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

Phan Gia Anh Thư
Columbia University Teachers College & Temple University, agp2132@tc.columbia.edu
&
Ngô Thanh Nhàn
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
readings of the same southern Vietnamese folk song, *Ru Con miền Nam* “Lullaby”. The results reveal that toneumes do manifest themselves distinctly in their syllameles, leaving space for inter-toneume expression and the artist’s idiosyncratic interpretation. For a tonal language, the lyric vs toneume/syllamelis 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 syllamelis 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, syllamelis, 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, vân.

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.

<table>
<thead>
<tr>
<th>• bồng “level”</th>
<th>• trắc “contour”</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>bồng</strong> [high]</td>
<td><strong>bồng</strong> [high]</td>
</tr>
<tr>
<td>Orthography: ○</td>
<td>Orthography: ○</td>
</tr>
<tr>
<td>Tone order: 1 [fo. 1333]</td>
<td>Tone order: 3 [fo. 830]</td>
</tr>
<tr>
<td><strong>Tone pitch:</strong> 444</td>
<td><strong>Tone pitch:</strong> /214/</td>
</tr>
<tr>
<td>Tone letter: 1</td>
<td>Tone letter: 1</td>
</tr>
<tr>
<td><strong>hội “low rising”</strong></td>
<td><strong>sắc “high rising”</strong></td>
</tr>
<tr>
<td>Orthography: ○</td>
<td>Orthography: ○</td>
</tr>
<tr>
<td>Tone order: 5 [fo. 1759]</td>
<td>Tone order: 5 [fo. 1759]</td>
</tr>
<tr>
<td><strong>Tone pitch:</strong> /345/</td>
<td><strong>Tone pitch:</strong> /345/</td>
</tr>
<tr>
<td>Tone letter: 1</td>
<td>Tone letter: 1</td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>através</th>
<th>huyền “low level”</th>
<th>ngã “creaky rising”</th>
<th>nẳng “creaky falling”</th>
</tr>
</thead>
<tbody>
<tr>
<td>[low]</td>
<td>Orthography: ◌̀</td>
<td>Orthography: ◌̃</td>
<td>Orthography: ◌̣</td>
</tr>
<tr>
<td></td>
<td>Tone order: 2 [fo. 1157]</td>
<td>Tone order: 4 [fo. 486]</td>
<td>Tone order: 6 [fo. 1414]</td>
</tr>
<tr>
<td></td>
<td>Tone pitch: /²²²/</td>
<td>Tone pitch: /⁴¹⁵/</td>
<td>Tone pitch: /²¹¹/</td>
</tr>
<tr>
<td></td>
<td>Tone letter: ↓</td>
<td>Tone letter: ˥</td>
<td>Tone letter: ࠉ</td>
</tr>
</tbody>
</table>

Table 1. Vietnamese tones and their representations.

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

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.

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. 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 năng /211/ in côi “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. 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, plural, syllameles, Sanskrit unit symbol स/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 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, Sankrit unit symbol त/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.

---

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.
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”\(^\text{13}\) 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 dao “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,\(^\text{14,15}\) nevertheless, the folk version is still a masterpiece and is the one most favored by the community.

```
“Gió mùa thu... mẹ ru mà con
ngủ...
Nam... canh dài...
Nam... canh dài, thắc dù vida năm...
Hỡi chàng... chàng ơi!
Hỡi người... người ơi!
Em nhớ tôi chàng.
Em nhớ tôi chàng!
Hãy nín... nín đi, con!
Hãy ngủ... ngủ đi, con!
Con hời mà con hối!
Con hối, con hối...
Con hối, con hối, hối con!”
```

Tentative translation:

1. Autumn breeze helps Mother to lull her baby to sleep.
2. Five times the timekeeper had struck for ten hours straight...
3. I have been up all ten.
4. Oh, lover… please, lover!
5. Oh, man… please, man!
6. I am thinking of you,
7. I really miss you!
8. Hush… don’t cry, baby!
9. Go to sleep… sleep well, baby!
10. Oh child, please hush, oh baby!
11. Oh child, please hush, oh baby!
12. 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 syllable units in the nPIC. There are 23 tone /ngang/ “high level”, 15 tone /huyễn/ “low level”, 12 tone /hỏi/ngã/ “concave tone”, 8 tone /sắc/ “high rising”, and 1 tone /nặng/ “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/: 

<table>
<thead>
<tr>
<th>1.1</th>
<th>1.2</th>
<th>1.3</th>
<th>1.4</th>
<th>1.5</th>
<th>1.6</th>
<th>1.7</th>
<th>1.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>gió /yó⁴⁴⁵/</td>
<td>mưa /mù⁴⁴⁵/</td>
<td>thủ /tu⁴⁴⁴/</td>
<td>mẹ /me²¹¹/</td>
<td>ru /ru⁴⁴⁴/</td>
<td>mà /ma²²²/</td>
<td>con /kɤn⁴⁴⁴/</td>
<td>ngú /nú²¹⁴/</td>
</tr>
<tr>
<td>[yó⁴⁴⁵ ə³⁴⁵]</td>
<td>[mù⁴⁴⁵ ə³⁴⁴]</td>
<td>[tu⁴⁴⁴ ə³⁴⁴]</td>
<td>[me²¹¹ ə³⁴⁴]</td>
<td>[ru⁴⁴⁴ ə³⁴⁴]</td>
<td>[ma²²² ə³⁴⁴]</td>
<td>[kɤn⁴⁴⁴ ə³⁴⁴]</td>
<td>[nú²¹⁴ ə³⁴⁵]</td>
</tr>
<tr>
<td>2.1</td>
<td>2.2</td>
<td>3.3</td>
<td>3.3</td>
<td>4.4</td>
<td>5.4</td>
<td>5.4</td>
<td>5.4</td>
</tr>
<tr>
<td>năm /næm⁴⁴⁴/</td>
<td>canh /kɤŋ⁴⁴⁴/</td>
<td>dài /dày²²²/</td>
<td>dàn /di²¹⁴/</td>
<td>vú /vúy²²²/</td>
<td>năm /næm⁴⁴⁴/</td>
<td>năm /næm⁴⁴⁴/</td>
<td>năm /næm⁴⁴⁴/</td>
</tr>
<tr>
<td>[næm⁴⁴⁴ ə²²² ə³⁴⁵]</td>
<td>[kɤŋ⁴⁴⁴ ə³⁴⁴]</td>
<td>[dày²²² ə³⁴⁴]</td>
<td>[dí²¹⁴ ə³⁴⁵]</td>
<td>[vúy²²² ə³⁴⁴]</td>
<td>[næm⁴⁴⁴ ə³⁴⁴]</td>
<td>[næm⁴⁴⁴ ə³⁴⁴]</td>
<td>[næm⁴⁴⁴ ə³⁴⁴]</td>
</tr>
<tr>
<td>3.3</td>
<td>4.3</td>
<td>4.3</td>
<td>5.3</td>
<td>6.3</td>
<td>6.3</td>
<td>7.4</td>
<td>7.4</td>
</tr>
<tr>
<td>năm /næm⁴⁴⁴/</td>
<td>canh /kɤŋ⁴⁴⁴/</td>
<td>dài /dày²²²/</td>
<td>chăng /ʧaŋ⁴⁴⁴/</td>
<td>người /n̥uə⁴⁴⁴/</td>
<td>người /n̥uə⁴⁴⁴/</td>
<td>cháng /ʧaŋ⁴⁴⁴/</td>
<td>cháng /ʧaŋ⁴⁴⁴/</td>
</tr>
<tr>
<td>[næm⁴⁴⁴ ə²²² ə³⁴⁵]</td>
<td>[kɤŋ⁴⁴⁴ ə³⁴⁴]</td>
<td>[dày²²² ə³⁴⁴]</td>
<td>[ʧaŋ⁴⁴⁴ ə³⁴⁴]</td>
<td>[n̥uə⁴⁴⁴ ə²²²]</td>
<td>[n̥uə⁴⁴⁴ ə²²²]</td>
<td>[ʧaŋ⁴⁴⁴ ə³⁴⁴]</td>
<td>[ʧaŋ⁴⁴⁴ ə³⁴⁴]</td>
</tr>
<tr>
<td>5.4</td>
<td>6.4</td>
<td>7.4</td>
<td>8.4</td>
<td>9.4</td>
<td>9.4</td>
<td>9.5</td>
<td>9.5</td>
</tr>
<tr>
<td>hợi /hɤy²¹⁴/</td>
<td>người /n̥uə⁴⁴⁴/</td>
<td>cháng /ʧaŋ⁴⁴⁴/</td>
<td>cháng /ʧaŋ⁴⁴⁴/</td>
<td>con /kɤn⁴⁴⁴/</td>
<td>con /kɤn⁴⁴⁴/</td>
<td>con /kɤn⁴⁴⁴/</td>
<td>con /kɤn⁴⁴⁴/</td>
</tr>
<tr>
<td>[hɤy²¹⁴ ə³⁴⁵]</td>
<td>[n̥uə⁴⁴⁴ ə²²²]</td>
<td>[ʧaŋ⁴⁴⁴ ə³⁴⁴]</td>
<td>[ʧaŋ⁴⁴⁴ ə³⁴⁴]</td>
<td>[kɤn⁴⁴⁴ ə³⁴⁴]</td>
<td>[kɤn⁴⁴⁴ ə³⁴⁴]</td>
<td>[kɤn⁴⁴⁴ ə³⁴⁴]</td>
<td>[kɤn⁴⁴⁴ ə³⁴⁴]</td>
</tr>
</tbody>
</table>

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/:
<table>
<thead>
<tr>
<th>11.1</th>
<th>11.2</th>
<th>11.3</th>
<th>11.4</th>
<th>11.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>con</td>
<td>/kɔn/</td>
<td>/vɔn/</td>
<td>con</td>
<td>coni444/</td>
</tr>
<tr>
<td>[kɔn/444 ᵃ⁴⁴]</td>
<td>[kɔn/244 ᵃ⁴⁴]</td>
<td>[kɔn/244 ᵃ⁴⁴]</td>
<td>/vɔn/</td>
<td>/vɔn/</td>
</tr>
<tr>
<td>12.1</td>
<td>12.2</td>
<td>12.3</td>
<td>12.4</td>
<td></td>
</tr>
<tr>
<td>con</td>
<td>/kɔn/</td>
<td>/kɔn/</td>
<td>con</td>
<td>coni444/</td>
</tr>
<tr>
<td>[kɔn/444]</td>
<td>[kɔn/444 ᵃ⁴⁴]</td>
<td>/kɔn/</td>
<td>[kɔn/244 ᵃ⁴⁴]</td>
<td></td>
</tr>
<tr>
<td>13.1</td>
<td>13.2</td>
<td>13.3</td>
<td>13.4</td>
<td>13.5</td>
</tr>
<tr>
<td>con</td>
<td>/kɔn/</td>
<td>con</td>
<td>con</td>
<td>con</td>
</tr>
<tr>
<td>[kɔn/444]</td>
<td>[kɔn/444 ᵃ⁴⁴]</td>
<td>[kɔn/444 ᵃ⁴⁴]</td>
<td>[kɔn/444 ᵃ⁴⁴]</td>
<td>/vɔn/</td>
</tr>
<tr>
<td>[kɔn/444 ᵃ⁴⁴]</td>
<td>[kɔn/444 ᵃ⁴⁴]</td>
<td>/vɔn/</td>
<td>[kɔn/444 ᵃ⁴⁴]</td>
<td>[kɔn/444 ᵃ⁴⁴]</td>
</tr>
</tbody>
</table>

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♯3−18c (18c under F♯3) to F♯3+2c (2c over F♯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♯3−18c — F♯3+2c — Bin Pitch: F♯3−23c — F♯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.
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 Sàigòn. 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 syllables स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्स्ः for lyric syllable #8.2. The patterns of pitches are shown in Figure 4. The followings are noted:

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 sặc /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.

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

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

1. Lining up the phonemic and phonetic manifestations of tones onto the corresponding syllableis. For example, the tone sáng “high rising” of word 1.1 gió “wind; breeze”, phonemic /yɔ̂345/ and phonetic [yɔ̂444 ɔ̂345] 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 [ɔ̂345] in Figures 5 and 6.

2. Finding the distinct patterns of tone pitches on the syllableic pitch contours. This
is generally feasible with adjustments against *vibratos* (wavering of pitch\(^{17}\)), spikes (sudden sharp rise in pitch) and dips (sudden sharp drop in pitch).

![Figure 6: Toneumes (yellow arrows) in each syllamelis (between purple vertical bars).](http://vietcenter.cla.temple.edu/mpic/Ru_Con_nPIC_tones-in-syllameles.jpg)

Ignoring the effects of 110 *vibratos*, 16 spikes and 4 dips, a toneume is identified in each syllamelis. They are fully demarcated on the nPIC of *Ru Con*, shown at [http://vietcenter.cla.temple.edu/mpic/Ru_Con_nPIC_tones-in-syllameles.jpg](http://vietcenter.cla.temple.edu/mpic/Ru_Con_nPIC_tones-in-syllameles.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 syllameles 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 syllameles स\(_{2.2}\) and स\(_{2.3}\), or between toneumes त\(_{2.1}\) and त\(_{2.2}\) in Figure 6).

There are four types of congruences:

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 7\(^{18}\).

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 \([211]\)).

---

18 Specific data is shown at http://vietcenter.cla.temple.edu/mpic/rucontoneumes.php?resid=444.
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. 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 syllable, 10 on the second syllable, and 21 on both syllables, ignoring 6 disrupted transitions due to rests and breath takings.

A toneume can move horizontally (time) and vertically (pitch) within its syllable 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 \( t_{2.1}^{444} t_{2.2}^{444} t_{2.3}^{222} \) show \( t_{2.1}^{444} t_{2.2}^{333} t_{2.3}^{222} \), where toneume \( t_{2.2}^{333} \) is lowered to accommodate the transition to toneume \( t_{2.3}^{222} \).

3. **Singers’ idiolectal toneumes** demonstrate changes in the contour of the tone pitches:
— toneumes szęc \[345\] “high rising” and höi \[214\] “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 höi \[214\] “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-Oân* “mourning” mode of Figure 11.\(^{20}\) This analysis independently confirms the conclusion in Ngô & Phan (8/2016), especially strongest regarding F\#\(^4\) and G\(^4\).

---


\(^{20}\) Nguyên, Phú Phong, op. Cit. p. 255.
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 bực 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.”
studies can be made. In this paper, Ru con miền Nam, a lullaby from southern Vietnam, was used to study the interactions of the song melody and linguistic tones. The nPIC of Ru con miền 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 toneeme within its corresponding syllable, and an associated, often ignored, inter-toneeme. Since Vietnamese tones are not spoken with vibratos, the extensive vibratos of both toneemes and inter-toneemes 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


SHORT BIOGRAPHIES

Phan Gia Anh Thư, 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.


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

Thử phân tích mối tương tác giữa thanh điệu tiếng Việt và nhạc dân ca Việt Nam

25 Hai tác giả cám ơn nhóm Mekong Traditional Vietnamese Instrumental Ensemble Troupe (MTVIET) nhờ các sáng kiến xây ra trong buổi họp ngày 4 tháng 3 năm 2016.
Tóm tắt

Bài nghiên cứu này mô tả tương tác giữa 6 thanh trong lời và giai điệu dân ca Việt Nam. Dựa trên nền hệ thống phổ quát: chuỗi âm thanh của ngôn ngữ và giai điệu quyện vào nhau trong một giọng hát dân ca, các hệ lý thuyết nhạc và ngôn ngữ học phương Tây và Việt Nam hỗ trợ thể cùng linh hoạt. Điều tra khoa học bắt đầu bằng một biểu đồ biến thiên cao/cường độ có minh họa (notated pitch/intensity contour, nPIC) của một bản dân ca thu thanh, xuất phát từ đỉnh âm tần, trên đó, bất tiếng (hay từ) được định vị, bản tự động. Đơn vị này gọi là xướng âm tiết (syllamelis, syllameles). Cùng trên nền nPIC, một đồng nhạc được đặt lên. Để định được đấy dựa vào một số tương tác giữa thanh và giai điệu, chúng tôi đề cử các khái niệm: đơn vị xướng âm tiết, đơn vị xướng thanh, tính kế liền, và tính tương hợp (nghĩa là, một thanh có được thể hiện toàn vẹn trong giai điệu không). Một đơn vị xướng thanh (tone+neume → toneume) nằm gọn trong đơn vị xướng âm tiết bằng các nét đặc trưng ngôn ngữ học của nó. Một đơn vị âm học liên xướng thanh (inter-toneume) xuất hiện để là đơn vị cầu nối giữa hai đơn vị xướng thanh kề liền nhau. Một hệ tương tác phức tạp đã được tìm ra: (1) giữa thanh điệu các miền và đơn vị xướng thanh, (2) giữa hai đơn vị xướng thanh và hiện tượng âm học chuyển thanh giữa hai tiếng, (3) các thanh và đơn vị xướng thanh của giọng địa phương và đặc thù của nghệ sĩ, và (4) giữa giọng rung của nghệ sĩ và thanh điệu cũng như hệ âm giai, điệu và hơi do sách vở dân nhạc mô tả. Phân tích sơ khởi trên dựa vào 412 đơn vị xướng âm tiết nPIC của 7 tệp thu thanh của giọng nam nữ 3 miền và giọng hát của nghệ sĩ bài Ru Con miền Nam. Kết quả cho thấy các đơn vị xướng thanh thường có đủ nét đặc trưng trong mỗi đơn vị xướng âm tiết, và những chỗ còn lại cho đơn vị liên xướng thanh và phương ngữ của nghệ sĩ. Một số ngôn ngữ có thanh điệu, tương hợp thuận giữa lời và các đơn vị
xướng thanh/xướng âm tiết thường được ưa chuộng, trong khi tương hợp nghịch là nơi nghệ sĩ thể hiện cá tính của mình. Phương pháp đề nghị trong nghiên cứu này: nền nPIC với định vị xướng âm tiết, và vai trò của thanh và rung hệ thống trong giai điệu, giúp cho các nhà giáo dụng tự do các ca khúc trên mạng và soạn bài giảng chính xác và nhanh chóng cũng như bài tập đầy thách thức cho người học. Phương pháp này cũng giúp người học ở bất cứ đâu hiểu ngay cấu trúc của ngôn ngữ và âm nhạc.

Key words: biểu đồ biến thiên cao/cường độ nPIC, cao độ trong mạch giai điệu kpic, thanh, tương đồng, liên thanh điệu, đơn vị xướng âm tiết, đơn vị xướng thanh, rung.

Paper’s ID: VS2.032P1