Mixed Design ANOVA

Final obstruent voicing in whispered speech

Marita Everhardt
Methodology & Statistics Linguistics Research

26 April 2016
Phonated voicing contrast

- e.g. *beat vs bead*

- **English:**
  - Primary cue vs secondary cues
  - Preceding vowel length
    - Ratio of 2:1 to 3:2  
      (Hogan & Rozsypal, 1980; Raphael et al., 1975)
    - Physiological effort vs linguistically determined  
      (Sharf, 1964)
    - More significant cue  
      (Raphael et al., 1975)

- **Dutch:**
  - Devoicing  
    (Ernestus & Baayen, 2003)
  - No vowel length difference  
    (Jongman et al., 1992)

- **English as an L2:**
  - Identifying L2-sounds as L1-categories  
    (Flege, 1987)
What’s in a whisper?

- Vocal fold vibration vs air exhalation (Collins & Mees, 2003)

Different acoustic characteristics

- Reduced amplitude of voiced segments (Ito et al., 2005)
- Amplitude vowels < amplitude consonants (Ito et al., 2005)
Whispered voicing contrast

- Contrast can still be identified (Dannenbring, 1980; Tartter, 1989)
- Learned as part of language structure (Sharf, 1964)
- Heavy reliance (Higashikawa et al., 2003; Sharf, 1964)
  - Vowels are longer (Sharf, 1964)
Aim of the study

› Investigate use of secondary acoustic cues
  • Whispered vs phonated
  • L1 vs L2

› Production task
Participants

› Two groups

› Dutch L2 speakers of English
  • N=20 (10 male, 10 female)
  • (University) students
  • “normal” exposure to English

› L1 speakers of English
  • N=8 (3 male, 5 female)
  • (University) students and teachers
  • Living in The Netherlands

› No hearing or reading disabilities
Stimuli

› Monosyllabic English minimal pairs
  • e.g. *beat* vs *bead*
  • pseudo words
  • phonated vs whispered

› Obstruents
  • fricatives (/f, v, s, z/)
  • stops (/p, b, t, d/)

› Vowels
  • KIT
  • FLEECE

› Total of 32 target words
Analysis

› Boundary placements
  • Praat
  • Phonated vs whispered
Whispered stop

Methodology

<table>
<thead>
<tr>
<th>Time (s)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>32.61</td>
<td>32.95</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time (s)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>32.6062751</td>
<td>32.9465156</td>
</tr>
</tbody>
</table>

pip
Phonated fricative

sis

Time (s)
Whispered fricative

 sis

Time (s)
Analysis

› Boundary placements
  • Praat
  • Phonated vs whispered

› Vowel length measurements
  • Praat (script)

› Mixed-design ANOVA
Mixed-design ANOVA

- Dependent variable
  - Mean Vowel Length (ms)

- Independent variables
  - Between subject
    - Group (NS vs NNS)
  - Within subject
    - Whisper (+Wh vs -Wh)
    - Obstruent (+Ct vs -Ct)
    - Voicing (+V vs -V)
Mixed-design ANOVA

str(WhisperData)

'data.frame': 224 obs. of 6 variables:
$ Participant : Factor w/ 28 levels "F1AH36","F1AM27",...: 1 1 1 1 1 1 1 1 2 2 ...
$ Group : Factor w/ 2 levels "NNS","NS": 2 2 2 2 2 2 2 2 2 2 ...
$ Whisper : Factor w/ 2 levels "+Wh","-Wh": 1 1 1 1 2 2 2 1 1 ...
$ Obstruent : Factor w/ 2 levels "+Ct","-Ct": 1 1 2 2 1 1 2 2 1 1 ...
$ Voicing : Factor w/ 2 levels "+V","-V": 1 2 1 2 1 2 1 2 1 2 ...
$ Vowel : num 257 142 311 258 262 ...

head(WhisperData, 8)

<table>
<thead>
<tr>
<th>Group</th>
<th>Participant</th>
<th>Whisper</th>
<th>Obstruent</th>
<th>Voicing</th>
<th>Vowel</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NS</td>
<td>F1AH36</td>
<td>+Wh</td>
<td>+Ct</td>
<td>+V</td>
</tr>
<tr>
<td>29</td>
<td>NS</td>
<td>F1AH36</td>
<td>+Wh</td>
<td>+Ct</td>
<td>-V</td>
</tr>
<tr>
<td>57</td>
<td>NS</td>
<td>F1AH36</td>
<td>+Wh</td>
<td>-Ct</td>
<td>+V</td>
</tr>
<tr>
<td>85</td>
<td>NS</td>
<td>F1AH36</td>
<td>+Wh</td>
<td>-Ct</td>
<td>-V</td>
</tr>
<tr>
<td>113</td>
<td>NS</td>
<td>F1AH36</td>
<td>-Wh</td>
<td>+Ct</td>
<td>+V</td>
</tr>
<tr>
<td>141</td>
<td>NS</td>
<td>F1AH36</td>
<td>-Wh</td>
<td>+Ct</td>
<td>-V</td>
</tr>
<tr>
<td>169</td>
<td>NS</td>
<td>F1AH36</td>
<td>-Wh</td>
<td>-Ct</td>
<td>+V</td>
</tr>
<tr>
<td>197</td>
<td>NS</td>
<td>F1AH36</td>
<td>-Wh</td>
<td>-Ct</td>
<td>-V</td>
</tr>
</tbody>
</table>
Boxplot

<table>
<thead>
<tr>
<th></th>
<th>NNS</th>
<th>NS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><img src="image1" alt="Boxplot NNS" /></td>
<td><img src="image2" alt="Boxplot NS" /></td>
</tr>
</tbody>
</table>

**Mean Vowel Length (ms)**

- Whisper
  - +Wh
  - −Wh

**Voicing**

- +V
- −V
Boxplot (with obstruents)

Mean Vowel Length (ms)

<table>
<thead>
<tr>
<th>Voicing</th>
<th>NNS</th>
<th>NS</th>
</tr>
</thead>
<tbody>
<tr>
<td>+V</td>
<td>Whisper +</td>
<td>Whisper -</td>
</tr>
<tr>
<td>-V</td>
<td>Whisper -</td>
<td>Whisper +</td>
</tr>
</tbody>
</table>

Whisper

-Wh

+Wh

Data
Assumptions - Normality

> Shapiro-Wilk’s test

<table>
<thead>
<tr>
<th></th>
<th>Shapiro-Wilk</th>
<th>N</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>+Wh+Ct+V</td>
<td>NS 0.944</td>
<td>8</td>
<td>0.647</td>
</tr>
<tr>
<td></td>
<td>NNS 0.961</td>
<td>20</td>
<td>0.566</td>
</tr>
<tr>
<td>+Wh+Ct-V</td>
<td>NS 0.971</td>
<td>8</td>
<td>0.909</td>
</tr>
<tr>
<td></td>
<td>NNS 0.964</td>
<td>20</td>
<td>0.636</td>
</tr>
<tr>
<td>+Wh-Ct+V</td>
<td>NS 0.938</td>
<td>8</td>
<td>0.589</td>
</tr>
<tr>
<td></td>
<td>NNS 0.937</td>
<td>20</td>
<td>0.215</td>
</tr>
<tr>
<td>+Wh-Ct-V</td>
<td>NS 0.980</td>
<td>8</td>
<td>0.962</td>
</tr>
<tr>
<td></td>
<td>NNS 0.942</td>
<td>20</td>
<td>0.267</td>
</tr>
<tr>
<td>-Wh+Ct+V</td>
<td>NS 0.880</td>
<td>8</td>
<td>0.190</td>
</tr>
<tr>
<td></td>
<td>NNS 0.940</td>
<td>20</td>
<td>0.236</td>
</tr>
<tr>
<td>-Wh+Ct-V</td>
<td>NS 0.800</td>
<td>8</td>
<td><strong>0.029</strong></td>
</tr>
<tr>
<td></td>
<td>NNS 0.907</td>
<td>20</td>
<td>0.056</td>
</tr>
<tr>
<td>-Wh-Ct+V</td>
<td>NS 0.944</td>
<td>8</td>
<td>0.648</td>
</tr>
<tr>
<td></td>
<td>NNS 0.953</td>
<td>20</td>
<td>0.407</td>
</tr>
<tr>
<td>-Wh-Ct-V</td>
<td>NS 0.933</td>
<td>8</td>
<td>0.543</td>
</tr>
<tr>
<td></td>
<td>NNS 0.942</td>
<td>20</td>
<td>0.263</td>
</tr>
</tbody>
</table>
Theoretical Quantiles
Sample Quantiles

+Wh+Ct+V

+Wh+Ct−V

+Wh−Ct+V

+Wh−Ct−V
Assumptions - Homogeneity

› Variance of between-subject
› Levene's test

  • `leveneTest(WhisperData$Vowel ~ WhisperData$Group)

Levene's Test for Homogeneity of Variance (center = median)

<table>
<thead>
<tr>
<th>Df</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>group</td>
<td>1</td>
<td>2.1806</td>
</tr>
<tr>
<td>222</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Assumptions - Sphericity

Variance of within-subject
- Only 2 levels
- “a within-subject factor with df=1 will always satisfy the sphericity condition” (Rietveld & van Hout, 1993).
Mixed-design ANOVA

\[ \text{WhisperModel} = \text{ezANOVA(data=WhisperData, dv=(Vowel), wid=(Participant), between=(Group), within=(Whisper, Obstruent, Voicing), type=3, detailed=TRUE)} \]

<table>
<thead>
<tr>
<th>Effect</th>
<th>DFn</th>
<th>DFd</th>
<th>SSn</th>
<th>SSd</th>
<th>F</th>
<th>p</th>
<th>p&lt;.05</th>
<th>ges</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>1</td>
<td>26</td>
<td>9,027597E+06</td>
<td>148467,872</td>
<td>1,580931E+03</td>
<td>8,133055E-25</td>
<td>*</td>
<td>9,774751E-01</td>
</tr>
<tr>
<td>Group</td>
<td>1</td>
<td>26</td>
<td>8,448149E+02</td>
<td>148467,872</td>
<td>1,479457E-01</td>
<td>7,036345E-01</td>
<td>*</td>
<td>4,044565E-03</td>
</tr>
<tr>
<td>Whisper</td>
<td>1</td>
<td>26</td>
<td>3,041496E+04</td>
<td>148467,872</td>
<td>1,479457E-01</td>
<td>7,036345E-01</td>
<td>*</td>
<td>4,044565E-03</td>
</tr>
<tr>
<td>Obstruent</td>
<td>1</td>
<td>26</td>
<td>1,659554E+05</td>
<td>13476,766</td>
<td>3,201687E+02</td>
<td>3,889494E-16</td>
<td>*</td>
<td>2,276988E-01</td>
</tr>
<tr>
<td>Group:Whisper</td>
<td>1</td>
<td>26</td>
<td>6,133432E+04</td>
<td>9669,445</td>
<td>1,649208E+02</td>
<td>9,187883E-13</td>
<td>*</td>
<td>7,741438E-05</td>
</tr>
<tr>
<td>Group:Obstruent</td>
<td>1</td>
<td>26</td>
<td>1,610589E+01</td>
<td>22644,198</td>
<td>1,849274E-02</td>
<td>8,928782E-01</td>
<td>*</td>
<td>9,683958E-05</td>
</tr>
<tr>
<td>Group:Voicing</td>
<td>1</td>
<td>26</td>
<td>2,014766E+01</td>
<td>13476,766</td>
<td>3,886979E-02</td>
<td>8,452400E-01</td>
<td>*</td>
<td>1,132450E-01</td>
</tr>
<tr>
<td>Whisper:Obstruent</td>
<td>1</td>
<td>26</td>
<td>4,250359E+03</td>
<td>4883,006</td>
<td>2,263142E+01</td>
<td>6,378791E-05</td>
<td>*</td>
<td>2,002222E-02</td>
</tr>
<tr>
<td>Whisper:Voicing</td>
<td>1</td>
<td>26</td>
<td>2,415395E+02</td>
<td>2460,529</td>
<td>2,552308E+00</td>
<td>1,222171E-01</td>
<td>*</td>
<td>1,159724E-03</td>
</tr>
<tr>
<td>Obstruent:Voicing</td>
<td>1</td>
<td>26</td>
<td>1,260756E+04</td>
<td>4297,582</td>
<td>7,627465E+01</td>
<td>3,284698E-09</td>
<td>*</td>
<td>5,714104E-02</td>
</tr>
<tr>
<td>Group:Whisper:Obstruent</td>
<td>1</td>
<td>26</td>
<td>5,539962E+02</td>
<td>4883,006</td>
<td>2,949802E+00</td>
<td>9,777838E-09</td>
<td>*</td>
<td>2,655964E-03</td>
</tr>
<tr>
<td>Group:Whisper:Voicing</td>
<td>1</td>
<td>26</td>
<td>1,071175E+02</td>
<td>2460,529</td>
<td>1,131893E+00</td>
<td>2,971535E-01</td>
<td>*</td>
<td>5,146445E-04</td>
</tr>
<tr>
<td>Group:Obstruent:Voicing</td>
<td>1</td>
<td>26</td>
<td>7,132076E+03</td>
<td>4297,582</td>
<td>4,314844E+01</td>
<td>5,770960E-07</td>
<td>*</td>
<td>3,314719E-02</td>
</tr>
<tr>
<td>Whisper:Obstruent:Voicing</td>
<td>1</td>
<td>26</td>
<td>6,294280E+02</td>
<td>2132,346</td>
<td>7,674707E+00</td>
<td>1,019996E-02</td>
<td>*</td>
<td>3,016508E-03</td>
</tr>
<tr>
<td>Group:Whisper:Obstruent:Voicing</td>
<td>1</td>
<td>26</td>
<td>1,142286E+03</td>
<td>2132,346</td>
<td>1,392806E+01</td>
<td>9,366409E-04</td>
<td>*</td>
<td>5,460935E-03</td>
</tr>
</tbody>
</table>
Voicing - Whisper

Results

$F(1, 26) = 2.55, p = .122, \eta^2 = 1.16 \times 10^{-03}$
Voicing - Group

\[ F(1, 26) = 7.14 \times 10^1, \quad p < .01, \quad \eta^2 = 1.13 \times 10^{-1} \]
Voicing - Whisper - Group

\[ F(1, 26) = 1.13 \times 10^{+00}, p = .297, \eta^2 = 5.15 \times 10^{-04} \]
Voicing - Obstruent - Group

\[ F(1, 26) = 4.31 \times 10^{-1}, p < .01, \eta^2 = 3.31 \times 10^{-2} \]
Voicing - Obstruent - Whisper

F(1, 26) = 7.67e+00, p < .05, \eta^2 = 3.02e-03
Conclusions

› Whispered vs phonated speech
  • No difference

› L1 vs L2 speakers of English
  • Significant difference

› Future research
  • Perception
  • Other cues
Thank you for your attention
References

› Sharf, D. J. (1964) Vowel duration in whispered and in normal speech. Language and Speech, 7(2), 89-97. doi:10.1177/002383096400700204