What does emphatic lengthening tell us about binary length distinctions?

Aaron Braver*
TTU Cognitive Science Brown Bag

* Based on joint work with Shigeto Kawahara, Keio University

WHAT IS BINARY?
Why do linguists—and others—care?

Ferdinand de Saussure
• Binary oppositions are “the means by which units of language have value or meaning; each unit is defined against what it is not” (Fogarty 2005)

Roman Jakobson
• “The binary opposition is the child’s first logical operation” (Jakobson and Halle, 1956:47)
English Stop Consonants

- p, t, k, b, d, g

- [+voice]
  - p, t, k

- [-voice]
  - b, d, g

English Stop Consonants

- k, g

- [+velar]

- [-velar]

- p, t, b, d

No trickery allowed

- All vowels

- [+short]

- [-short]

- [+long]

- [-long]
Binary duration contrasts

• **Vowel length**
  – Japanese:

<table>
<thead>
<tr>
<th></th>
<th>obasan</th>
<th>obaasan</th>
</tr>
</thead>
<tbody>
<tr>
<td>“aunt”</td>
<td>“old lady”</td>
<td></td>
</tr>
<tr>
<td>ki</td>
<td>“tree”</td>
<td>kii</td>
</tr>
<tr>
<td>see</td>
<td>“height”</td>
<td>see</td>
</tr>
<tr>
<td>o</td>
<td>“tail”</td>
<td>oo</td>
</tr>
<tr>
<td>fu</td>
<td>“gluten”</td>
<td>fuu</td>
</tr>
</tbody>
</table>

• **Consonant length:**

<table>
<thead>
<tr>
<th></th>
<th>sacked</th>
<th>sakka</th>
</tr>
</thead>
<tbody>
<tr>
<td>“hill”</td>
<td>“author”</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>fatt o</th>
<th>fatt o</th>
</tr>
</thead>
<tbody>
<tr>
<td>“fate”</td>
<td>“fact”</td>
<td></td>
</tr>
</tbody>
</table>

Estonian: an exception?

• sata “hundred”
• saata “send!”
• saaata “to get”

• “Faced with a three-way surface contrast, a blatant *prima facie* insult to the phonological number two,” a number of authors have sought ways to say “this doesn’t count” (Prince 1980).

Why are length contrasts binary?

• **Option 1:** phonology just is binary

• **Option 2:** it’s hard to produce greater (e.g. ternary, quaternary...) distinctions

• **Option 3:** it’s hard to perceive more fine-grained distinctions

Emphatic lengthening

• That lecture was so boring
  sooo
  sooooo
  sooooooo
Emphatic lengthening in Japanese

- Adjectives lengthen their stem-final vowel to show emphasis

\[ \text{ita} + i = \text{itai} \]

\text{pain} \quad \text{adj.} \quad \text{painful} \quad \text{stem} \quad \text{suffix} \quad \text{adjective}

Procedure

- 7 female native Japanese speakers
- Shown stimuli in carrier sentences, 10 repetitions, randomized

\[ (6 \text{ adjectives} \times 6 \text{ emphasis levels} \times 10 \text{ blocks}) \]

An example...

A speaker’s production of “itai”, level 5 emphasis

<table>
<thead>
<tr>
<th>Japanese orthography</th>
<th>Transcription</th>
<th>Condition</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. いいたい</td>
<td>[itai]</td>
<td>no emphasis</td>
<td>‘painful’</td>
</tr>
<tr>
<td>b. いたーい</td>
<td>[itai]</td>
<td>level 1 emphasis</td>
<td>‘painful’ (emphatic)</td>
</tr>
<tr>
<td>c. いたーーい</td>
<td>[itaaai]</td>
<td>level 2 emphasis</td>
<td>‘painful’ (very emphatic)</td>
</tr>
<tr>
<td>d. いたーーーい</td>
<td>[itaaaaai]</td>
<td>level 3 emphasis</td>
<td>‘painful’ (very very emphatic)</td>
</tr>
<tr>
<td>e. いたーーーーい</td>
<td>[itaaaaaaai]</td>
<td>level 4 emphasis</td>
<td>‘painful’ (very * 3 emphatic)</td>
</tr>
<tr>
<td>f. いたーーーーーい</td>
<td></td>
<td>level 5 emphasis</td>
<td>‘painful’ (very * 4 emphatic)</td>
</tr>
</tbody>
</table>
Stats

- No pairwise comparisons, to avoid Type I error:
  - 6 emphasis levels * 7 speakers (* 3 vowel types)
- Post-hoc linear regressions
- 95% CI error bars

The 'best' speaker

\[ r = 0.89 \]

Second best

\[ r = 0.81 \]

More average

\[ r = 0.76 \]
The worst...

\[ r = 0.41 \]

Some things to notice...

- The “worst” speakers had the smallest range (533 ms for the worst vs. 975 ms for the best)
- All speakers showed a qualitative, binary distinction between no-emphasis and level 1

But...

- Japanese has a binary duration contrast
  - Does that make them better?
  - Does that make them more binary?

Experiment 2: English

- 7 target intensifier words:
  - very
  - too
  - way
  - super
  - mad
  - Really
- Placed in a carrier sentence:
  - That guy is sooo creepy
Emphasis levels

- 6 levels of emphasis, based on orthography:

<table>
<thead>
<tr>
<th>No emphasis</th>
<th>so</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1</td>
<td>soo</td>
</tr>
<tr>
<td>Level 2</td>
<td>sooo</td>
</tr>
<tr>
<td>Level 3</td>
<td>sooooo</td>
</tr>
<tr>
<td>Level 4</td>
<td>soooooo</td>
</tr>
<tr>
<td>Level 5</td>
<td>sooooooo</td>
</tr>
</tbody>
</table>

Procedure

- 8 female native English speakers
- Shown stimuli in carrier sentences, 10 repetitions, randomized

Results

- All speakers show correlation significant to $p < 0.001$ between emphasis level and duration

The star pupil

A speaker's production of "too", level 5 emphasis

The correlation coefficient $r = 0.75$
Some things to notice...

- The “worst” speakers had the smallest range

- All speakers showed a qualitative, binary distinction between no-emphasis and level 1

Why are length contrasts binary?

- Option 1: phonology just is binary

- Option 2: it’s hard to produce greater (e.g. ternary, quaternary...) distinctions

- Option 3: it’s hard to perceive more fine-grained distinctions
And now, a detour…

Vowel inventory size

Vowels of the world

Vowels of American English
The simplest vowel system

Inuktitut, Yidin

Always [i, a, u].

The default: 5 vowels

Spanish, Japanese

Usually [a, i, e, o, u] or their [+lax] counterparts.

7 vowel system

Italian, Yoruba

An improbable vowel system
The vowel dispersion principle

Most vowel systems tend to be evenly distributed in terms of perceptual space

[Diagram showing vowel pronunciation with dimensions of Front, Central, and Back, and High, Mid, and Low, with symbols representing vowel sounds like i, u, etc.]

So...

- Vowel quality is diffuse throughout perceptual space
- So is vowel length
- Like vowel quality, it’s a general trend, not a universal

Experiment 3: English listeners

- 24 native English speakers
- Did not participate in previous study
Stimuli

- Tokens selected from “top” 3 English speakers
- 3 speakers * 3 items * 6 emphasis levels
- Blocked by speaker, randomized within blocks

Confusion matrix

<table>
<thead>
<tr>
<th>Level of Stimulus</th>
<th>Listeners' response</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.39 0.14 0.14 0.25 0.24</td>
</tr>
<tr>
<td>1</td>
<td>35.06 5.80 2.92 1.18 1.21</td>
</tr>
<tr>
<td>2</td>
<td>12.21 28.50 20.98 11.17 4.41</td>
</tr>
<tr>
<td>3</td>
<td>5.44 33.19 35.11 32.93 26.57 21.37</td>
</tr>
<tr>
<td>4</td>
<td>2.94 19.97 26.10 34.42 16.62 10.23</td>
</tr>
<tr>
<td>5</td>
<td>0.98 6.85 11.83 18.41 28.38 34.16</td>
</tr>
</tbody>
</table>

% response per stimulus level