TONE CORRESPONDENCES AND TONOGENESIS IN VIETIC

Koichi Honda
School of Language Studies, ANU
Koichi.Honda@anu.edu.au

This paper\(^1\) aims to examine tone correspondences in Vietic languages, namely Vietnamese, Ruc and Arem, and attest Haudricourt’s tonogenesis hypothesis.

1. Haudricourt’s Tonogenesis hypothesis of Vietnamese

Haudricourt’s Tonogenesis hypothesis of Vietnamese (1954) is well known. It not only put an end to the controversy on what language family Vietnamese belongs to, but also proposed a model of the development of tones in a language. Even after fifty years since its publication, it is generally accepted and still very influential among linguists. The hypothesis is usually explained as follows.

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
\text{pa} = \text{pa} = \text{pa} = \text{ba} & \text{paX} > \text{pá} = \text{pá} > \text{bá} & \text{pah} > \text{pà} > \text{pà} > \text{bà} \\
\text{ba} = \text{ba} > \text{pà} > \text{bà} & \text{baX} > \text{bá} > \text{pà} > \text{bà} & \text{bah} > \text{bà} > \text{pà} > \text{bà} \\
\end{array}
\]

In stage 1, Vietnamese started as a toneless language. In stage 2, three pitch contours were formed depending on the coda types: level pitch from null coda, rising pitch from stop coda including glottal stop, and falling pitch from voiceless fricative coda. In stage 3, they split into higher and lower series according to [+/- voice] feature of onsets, and the number of tones doubled causing onsets of lower series devoiced. In stage 4, voiceless onsets were voiced again, though this stage is not relevant to the number of tones. This model may be called a **consonantly-based model** of tonogenesis.

However, simple questions arise. Tone is a phonological category, and it is realized and perceived by phonetic pitch. This pitch is encoded acoustically by fundamental frequency (F0), and this F0 cannot exist without voicing. If I accept the consonantly-based model, such syllables as [ai] or [au] without consonants cannot have tones. And specifically, how can a change from [bà] to [pa] between stage 2 and 3 explained? This change from voiced high rising pitch to voiceless low pitch does not seem to be plausible.

My question leads to a certain confidence that it is not consonants themselves, but interactions between various aspects of phonation that determines the pitch. This kind of idea is proposed by Thurgood (2001) in the name **laryngeally-based model** of tonogenesis.

---

\(^1\) This paper is the main result from my Master of Linguistics sub-thesis at the ANU (2004). I’m indebted to my supervisor Dr. Paul Sidwell for his enormous efforts, and lots of comments on it by Dr. Mark Alves and Dr. Phil Rose. I’m also indebted to Dr. Graham Thurgood for a new concept specifically applied in this paper.

2. Ruc and Arem – Vietic languages
I’m going to focus on the languages of Arem, Ruc and Vietnamese. The term “Vietic” is used to be known as “Viet-Muong”. Some scholars such as Ferlus prefer to use the latter name.

Population of Ruc and Arem is extremely small. According to the census in 1985, it is Arem 76, Ruc 125, Mälêng 715, Mây 715, Sách 625 (Phong et al. 1988). Another source shows Chìt 2400 (VNA 1996). The name Chìt is used by the government to represent these extremely minor ethnic people settled in the North-Central area of Vietnam.

The survey of Ruc language started in 1986 by Russian and Vietnamese linguists (Loi 1993). However, the report is not available at hand. Available materials at hand are the following three.

- Nguyễn Phú Phong, Trần Trí Đổi and M. Ferlus (1988) on Ruc
- Nguyễn Văn Lợi (1993) on Ruc
- Michel Ferlus (1997, original version 1991) on Arem and Ruc

My study is based on these sources and not based on the field work of my own.

In the SEALS Conference, Mark Alves and Michel Ferlus have contributed papers on Ruc language in 1997, 1999 and 2001. These papers are also cited.

3. Lexicostatistic Data
The following lexicostatistic data show lexical distance between the languages of Vietic and Katuic.

(3) Compared Vietic and Katuic lexica (Samarina, I.V. 1989, cited in Loi 1993)

<table>
<thead>
<tr>
<th></th>
<th>Arem</th>
<th>Pakoh</th>
<th>Taoih</th>
<th>Bru</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ruc</td>
<td>66.5</td>
<td>71.0</td>
<td>72.0</td>
<td>63.0</td>
</tr>
<tr>
<td>Poong</td>
<td>53.0</td>
<td>51.5</td>
<td>35.5</td>
<td>33.0</td>
</tr>
<tr>
<td>Muong</td>
<td>50.5</td>
<td>50.0</td>
<td>27.0</td>
<td>26.5</td>
</tr>
<tr>
<td>Viet</td>
<td>74.0</td>
<td>51.5</td>
<td>45.0</td>
<td>24.0</td>
</tr>
</tbody>
</table>

As is shown, Vietnamese shares nearly 50% common lexica with Ruc and Arem, while the rates are less than 30% with Katuic languages.

4. Tone Description
Tone description is crucially important but is a bit complicated because they are not shown by pitch alone but voice quality differences are involved in Vietic languages.
Tone correspondences and tonogenesis in Vietic

4-1. Standard Northern Vietnamese: 6 tones
Standard Northern Vietnamese is well known to have six tones, in Vietnamese name ngang, sắc, hoi, huyền, năng, ngã.

<table>
<thead>
<tr>
<th></th>
<th>(plain)</th>
<th>(abrupt)</th>
<th>(contour)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ngang (1A)</td>
<td>2. sắc (1B)</td>
<td>3. hoi (1C)</td>
</tr>
<tr>
<td>4</td>
<td>huyền (2A)</td>
<td>5. năng (2B)</td>
<td>6. ngã (2C)</td>
</tr>
</tbody>
</table>

They are grouped into historical higher and lower series, and each two tones make three pairs: plain tones, abrupt tones and contour tones. Besides differences in pitch, năng tone is accompanied by glottal stop, and ngã tone has creaky voice. Tone names in brackets, which I call **comparative tone numbers**, will be used to show tone correspondences more explicitly; the numbers (1 and 2) indicate higher and lower series, and the alphabets (A, B and C) indicate three contour types.

4.2 Ruc: 4 tones
Ruc is reported to have four tones.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>High level pitch</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Low level pitch (with breathy voice) or falling pitch (with breathy voice)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Rising pitch (with optional glottal stop)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Concave pitch or falling pitch (with optional glottal stop)</td>
<td></td>
</tr>
</tbody>
</table>

According to Loi (1993), tone 1 has high level pitch, and tone 2 has two variations in free variation, accompanied by breathy phonation. Tone 3, rising pitch can accompany glottal stop optionally. Tone 4 is reported to have two free variations, concave pitch and low falling pitch, and the falling pitch is accompanied by glottal stop optionally. Phong et al. (1988) describes six tones for Ruc, which counts tones 3 and 4 with and without glottal stop as an individual one.

4.3 Arem: Toneless
Arem is reported to be toneless, but to have breathy phonation, and post-vocalic laryngeal constriction (Ferus 1997, 2001).

4.4 Research Questions
Here are some questions. According to Haudricourt, the tones developed from $0 \geq 3 > 6$. How are they different from Ruc’s 4-tone system? What are the tone correspondences among Arem, Ruc and Viet?

5. Tone Developments – A tentative tree diagram
I propose a tentative tree diagram of Vietic from a viewpoint of the development of tones.
(6) *Proto-Vietic (toneless, contrastive phonation types)

Arem *Proto-RucViet (4 tones)
Ruc *Proto-VietMuong (6 tones)

South Viet, Central Viet (5 tones) Muong, N-C Viet (5 tones) North Viet (6 tones)

I assume the tones developed from 0 > 4 > 6, in the same way as Arem, Ruc and Viet. 5-tone systems of Southern, Central, North-Central varieties of Vietnamese as well as Muong are assumed to be the result of merger of two tones of the *Proto-VietMuong, though categorisation of the two tones are not the same.

6. Word list
Word list is attached as an Appendix to this paper. I have found 109 cognate sets from the three materials and made them in a comparative list. On top from the left, after numbers and English glossary, there are transcriptions of Arem in Ferlus (1997), Ruc in Ferlus (1997), Ruc in Loi (1993) and Ruc in Phong et al. (1988) as in the original source. Next three columns are Vietnamese orthography and comparative tone numbers. In the last column, Ferlus’ 1997 reconstruction is shown as a reference. On left from the top, the 109 lexica are categorised into five groups by their coda types: (1) null coda, (2) fricative coda, (3) liquid coda, (4) stop codas, and (5) nasal codas. The colour further categorises the lexica within the coda types: coloured cells show basic regular correspondences, and colourless cells show irregular correspondences. From next section, I will demonstrate several typical examples of regular and irregular correspondences.

7. Regular Tone Correspondences (79/109 cases)
I will explain the regular tone correspondences according to the six tones of Standard Northern Vietnamese. This regular correspondence applies to 79 cases out of 109 lexica. The number after Ruc and Viet IPA transcription refers to its tone.

7.1 Ngang tone (1A)

<table>
<thead>
<tr>
<th>No.</th>
<th>Gloss.</th>
<th>Arem</th>
<th>Ruc</th>
<th>Viet</th>
<th>*P-V(Ferlus)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>‘three’</td>
<td>па:</td>
<td>па: 1</td>
<td>ba: 1A</td>
<td>па:</td>
</tr>
<tr>
<td>47</td>
<td>‘fly (verb)’</td>
<td>пал</td>
<td>пал 1</td>
<td>bāj 1A</td>
<td>пəр</td>
</tr>
<tr>
<td>78</td>
<td>‘cooked rice’</td>
<td>кəм</td>
<td>кəм 1</td>
<td>кəм 1A</td>
<td>кəм</td>
</tr>
</tbody>
</table>

Vietnamese tone 1A corresponds to Ruc tone 1 in null coda, liquid coda and nasal coda syllables. Liquid coda in Arem and Ruc corresponds to the glide coda [j] in Vietnamese.