Making Speech More Distinct

Presented by:
Donald J. Schum, PhD
Vice President, Audiology & Professional Relations
Oticon, Inc.

Moderated by:
Carolyn Smaka, Au.D., Editor-in-Chief, AudiologyOnline

TECHNICAL SUPPORT
Need technical support during event?
Please contact us for technical support at
800-753-2160

CEUs
CEU Total Access members can earn continuing education credit for participation in this course. Be sure to take the outcome measure following course completion to earn your CEUs, or contact us for more information or assistance: 800-753-2160

Making Speech More Distinct

Donald J. Schum, PhD
Vice President, Audiology & Professional Relations
Oticon, Inc.

DJS@Oticonusa.com
<table>
<thead>
<tr>
<th>Agenda</th>
</tr>
</thead>
</table>
| - The Nature of Speech Understanding  
- The Acoustics of Speech in Conversation  
- Naturally Produced Enhanced Speech  
- Requirements for Computer Enhancement  
- Some Past Examples  
- Looking Forward |

### Speech Importance Function

<table>
<thead>
<tr>
<th>Frequency (Hz)</th>
<th>Speech Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>200-20</td>
<td>1</td>
</tr>
<tr>
<td>201-30</td>
<td>2</td>
</tr>
<tr>
<td>301-40</td>
<td>3</td>
</tr>
<tr>
<td>401-50</td>
<td>4</td>
</tr>
<tr>
<td>501-60</td>
<td>5</td>
</tr>
<tr>
<td>601-70</td>
<td>6</td>
</tr>
<tr>
<td>701-80</td>
<td>7</td>
</tr>
<tr>
<td>801-90</td>
<td>8</td>
</tr>
<tr>
<td>901-100</td>
<td>9</td>
</tr>
<tr>
<td>110-120</td>
<td>10</td>
</tr>
<tr>
<td>121-130</td>
<td>11</td>
</tr>
<tr>
<td>131-140</td>
<td>12</td>
</tr>
<tr>
<td>141-150</td>
<td>13</td>
</tr>
<tr>
<td>151-160</td>
<td>14</td>
</tr>
<tr>
<td>161-170</td>
<td>15</td>
</tr>
<tr>
<td>171-180</td>
<td>16</td>
</tr>
<tr>
<td>181-190</td>
<td>17</td>
</tr>
<tr>
<td>191-200</td>
<td>18</td>
</tr>
<tr>
<td>201-300</td>
<td>19</td>
</tr>
<tr>
<td>301-400</td>
<td>20</td>
</tr>
<tr>
<td>401-500</td>
<td>21</td>
</tr>
<tr>
<td>501-600</td>
<td>22</td>
</tr>
<tr>
<td>601-700</td>
<td>23</td>
</tr>
<tr>
<td>701-800</td>
<td>24</td>
</tr>
<tr>
<td>801-900</td>
<td>25</td>
</tr>
<tr>
<td>901-1000</td>
<td>26</td>
</tr>
</tbody>
</table>
She said, "Give me a kiss". See Fred give me a fish.

Examples of Dynamic Based Contrasts

- fricatives versus affricates: sh vs. ch
- plosives versus fricatives: d vs. z
- dipthongs versus vowels: oi vs. o
- CV & VC transitions
"Why chew my shoe"

Envelope of Onset
“ch”

“sh”

CV F2 Transition

“pool”

“tool”


Agenda

- The Nature of Speech Understanding
- The Acoustics of Speech in Conversation
- Naturally Produced Enhanced Speech
- Requirements for Computer Enhancement
- Some Past Examples
- Looking Forward

Speech is Movement

Slip knot
How Do We Talk?

- Minimal Effort
  - Minimal Contrasts
  - Minimal Durations
  - Reductions
- Only as clear as we need to be
  - General human tendency: Conservation of energy
- We adjust to environment, but . . .
- Assume normally hearing listener
Speaker-to-Speaker Variability
Agenda

- The Nature of Speech Understanding
- The Acoustics of Speech in Conversation
- Naturally Produced Enhanced Speech
- Requirements for Computer Enhancement
- Some Past Examples
- Looking Forward

Clear Speech

- Natural acoustic changes which occur when a talker attempts to produce speech which is precise and accurate.
## Clear Speech: Research History

- Picheny, Braid, & Durlach: Mid 1980’s
  - coined the term, contrasted to “conversational style speech”
  - documented acoustic changes
  - documented intelligibility improvement

## Acoustic Changes

- Slower rate of speech
- More frequent and longer pauses
- Longer phoneme durations (consonants & vowels)
- More released word-final stops
- Greater differentiation of vowels
- Improved Consonant/Vowel ratio

---

"Tuborg is a famous Danish beer. So is Carlsberg."  
Conversational vs. Clear

- Slower Rate  
- Higher Level  
- More & Longer Pauses
“So is Carlsberg.”

Conversational

Clear

• Longer, more intense consonants (improved C/V)

Vowel Triangle

Low 2nd Formant

High 2nd Formant

/i:/

/i/ Low 1st Formant

/æ/ High 1st Formant

Back

Front

Effect on Intelligibility?

At least 15 to 20% in word recognition
Schum (1996) Study

- Elderly & young, inexperienced talkers
- 5-10 minutes of training and practice
- Materials played back in noise to patients with SNHL (mild-mod sloping)

Word Recognition Improvement

Average ≈ 19%

Released Word-final Stops

- Conversational
- Clear

PercentReleased

0 10 20 30 40 50 60 70 80 90 100

Young Elderly
Agenda

- The Nature of Speech Understanding
- The Acoustics of Speech in Conversation
- Naturally Produced Enhanced Speech
- Requirements for Computer Enhancement
- Some Past Examples
- Looking Forward

Automatic Speech Recognition

Accurate Automatic Speech Recognition
Robust
Automatic Speech Recognition

Automatic Speech Recognition:
How good is it?

How good is it?

0 dB S/N  +5 dB S/N  +10 dB S/N  +15 dB S/N  +20 dB S/N
1. The boy fell from the window.
2. The wife helped her husband.
3. Big dogs can be dangerous.
4. Her shoes were very dirty.
5. The player lost a shoe.
6. Somebody stole the money.
7. The fire was very hot.
8. She’s drinking from her own cup.
10. The car was going too fast.

+5 dB S/N Babble

1. The boy know from
2. the wife helped her husband
3. and gods in the game
4. or shoes.
5. Her player lost issue
6. somebody still for
7. my player was very
8. he's drinking from her own
9. secure gain from
10. the car was going to that

Sentences in Quiet

1. The boy fell from the window.
2. To myself from the window
3. the life helped her husband
4. big dogs can be dangerous
5. her shoes were very dirty
6. the player lost a shoe
7. somebody stole the money
8. the fire was very hot
9. cheese drinking from her own cop
10. the picture came from a book

Davis, 2002
Real Time
Automatic Speech Recognition

Multi-dimensional Enhancement Scheme
Natural Clear Speech

- Individual correlations between 15 x 3 acoustic variables and Clear Speech Performance benefit for 20 talkers
- Best single correlation .45
- Most < .2
- ... Stepwise analysis drove correlations to > .9
- ... but needed 9 or 10 dimensions

Naturalness

(“The Avatar Effect”)

Agenda

- The Nature of Speech Understanding
- The Acoustics of Speech in Conversation
- Naturally Produced Enhanced Speech
- Requirements for Computer Enhancement
- Some Past Examples
- Looking Forward
Clear Speech as a Signal Processing Technique

Can a computer based hearing instrument be taught to produce Clear Speech?

Can Clear Speech by Created by a DSP based Hearing Instrument?

- Naturally produced Clear Speech is an extremely complex combination of acoustic changes
  - Not just... increased C/V ratio
  - ... longer consonant durations
  - ... more pauses
  - ... changes in intonation contours (linguistically meaningful)

- DSP approach unlikely to mimic this effect
  - Finding/manipulating phonemes is difficult
  - Time marches on
  - Manipulating one or two dimensions is not enough
Phoneme Enhancement

/i/

/i/
Please jot down how much change I need.

Actual effect of Fast Acting Compression
Please jot down how much change I need.