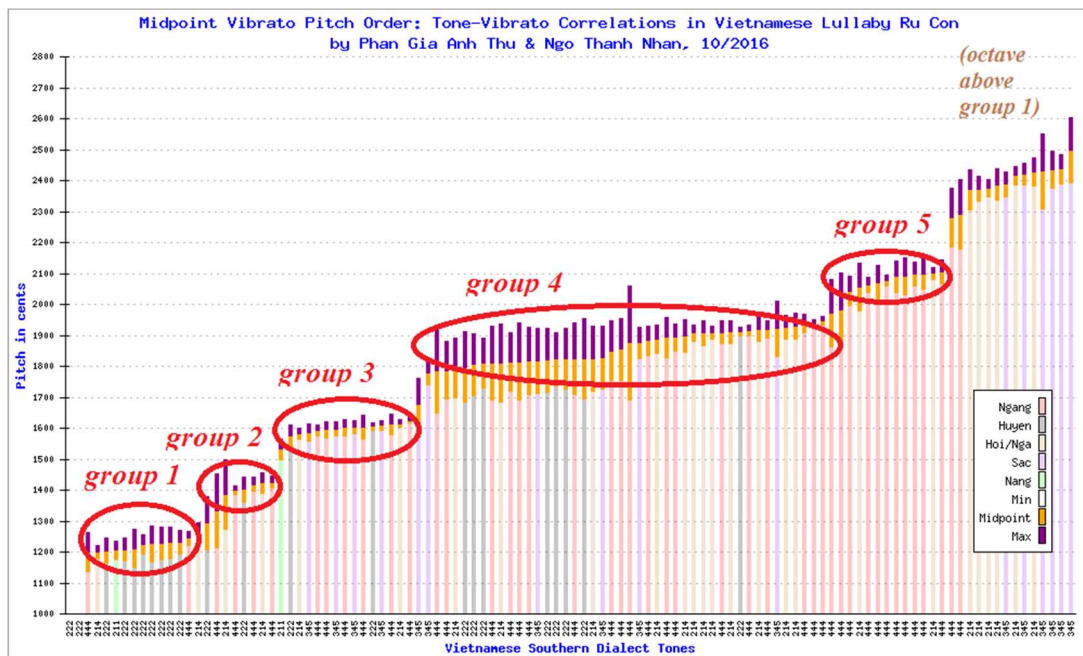


Figure 10. Grouping of *vibratos* in *Ru Con* according to their midpoints.

¹⁹ See full data at. <http://localhost/projects/mpic/vibs.php?resid=2>.

²⁰ Nguyễn, Phú Phong, op. Cit. p. 255.

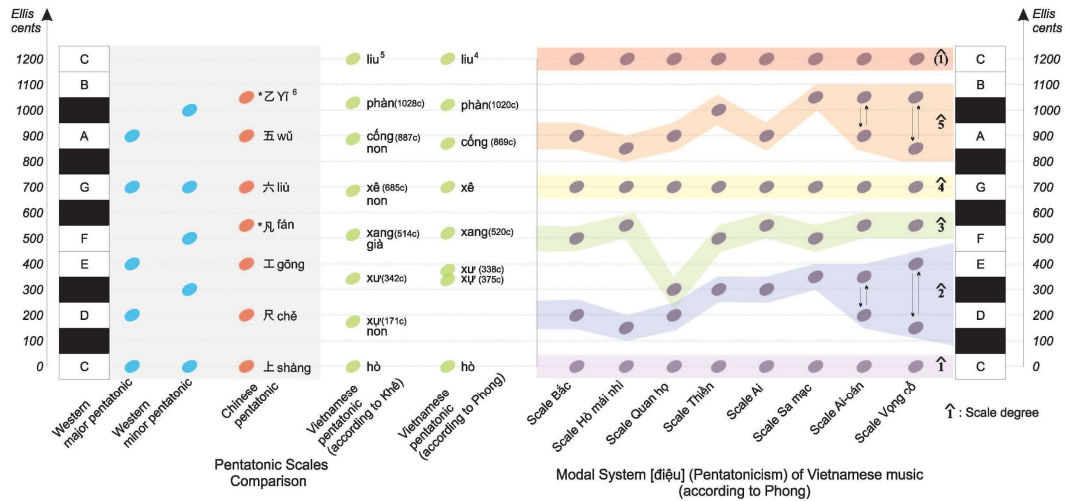


Figure 11: Pentatonic scales used in the Vietnamese artistic tradition.^{21,22,23,24}

Vibratos are said to be a systematic classificatory feature of Vietnamese traditional and folk music. For that reason, they must be measured and highlighted methodically, with data such as duration, onset point, number of circles, highest pitch, lowest pitch, amplitude, coda, etc.

One may surmise from the data on *vibratos* that there are no pentatonic scales which have *vibratos* on all scale degrees. However, a larger repertoire of folk lullabies is needed to reveal the full role of vibratos in Vietnamese modes and airs.

E. Conclusion

The MTVIET analysis of Vietnamese folk songs intends to devise a reliable process for teachers of Vietnamese traditional and folk music, while exposing the music *finesse*. The foundation of this process is a pitch/intensity vs time contour overlaid by onsets of syllameles and a music staff, called the *notated PIC graph* (nPIC), a visual representation of a performed version of a musical piece. It serves as a platform on which universal

²¹ Trần Văn Khê, in “Vài ý kiến về thất cung thiên nhiên và việc dùng comma để đo cung bậc trong nhạc Việt [Some opinions on the natural octave and the use of comma to measure steps in Vietnamese music],” *Nghiên Cứu Việt Nam*, Huế, No. 3 Fall 1966, p. 10.

²² Extracted from Nguyễn Phú Phong (2008), p. 253 with “[a]verage deviation: 20 cents.” He writes, “Finely adjusted intervals, even microtones, are typical of folk songs. Two to twelve tones may be selected from the twelve-tone system of tuning available in Vietnamese music.”

²³ Cf. https://en.wikipedia.org/wiki/Gongche_notation 工尺譜 Gongche [*công xê phổ*] was invented by the Tang Dynasty [*nhà Đường*, 唐朝 6/18/618 - 6/1/907] and became popular by the Song Dynasty [*nhà Tống*, 宋朝 960 - 1279]. 凡 *fán* is simply characterized as “between F and F#”, and 乙 *yǐ*, “between ti and ti.”

²⁴ Phong (2008), p. 255.

studies can be made. In this paper, *Ru con miên Nam*, a lullaby from southern Vietnam, was used to study the interactions of the song melody and linguistic tones. The nPIC of *Ru con miên Nam* consists of 59 syllables, broken into 13 phrases, and with an extensive presence of 110 *vibratos*. Each linguistic tone pitch is melodically represented by a toneume within its corresponding syllable, and an associated, often ignored, inter-toneume. Since Vietnamese tones are not spoken with *vibratos*, the extensive *vibratos* of both toneumes and inter-toneumes provide a clue to tracking Vietnamese modal/air systems as well as the idiolectal style of the artist.

The nPIC graph binds music researchers to raw data. It gives us measurable information about silence, pitch and rest duration, spikes, dips, *vibratos*, and the intensity of each pitch, etc. This study seems to call for both a fully automated nPIC generator, with a syllable onset detector using phonetic knowledge as well as a *vibrato* detector, and for a series of similar analyses of the same lullaby by different artists, as well as different songs of the same mode/air systems (*a priori* classified by grandmasters). We shall leave these subjects for future studies in Vietnamese folk music.

References

- Brossier, P., J.P. Bello, J.P. & Plumbley, M.D. (2004). "Real-time temporal segmentation of note objects in music signals," in *Proceedings of the International Computer Music Conference*, Miami, Florida, 2004, 458-461.
- Chao, Y.R. (1920). "A system of tone letters," *Le maitre phonétique*, 45: 24-27.
- Cannam, Chris and Queen Mary. (2015). *Sonic Visualiser*, release 2.5, © 2005–2015, University of London, from <http://www.sonicvisualiser.org/>.
- Benetos, Emmanouil & Dixon, Simon. 2012. A Shift-Invariant Latent Variable Model for Automatic Music Transcription, *Computer Music Journal* 36:4 (2012), 81-94.
- Simon Dixon. (2006). Onset Detection Revisited, in *Proceedings of the 9th International Conference on Digital Audio Effects (DAFx-06)*, Montreal, Canada.
- Han, M.S. & Kim, K.O. (1974). "Phonetic variation of Vietnamese tones in disyllabic utterances," *Journal of Phonetics*, 2.3: 223-232.
- Hoàng, Tuệ & Hoàng, Minh. (1976). Remarks on the phonological structure of Vietnamese," *Vietnamese Studies*, 40:65-95.
- Jähnichen, Gisa. (2014). [Melodic relativisation of speech tones in classical Vietnamese singing: the case of Many Voices](#), *Jahrbuch des Phonogrammarchivs der Österreichischen Akademie der Wissenschaften*, 4. Göttingen: Cuvillier, pp. 180–194.
- Miller, Terry & Williams, Sean. Eds. (2008). *The Garland Handbook of Southeast Asian Music*. New York: Routledge.
- Ngô, Thanh Nhân. 1984. *The syllabeme and patterns of word formation in Vietnamese*.

- Ph.D. thesis, New York University.
- Ngô, Thanh Nhân & Phan, Gia Anh Thu. 8/2016. [A contribution to teaching Vietnamese music: key notes in context and pitch contour map](#), in *Proceedings of ICESI 2016 International Conference on Education and Social Integration*, pp. 144-156, ISBN 978-0-9945391-06, Temple University Center for Vietnamese Philosophy, Culture & Society and Tôn Đức Thắng University, Hồ Chí Minh City, Vietnam.
- Nguyễn, Thuyết Phong. (1998). Vietnam. In Miller, T., & Williams, S. (Eds). *The Garland Encyclopedia of World Music: Southeast Asia*, 4:444-517. New York, NY: Garland Publishing, Inc.
- Nguyễn, Văn Lợi and Edmondson, J.A. (1998). "Tones and voice quality in modern northern Vietnamese: instrumental case studies", in *The Mon-Khmer Studies Journal*, vol. 28, pp. 1-18.
- Rodger, Matthew W.M. & Craig, Cathy M. (2013). "Moving with Beats and Loops: The Structure of Auditory Events and Sensorimotor Timing," *10th International Symposium on Computer Music Multidisciplinary Research (CMMR)*, Marseille, France.
- Trần, Văn Khê. (1962). *La musique vietnamienne traditionnelle*. Paris, Presses Universitaires de France.
- Trần, Văn Khê. (1967). *Les traditions musicales: Viêt-nam*. Institut International d'Études Comparatives de la Musique. Paris: Buchet-Chastel.



SHORT BIOGRAPHIES

Phan Gia Anh Thu, M.A. *Music Education*, Teachers College - Columbia University (2015), is a piano teacher at Teachers College, Columbia University (2014-present), and an Artist-in-residence at Mekong NYC, focusing on cultivating traditional identities for children in the Bronx, New York, through music, visualization, and traditional games (2016-present). She is also a *đàn bầu* [monochord] and *đàn kìm* [moon lute] musician of the *Mekong Traditional Vietnamese Instrumental Ensemble Troupe* (2013-present).
Contact by email at agp2132@tc.columbia.edu.

Ngô Thanh Nhân, Ph.D. *Linguistics*, a scholar of New York University *Linguistic String Project* (1984-present), a fellow and Adjunct Associate Director of Temple University *Center for Vietnamese Philosophy, Culture & Society* (Philadelphia, 2006-present), and a teacher of *Đàn tranh Ensemble* [16-string zither] at the *Folk Arts – Cultural Treasures Charter School*, Grades 3-8 (Philadelphia, 2009-present), and the *Mekong Traditional Vietnamese Instrumental Ensemble Troupe* (New York, 2013-present).

Contact by email at nhan@cs.nyu.edu or nhan@temple.edu.