The role of duration in the clear speech vowel intelligibility benefit

Sarah Hargus Ferguson, Ph.D., CCC-A
University of Utah

Diane Kewley-Port, Ph.D.
Indiana University
Acknowledgments

• Current project
  – NIDCD-R03-08886 (PI: S.H. Ferguson)

• Development of Ferguson database
  – NIDCD-R01-02229 (PI: D. Kewley-Port)

• Schiefelbusch SPLH Clinic at KU

• Doug Kieweg, Leela Parimisetty, Kyung Ae Keum, Billy Speer, Jessica Stamey
Clear speech

• The style of speech a talker produces when his/her communication partner has a hearing loss

• How it’s elicited: Talkers read written materials in two conditions
  (1) In a conversational manner
  (2) As if talking to hearing-impaired listener
Clear speech

- It’s more intelligible than conversational speech
  - In most studies, normal-hearing and hearing-impaired listeners have enjoyed a similar CL benefit (range: 11-35 points)
What makes clear speech clear?

• **Acoustic characteristics** (e.g., Picheny et al., 1986):
  – It’s *slower* (more/longer pauses, longer individual phonemes)
  – Voice *pitch varies* more
  – Final *consonants are released* more often
  – Consonants have *more energy*
  – Vowels are *more distinct from each other*
What makes clear speech clear?

• Which of these changes actually contribute to the superior intelligibility of clear speech?

• Ways to study this
  (1) Signal-processing approach
    • i.e., make conversational speech “clear”
  (2) Talker-differences approach
    • i.e., compare talkers among whom the clear speech advantage varies, or look at within-talker differences
What makes clear *vowels* clear?

• Why vowels?
  – Limits the pool of potential acoustic correlates
  – They’re extremely important to overall intelligibility
Vowel acoustics

Steady-state formant frequencies
Vowel acoustics

High vowels

Low vowels

Front vowels

Back vowels
Vowel acoustics

Ferguson & Kewley-Port 2002
Vowel acoustics

Dynamic formant movement
Vowel acoustics

Duration

bead (305 ms)

bid (275 ms)
What makes clear *vowels* clear?

- **Talker-differences studies**
  - Using talkers selected from *Ferguson Clear Speech database* (2004), compared acoustic changes in talkers with big CL benefit (BB) versus NO CL benefit (NB)
    - Ferguson & Kewley-Port (2007): YNH listeners
    - Ferguson (2009): EHI listeners
Ferguson Clear Speech Database
BB versus NB talkers

![Bar chart showing comparison between BB and NB talkers. The chart displays RAU, EHI listeners on the y-axis. The x-axis represents BB and NB talkers. The chart shows a comparison between CL and CO listeners.]
What makes clear *vowels* clear?

- **Talker-differences** studies
  - For both listener groups, talkers who produced a larger CL benefit for vowels
  - Showed *greater vowel space expansion*
Vowel space expansion
What makes clear *vowels* clear?

- **Talker-differences studies**
  - For both listener groups, talkers who produced a larger CL benefit for vowels
    - Showed *greater vowel space expansion*
    - Showed *larger increases in vowel duration*
Increased vowel duration

![Bar chart showing increased vowel duration for BB and NB]

- BB: Duration (ms) for CL and CO
- NB: Duration (ms) for CL and CO
The current project

• Goals
  – Assess the role of duration, independent of steady-state formant changes, in increasing the intelligibility of vowels in clear speech
  – Assess whether this role depends on
    • Talkers’ conversational vowel durations
    • Listener age/ hearing status
Materials

- Talkers

<table>
<thead>
<tr>
<th></th>
<th>CL DUR</th>
<th>CON DUR</th>
<th>DIFF DUR</th>
<th>CL EHI RAU</th>
<th>CO EHI RAU</th>
<th>DIFF RAU</th>
<th>DUR ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>F06</td>
<td>288</td>
<td>173</td>
<td>115</td>
<td>85</td>
<td>73</td>
<td>12</td>
<td>1.7</td>
</tr>
<tr>
<td>M05</td>
<td>252</td>
<td>155</td>
<td>97</td>
<td>89</td>
<td>79</td>
<td>10</td>
<td>1.6</td>
</tr>
<tr>
<td>F19</td>
<td>336</td>
<td>306</td>
<td>30</td>
<td>92</td>
<td>91</td>
<td>1</td>
<td>1.1</td>
</tr>
<tr>
<td>M19</td>
<td>345</td>
<td>288</td>
<td>57</td>
<td>91</td>
<td>89</td>
<td>2</td>
<td>1.2</td>
</tr>
</tbody>
</table>
Materials

• /ʊ ʊ ə, ə, ʌ, æ, ɒ, ə/ in /bVd/ context
• For each talker, used STRAIGHT to create 8 sets of vowels with average durations of 125-300 ms in 25-ms steps – 125, 150, 175, 200, 225, 250, 275, & 300 ms
• Example: F06 (BB; mean CON dur = 173 ms)
Listeners

- Two groups, n = 20 each
  - YNH, aged 18-25 years
    - Screened at 20 dB HL for 250-8000 Hz
  - EHI, aged 66-88 years
    - Mild-to-moderate sloping SNHL
    - Passed MMSE (mean = 28.8)
    - Had successfully participated in previous vowel intelligibility study
Procedures

• /bVd/ words presented at 70 dB SPL in 12-talker babble
  • SNR = -10 dB for YNH, -3 dB for EHI
• Listeners identified vowels by choosing from 10 categories shown on the computer screen
• Everyone was familiarized with vowel ID task beforehand
Data analysis

- Two sets of **mixed 3-way ANOVAs**
  - **Listener-based**: Listener scores, averaged across the vowels, for each talker and set
  - **Vowel-based**: Vowel scores, averaged across the listeners in each group, for each talker and set
  - For both, percent scores were converted to RAU before analysis
Results

• Main effect of duration was significant
  – Intelligibility ↑ as duration ↑

• Main effect of listener group was significant
  – But it interacted with the main effect of talker
  – Importantly, it did not interact with duration
Results

• Main effect of **talker** was significant in the listener-based ANOVA, but not the vowel-based one
  – This was also true for the **talker X duration** interaction

• The **three-way interaction was never significant**
The diagram represents the relationship between mean duration and percent correct for two conditions: YNH and EHI. The x-axis represents the mean duration, ranging from 100 to 325, while the y-axis represents the percent correct, ranging from 40 to 100. The data points for YNH are indicated by red squares, and the data points for EHI are indicated by blue diamonds. The green squares highlight specific data ranges for each condition.
F06

Percent correct

Mean duration

YNH
EHI
Discussion

• When conversational vowels are long to begin with, making them shorter causes intelligibility to decrease

• This duration effect is independent of listener group

• Formant frequency changes improved performance for YNH but not EHI
Discussion

• When conversational vowels are short to begin with, making them longer causes intelligibility to increase – almost to the levels found for clear vowels!

• This duration effect is independent of listener group
Tentative conclusions

• Increases in vowel duration affect YNH and EHI listeners similarly.

• However, the importance of formant frequency changes seems to differ for the two groups.
So, what do you think?