Covariation of voice onset time: a universal aspect of phonetic realization

Eleanor Chodroff¹, Alessandra Golden², and Colin Wilson²

¹Northwestern University, Department of Linguistics
²Johns Hopkins University, Department of Cognitive Science
Introduction

Extensive cross-linguistic variation in the realization of speech sounds

- Vowel formants
  e.g., Disner 1978, Lindau 1978, Manuel 1990

- Fricative COG
  e.g., Gordon 2002

- Vowel f0
  e.g., Whalen and Levitt, 1995

- Stop VOT
  e.g., Maddieson 1997, Cho & Ladefoged 1999
Cross-linguistic phonetic variation

\[
[k^h]
\]

**DORSAL**
- + spread glottis
- - continuant
...

**Phonetic implementation**

**TONGUE BODY**
- closed, velar

**GLOTTIS**
- spread

Cross-linguistic phonetic variation

[pʰ]

LABIAL
– continuant...
+ spread glottis

[kʰ]

DORSAL
– continuant...
+ spread glottis

[th]

CORONAL
– continuant...
+ spread glottis

What is the relational structure of cross-linguistic phonetic variation?

[pʰ] mean VOT range: 63 to 83 ms

[th] mean VOT range: 50 to 150 ms

[kʰ] mean VOT range: 73 to 154 ms

1) Do the VOTs of [pʰ], [tʰ], and [kʰ] vary independently of one another?
2) Is there consistency in the ordinal ranking of \([p^h]\), \([t^h]\), and \([k^h]\)?

\[\text{VOT}[p^h] < (\text{VOT}[t^h]) < \text{VOT}[k^h]\]

e.g., Maddieson 1997, Cho & Ladefoged, 1999
3) Is there a consistent linear relationship among [pʰ], [tʰ], and [kʰ]?

- Linear relationship is a simple type of patterned covariation
- Could imply ordinal relation (e.g., \( \text{VOT}[k^h] = \text{VOT}[p^h] + x, x \approx 17 \text{ ms} \))
Outline

1. Introduction

2. Cross-linguistic VOT survey

3. Uniformity constraint

4. Discussion

5. Future Directions
Cross-linguistic VOT survey

Large collection of previously reported stop VOT values

Examine relational structure of VOT among stops that have the same laryngeal feature specification*

* not just [+spread glottis], but also [-spread glottis], [-voice], [+voice], etc.
Methods

Examined ~350 theses, articles, grammars, and manuscripts
Collected stop VOT values from 164 sources

113 languages (149 dialects)
36 language families

Removed:
• Breathy / voiced aspirated
• Glottalized / ejective
• Tense (Korean)
• Implosives
• Palatal stops
• Uvular stops

Removed:
• Child data
• Explicitly labeled bilingual data
• L2 data

1671 VOT values remained for analysis
Methods

Averaged VOT data points with shared place and voice within each study, resulting in **1079 data points**

<table>
<thead>
<tr>
<th>Language Family</th>
<th>Languages</th>
<th>Data points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indo-European</td>
<td>Afrikaans, Armenian (Eastern), Assamese, Bengali, Catalan, Croatian, Danish, Dutch, English, French, Gaelic (Scots), German, Greek (Modern), Hindi, Icelandic, Italian, Kurmanji, Marathi, Nepali, Norwegian, Pahari, Panjabi, Pashto, Persian, Polish, Portuguese (Brazilian), Portuguese (European), Russian, Serbian, Sindi, Spanish, Swedish, Welsh</td>
<td>557</td>
</tr>
<tr>
<td>Sino-Tibetan</td>
<td>Bunun, Burmese, Cantonese, Fukienese, Galo, Hakha Lai, Hakka, Hokkien, Karen (Sgaw), Khonoma Angami, Kurtop, Mandarin, Stau, Taiwanese, Wu (Shanghainese)</td>
<td>106</td>
</tr>
<tr>
<td>Afro-Asiatic</td>
<td>Amharic, Arabic, Dahalo, Hebrew (Modern), Musey</td>
<td>41</td>
</tr>
<tr>
<td>Austronesian</td>
<td>Belep, Madurese, Malay, Tsou, Yapese</td>
<td>31</td>
</tr>
<tr>
<td>Niger-Congo</td>
<td>Bowiri, Igbo, Shekgalagari, Swati, Tswana, Zulu</td>
<td>39</td>
</tr>
<tr>
<td>Uralic</td>
<td>Finnish, Hungarian</td>
<td>21</td>
</tr>
<tr>
<td>Na-Dene</td>
<td>Apache (Western), Hupa, Navajo, Tlingit</td>
<td>19</td>
</tr>
</tbody>
</table>
## Methods

<table>
<thead>
<tr>
<th>Language Family</th>
<th>Languages</th>
<th>Data points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Korean</td>
<td>Korean</td>
<td>18</td>
</tr>
<tr>
<td>Tai-Kadai</td>
<td>Tai Khamti, Thai</td>
<td>18</td>
</tr>
<tr>
<td>Tupian</td>
<td>Arara, Munduruku</td>
<td>17</td>
</tr>
<tr>
<td>Dravidian</td>
<td>Tamil, Telegu</td>
<td>15</td>
</tr>
<tr>
<td>Quechuan</td>
<td>Quechua (Bolivian), Quechua (Cuzco), Quichua</td>
<td>15</td>
</tr>
<tr>
<td>Japanese</td>
<td>Japanese</td>
<td>14</td>
</tr>
<tr>
<td>Mayan</td>
<td>Itzaj Maya, Mam (Southern), Mopan Maya, Tzutujil, Yukateko Maya</td>
<td>14</td>
</tr>
<tr>
<td>Altaic</td>
<td>Azerbaijani, Turkish</td>
<td>12</td>
</tr>
<tr>
<td>Kartvelian</td>
<td>Georgian</td>
<td>12</td>
</tr>
<tr>
<td>Austro-Asiatic</td>
<td>Pnar, Remo</td>
<td>11</td>
</tr>
<tr>
<td>Oto-Manguean</td>
<td>Mazatec (Jalapa), Zapotec (Yalalog)</td>
<td>10</td>
</tr>
<tr>
<td>Burushaski</td>
<td>Burushaski</td>
<td>9</td>
</tr>
<tr>
<td>Algic</td>
<td>Ojibwe</td>
<td>6</td>
</tr>
<tr>
<td>Kordofanian</td>
<td>Moro</td>
<td>6</td>
</tr>
<tr>
<td>Muskogean</td>
<td>Chickasaw</td>
<td>6</td>
</tr>
<tr>
<td>Language Family</td>
<td>Languages</td>
<td>Data points</td>
</tr>
<tr>
<td>------------------------</td>
<td>------------------------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Northwest Caucasian</td>
<td>Kabardian</td>
<td>6</td>
</tr>
<tr>
<td>Pama-Nyungan</td>
<td>Warlpiri, Yan-Nhangu</td>
<td>6</td>
</tr>
<tr>
<td>Salishan</td>
<td>Montana Salish</td>
<td>6</td>
</tr>
<tr>
<td>Ticuna</td>
<td>Ticuna</td>
<td>6</td>
</tr>
<tr>
<td>Uto-Aztecan</td>
<td>Paiute (Northern), Ute</td>
<td>6</td>
</tr>
<tr>
<td>Wakashan</td>
<td>Kwakw'ala</td>
<td>6</td>
</tr>
<tr>
<td>Tucanoan</td>
<td>Waimaha</td>
<td>5</td>
</tr>
<tr>
<td>Eskimo-Aleut</td>
<td>Aleut (Eastern), Aleut (Western)</td>
<td>4</td>
</tr>
<tr>
<td>Chapacura-Wanham</td>
<td>Wari’</td>
<td>3</td>
</tr>
<tr>
<td>Creole</td>
<td>Hawaiian Creole</td>
<td>3</td>
</tr>
<tr>
<td>Ijoid</td>
<td>Defaka</td>
<td>3</td>
</tr>
<tr>
<td>Nakh-Dagestanian</td>
<td>Udi</td>
<td>3</td>
</tr>
<tr>
<td>Tangkic</td>
<td>Kayardild</td>
<td>3</td>
</tr>
<tr>
<td>Arauan</td>
<td>Banawa</td>
<td>2</td>
</tr>
</tbody>
</table>
Methods

Relied on primary source descriptions of the laryngeal specifications

Aggregate analyses

VOT categories
Negative: < 0 ms
Short-lag: > 0 ms and < 35 ms
Long-lag: > 35 ms

Kuhl & Miller 1975
Results

Variation in language-specific VOT means (ms)

Range: -161 to 117 ms

Range: -177 to 130 ms

Range: -144 to 154 ms

Median values

<table>
<thead>
<tr>
<th>category</th>
<th>labial</th>
<th>coronal</th>
<th>dorsal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative</td>
<td>-83 ms</td>
<td>-80 ms</td>
<td>-64 ms</td>
</tr>
<tr>
<td>Short-lag</td>
<td>14 ms</td>
<td>18 ms</td>
<td>30 ms</td>
</tr>
<tr>
<td>Long-lag</td>
<td>62 ms</td>
<td>65 ms</td>
<td>76 ms</td>
</tr>
</tbody>
</table>
Ordinal rankings

Place differences
Canonical order: VOT[labial] < VOT[coronal] < VOT[dorsal]

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Canonical order</th>
<th>Non-canonical order</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Place1 &lt; Place2</td>
<td>Place2 &lt; Place1</td>
</tr>
<tr>
<td>labial - coronal</td>
<td>76%</td>
<td>24%</td>
</tr>
<tr>
<td>coronal - dorsal</td>
<td>89%</td>
<td>11%</td>
</tr>
<tr>
<td>labial - dorsal</td>
<td>96%</td>
<td>4%</td>
</tr>
</tbody>
</table>

Linear relation

Aggregate analysis of language-specific VOT means (ms)

r = 0.98*

r = 0.97*

r = 0.97*
Linear relation

Long-lag VOT

\[ r = 0.83^* \]

\[ r = 0.78^* \]

\[ r = 0.79^* \]
Linear relation

Short-lag VOT

- **coronal**
  - $r = 0.66^*$

- **dorsal**
  - $r = 0.46^*$

- **labial**
  - $r = 0.62^*$
Linear relation

Negative VOT

\[ r = 0.95^* \]

\[ r = 0.90^* \]

\[ r = 0.84^* \]
Outline

1. Introduction
2. Cross-linguistic VOT survey
3. Uniformity constraint
4. Discussion
5. Future Directions
Uniformity constraint

Mapping from distinctive features to phonetic targets is not independent across segments within a language.
Uniformity constraint

Within the phonetic grammar of a language/talker, the phonetic targets corresponding to a phonological feature value \([\alpha F]\) are (ideally) identical for all segments that are specified \([\alpha F]\).
Uniformity constraint

Applied to long-lag stops:
Within a language/speaker, duration and timing of glottal opening gesture relative to stop closure interval should be uniform for all stops specified [+s.g.]

Maddieson 1997, Cho & Ladefoged 1999
Uniformity constraint

Previous research on VOT: Are place differences in VOT planned or automatic / mechanistic?

Several aerodynamic and biomechanical explanations for VOT variation by place of articulation

• Volume of cavity posterior and anterior to constriction
• Movement of articulators
• Extent of articulatory contact area
• Change of glottal opening area
• Fixed duration for glottal gesture timed relative to a single point in the closure

Maddieson 1997, Cho & Ladefoged 1999

Claim that differences are automatic presupposes that, for all stops within a laryngeal series, phonetic targets for the laryngeal feature are uniform

Westbury & Keating 1984, Keating 1985
Uniformity constraint

Can uniformity be reduced to other known effects and constraints on phonetic realization?

Talker physiology / aerodynamics

• Cross-linguistic evidence: even within a laryngeal subcategory (e.g., long-lag), it is **physically possible** to produce \([p^h]\) with a consistently longer VOT than \([k^h]\)

Perceptual dispersion

• VOTs of stop categories within a laryngeal series are **more similar** to one another than would be predicted by dispersion alone

Uniformity constraint

Applies strongly to languages and speakers, thereby ensuring cross-talker relational invariance / restricting individual differences

Each point = pair of VOT means (ms) for a speaker of American English

Chodroff & Wilson 2017
Summary

Strong evidence for a uniformity constraint operating on the phonetic implementation of stop consonant laryngeal features

- Evidence from VOT covariation cross-linguistically
- Evidence from VOT covariation across talkers of American English

Linear relation arises from underlying identity (or near-identity) in the phonetic implementation of laryngeal feature value within each series

- Uniform duration and timing of glottal gestures (abduction and adduction) relative to supralaryngeal closure
Future directions

Role of contrast

→ Does uniformity apply as strongly to ‘unpaired’ stops as to those with in minimal laryngeal contrasts (e.g., languages with /p t k/ but /b d/)

Examine cross-linguistic patterns for other features and segments

→ Is uniformity specific to stop VOT?
   Evidence from fricatives in American English and Czech Chodroff 2017

→ Do some languages deviate from uniformity (e.g., as the result of recent sound change)?

Relate to phonological theories of feature hierarchies

→ Identify natural classes (e.g., stops) strongly bound by uniformity
Thanks to...

Andries Coetzee (UMich), James Kirby (Edinburgh), Paul Morris (Iowa), Sharon Rose (UCSD) for generously providing data

Doug Whalen (CUNY, Haskins) Paul Smolensky (JHU, Microsoft)

Northwestern Phonatics Berkeley Phonetics and Phonology Forum ("Phorum")

Northwestern University Postdoctoral Professional Development Travel Grant

Thank you!