IS THERE A PROSODIC WORD IN VIETNAMESE?
PROSODIC CONSTITUENTS IN VIETNAMESE

Phạm 2008

(ι)

(PPh)

ω

φ

σ

μ

Schiering et al. 2010

(ι)

PPh

PPh

σ

(μ)
ROADMAP

- The word in Vietnamese
- Structure of the Vietnamese lexicon
  - Native vocabulary
  - Sino-Vietnamese
  - Loanwords
- Word stress in Vietnamese as evidence for the prosodic word (and the foot)
  - Previous proposals
  - An experiment
Traditional Vietnamese grammar has no notion of word
- Basic unit is the chữ, a monosyllabic morpheme

Problems with the definition of the Vietnamese word raised explicitly since the 1950s (Honey 1956; Thomas 1962; Thompson 1963, 1965; Cao 1985)
Traditional Vietnamese grammar has no notion of word
- Basic unit is the chữ, a monosyllabic morpheme

Thomas 1962 is the most explicit
- No grammatical word: just morphemes
- No prosodic word: just syllables and phonological phrases
- No orthographic word: Attempts at marking wordhood with hyphens in the orthography in South Vietnam (before 1975) were never consistent
AN OVERSIMPLIFIED SKETCH OF THE VN LEXICON

• Native vocabulary is largely composed of monosyllabic roots
• No affixation, but lots of compounds
  - Coordinate compounds (dvandva) are not headed and reversible
    bàn ghế [table + chair] ‘furniture’
    ông bà [gr.father+gr.mother] ‘grandparents’
  - Subordinate compounds are left-headed, just like phrases
    sân bay [yard + fly] ‘airport’
    cá heo [fish + pig] ‘dolphin’
COMPOUNDS OR PHRASES?

• Native compounds obey the same morphosyntactic processes as phrases (Noyer 1998)
  - Left-headedness and rightwards recursivity

Hoa hồng [flower + pink] ‘rose’ or ‘pink flower’
Sân bay [yard + fly] ‘airport’ or ‘flying yard’

- Nguyễn and Ingram (2007a) have shown that there is no phonetic difference between compounds and corresponding phrases
  • Except in contrastive condition
REDUPLICATION, REAL AND FALSE

• Productive reduplication: either right or left-headed
  \textit{mạnh} ‘strong’
  \textit{mạnh mể} ‘very strong’
  \textit{mành mạnh} ‘strong-ish’

• False reduplication (lexicalized)
  \textit{nhát} ‘timid’ \textit{nhút nhát} ‘very timid’
  \textit{nóng}‘hot’ \textit{nóng nực} ‘very hot’

• Ideophones: polysyllabic morphemes?
  \textit{lung tung} ‘confused, disordered’
  \textit{oàm oạp} ‘like waves breaking on the shore’
WHAT ABOUT SINO-VIETNAMESE?

• Sino-Vietnamese subordinate compounds are right-headed
  $\text{tuần lộc}$ [docile? + deer?] ‘reindeer’
  $\text{quốc ngữ}$ [country + language] ‘national script’

• Probably fossilized compounds
  - Meaning usually, but not always, opaque
  - Yet morphemic structure roughly transparent (1 syllable = 1 morpheme)

• Speakers shouldn’t be able to distinguish subordinate and coordinate compounds
UNAMBIGUOUSLY POLYSYLLABIC:
LOANWORDS

• Borrowed morphemes can be polysyllabic
  - French
    so cua  < roue de secours  ‘back-up suitor’
    ban công  < balcon  ‘balcony’
    la va bô  < lavabo  ‘sink’
    phô tô cóp bi  < photocopie  ‘photocopy’
  - Place names from ethnic minority languages
    Sài Gòn  < Khmer prɛj nokor  ‘City of the forest’
    Phan Rang  < Cham panduranga
3 POSITIONS ON POLYSYLLABLE STRESS IN VN

• Morphosyntactically defined (Cao 1978; Ngô 1984)
  – Coordinative compounds: both syllables stressed
  – Subordinative compounds: non-head (SV: left, NV: right)
  – Reduplicants: base
  – Loanwords: first syllable

• Rightmost stress (Nguyễn and Ingram 2007a; 2007b, Phạm 2008)
  – Not phrase-final lengthening (Phạm: final iambic foot)

• No word stress, only phrasal
  – Final syllable of phrase (Thomas 1962, Hoàng and Hoàng 1975)
  – Alternating stress on phrases (Thomas 1962)
  – Thompson (1963, 1965) has a complex, but contradictory system
STRESS EXPERIMENT:
NON-REDUPLICATED FORMS

• Word list containing disyllabic forms with all 5 Southern VN tones embedded in naturalistic sentences
  - Inventory  Ngô       Ng.+Ingram
    • NV coordinative compounds (2 orders)  Two stresses      Final stress
    • NV subordinative compounds           Final stress      Final stress
    • Opaque disyllables (2 SV coord., 1 SV sub., 1 French, 1 Khmer)  Initial stress  Final stress

• 2 contexts, to tease apart the effect of prosodic breaks
  - Phrase-final
  - Phrase-medial

Ex:  Nhà ba má Tuấn có nhiều bàn ghế (đẹp).
‘Tuấn’s parent’s house has a lot of (nice) furniture’
STRESS EXPERIMENT: PROTOCOL

• 8 speakers (4 male, 4 female), all university students (not in linguistics), all native speakers of Southern Vietnamese

• Recorded in a sound-proof booth in Hồ Chí Minh City (Saigon)

• Read the wordlist 5 times, with breaks between blocks

• Were given the list a few minutes before the experiment
  - Familiar with sentences
  - Recordings sound very natural (not reading style)
  - However, many sentences were mistaken for others
    • In coordinate compounds, it was difficult to get the least frequent order
STRESS EXPERIMENT

STATS

• Mixed Linear Model implemented in R (lmerTest package)

• Dependant variables
  - Syllable duration
  - Mean intensity of syllable
  - Mean f0 of syllable

• Fixed effects
  - Position (phrase-medial, phrase-final)
  - Word type (coordinate, reversed coordinate, subordinate, opaque)

• Random effects
  - Participants
  - Syllable
RESULTS
DURATION IN NON-REDUPLICATED WORDS

Fixed effects:

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>2.197e-01</td>
<td>1.3</td>
<td>1.6</td>
</tr>
<tr>
<td>PositionFinal</td>
<td>8.270e-01</td>
<td>2.9</td>
<td>2.8</td>
</tr>
<tr>
<td>WordtypeCoordYX</td>
<td>-7.992e-03</td>
<td>1.3</td>
<td>-6.3</td>
</tr>
<tr>
<td>WordtypeOp</td>
<td>-1.637e-02</td>
<td>1.3</td>
<td>-12.5</td>
</tr>
<tr>
<td>WordtypeSub</td>
<td>-4.276e-05</td>
<td>1.3</td>
<td>-3.3</td>
</tr>
<tr>
<td>order2</td>
<td>-1.351e-02</td>
<td>1.3</td>
<td>-10.4</td>
</tr>
<tr>
<td>PositionFinal:WordtypeCoordYX</td>
<td>1.117e-04</td>
<td>4.8</td>
<td>0.2</td>
</tr>
<tr>
<td>PositionFinal:WordtypeOp</td>
<td>-4.416e-03</td>
<td>4.4</td>
<td>-1.0</td>
</tr>
<tr>
<td>PositionFinal:WordtypeSub</td>
<td>2.474e-04</td>
<td>4.2</td>
<td>0.1</td>
</tr>
<tr>
<td>PositionFinal:order2</td>
<td>1.514e-02</td>
<td>4.1</td>
<td>0.4</td>
</tr>
<tr>
<td>WordtypeCoordYX:order2</td>
<td>2.819e-02</td>
<td>3.0</td>
<td>0.9</td>
</tr>
<tr>
<td>WordtypeOp:order2</td>
<td>5.040e-02</td>
<td>2.1</td>
<td>2.4</td>
</tr>
<tr>
<td>WordtypeSub:order2</td>
<td>6.467e-03</td>
<td>2.2</td>
<td>3.0</td>
</tr>
<tr>
<td>PositionFinal:WordtypeCoordYX:order2</td>
<td>4.163e-03</td>
<td>6.2</td>
<td>0.7</td>
</tr>
<tr>
<td>PositionFinal:WordtypeOp:order2</td>
<td>-1.034e-02</td>
<td>6.2</td>
<td>-0.2</td>
</tr>
<tr>
<td>PositionFinal:WordtypeSub:order2</td>
<td>-9.286e-03</td>
<td>5.9</td>
<td>-1.6</td>
</tr>
</tbody>
</table>

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’

R2 = 0.87
R2 of random factors and intercept = 0.30
RESULTS

INTENSITY IN NON-REDUPLICATED WORDS

Fixed effects:

<table>
<thead>
<tr>
<th>Term</th>
<th>Estimate</th>
<th>Std</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>66.50680</td>
<td></td>
</tr>
<tr>
<td>PositionFinal</td>
<td>-0.04043</td>
<td></td>
</tr>
<tr>
<td>WordtypeCoordYX</td>
<td>-1.19783</td>
<td></td>
</tr>
<tr>
<td>WordtypeOp</td>
<td>-0.19510</td>
<td></td>
</tr>
<tr>
<td>WordtypeSub</td>
<td>-0.50335</td>
<td></td>
</tr>
<tr>
<td>order2</td>
<td>-0.22114</td>
<td></td>
</tr>
<tr>
<td>PositionFinal:WordtypeCoordYX</td>
<td>1.34774</td>
<td></td>
</tr>
<tr>
<td>PositionFinal:WordtypeOp</td>
<td>0.21941</td>
<td></td>
</tr>
<tr>
<td>PositionFinal:WordtypeSub</td>
<td>0.10081</td>
<td></td>
</tr>
<tr>
<td>PositionFinal:order2</td>
<td>0.14744</td>
<td></td>
</tr>
<tr>
<td>WordtypeCoordYX:order2</td>
<td>1.35268</td>
<td></td>
</tr>
<tr>
<td>WordtypeOp:order2</td>
<td>-1.21426</td>
<td></td>
</tr>
<tr>
<td>WordtypeSub:order2</td>
<td>0.12877</td>
<td></td>
</tr>
<tr>
<td>PositionFinal:WordtypeCoordYX:order2</td>
<td>-1.62940</td>
<td></td>
</tr>
<tr>
<td>PositionFinal:WordtypeOp:order2</td>
<td>-0.09679</td>
<td></td>
</tr>
<tr>
<td>PositionFinal:WordtypeSub:order2</td>
<td>0.14295</td>
<td></td>
</tr>
</tbody>
</table>

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘  

R² = 0.678 
R² of random factors and intercept = 0.674
RESULTS

F0 IN NON-REDUPLICATED WORDS

Fixed effects:

| Term                        | Estimate | Std. Error | t value | Pr(>|t|) |
|-----------------------------|----------|------------|---------|---------|
| (Intercept)                 | 191.3817 | 12.8488    | 14.899  | < 2e-16 |
| PositionFinal               | -2.4842  | 0.5782     | -4.279  | 2.6e-05 |
| WordtypeCoordYX             | 13.9190  | 7.0462     | 1.978   | 0.051   |
| WordtypeOp                  | -4.1283  | 1.3336     | -3.105  | 0.0023  |
| WordtypeSub                 | 5.0617   | 2.0714     | 2.460   | 0.016   |
| order2                      | 7.6961   | 3.2032     | 2.408   | 0.017   |
| PositionFinal:WordtypeCoordYX | 4.0878  | 2.8796     | 1.416   | 0.163   |
| PositionFinal:WordtypeOp    | 3.9671   | 1.2435     | 3.190   | 0.0021  |
| PositionFinal:WordtypeSub   | 1.6844   | 0.9072     | 1.856   | 0.067   |
| PositionFinal:order2        | 0.4991   | 0.6892     | 0.724   | 0.470   |
| WordtypeCoordYX:order2      | -16.0659 | 7.5859     | -2.132  | 0.035   |
| WordtypeOp:order2           | -5.2155  | 1.9386     | -2.673  | 0.0084  |
| WordtypeSub:order2          | -12.6781 | 5.8651     | -2.158  | 0.034   |
| PositionFinal:WordtypeCoordYX:order2 | -7.2798 | 5.9369 | -1.223 | 0.227 |
| PositionFinal:WordtypeOp:order2 | -0.2013 | 1.0982 | -0.183 | 0.856 |
| PositionFinal:WordtypeSub:order2 | -2.8729 | 1.1743 | -2.451 | 0.015 |

---

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05

R2 = 0.898
R2 of random factors and intercept = 0.896
STRESS EXPERIMENT:
REDUPLICATED FORMS

• Reduplicated forms (all adjectives)
  - Inventory  \( Ngô \text{’} s \ predictions \)
    - Base in non-reduplicant (control)  \( Mới \ lạ \ ‘New \ and \ odd’ \)  \( Initial \ stress \)
    - Reduplicant + Base  \( Mới \ mới \ ‘rather \ new’ \)  \( Final \ stress \)
    - Base + Reduplicant in -eMôi mẻ  \( ‘very \ new’ \)  \( Initial \ stress \)
  - One base for each of the 5 Southern Vietnamese tones
  - All phrase-medial

• Compared with all other phrase-medial forms

• Same speakers, same methodology, same wordlist as previous experiment
RESULTS: REDUPLICATED FORMS ARE THE SAME AS OTHER DISYLLABLES
STRESS - SUMMARY

• No/little evidence of stress
  - No difference between most types of disyllables (contra Ngô 1984)
    • ...except Sino-Vietnamese compounds, which have longer 2nd syllables??

• Clear phrase-final lengthening (contra Nguyễn and Ingram 2007a, 2007b, pro Hoàng and Hoàng 1975; Thomas 1962)

• Absence of stress rules out the foot and most probably the prosodic word
ADDITIONAL EVIDENCE OF A PROSODIC WORD IN VIETNAMESE?

- No segmental processes sensitive to a word-size domain

- Cliticization (Pham 2008)
  - Extreme reduction of function word as evidence for the existence of a prosodic word
  - However, lexical words can also get dramatically reduced
PROSODIC AND GRAMMATICAL DOMAINS IN VN

- root-clauses, adjuncts, hesitations,…
- syntactic phrases, phases
- morphemes (words?)