Reconsidering Vowels as Mathematical and Statistical Entities

Douglas S. Bigham
douglas.s.bigham@gmail.com

American Dialect Society Annual Meeting
at the annual meeting of the Linguistic Society of America

6-9 January, 2011, Pittsburgh
A RETURN TO FIRST PRINCIPLES

○ What are VOWELS?
  ● perceptual entities
  ● contrastive items for the perception of DIALECTS
  ● contextually determined entities
  ● non-meaningful linguistic atoms

○ How do we investigate VOWELS?
  ● Sampling TOKENS from SPEAKERS of DIALECTS
  ● Measuring FORMANTS
  ● Comparing FORMANT values using STATISTICS

○ We must reconsider these methods.
SAMPLING TOKENS FROM SPEAKERS

- What and How Much information can we lose?
  - GOOSE – averaged across tokens and speakers
GOOSE — INDIVIDUAL SPEAKERS, TOKENS
GOOSE — INDIVIDUAL SPEAKERS, TOKENS
DIFFERENT WAYS OF VARYING
WHAT INFORMATION DO WE NEED?

- GOOSE varies from fully back to fully central
  - Variation is both WITHIN and BETWEEN individuals
  - Which form of variation is more important?
- GOOSE varies differently than other vowels
- Do these patterns of variation have meaning?

- What does it *mean* to average a percept?

- Is there an appropriate N for TOKENS or SPEAKERS?
  - A linguistics-driven statistical methodology...
CAN A VOWEL BE REDUCED TO F1xF2?

- Why do we use F1xF2?
  - Labov, Yaeger, Steiner (1972)

- DeLattre, Liberman, Cooper, & Gerstman (1952)
  - An Experimental Study of the Acoustic Determinants of Vowel Color; Observations on One- and Two-Formant Vowels Synthesized from Spectrographic Patterns
  - DLCG were measuring perception via Hz values
  - Modern sociophonetics measures Hz values...

- FORMANTS are continuous; VOWELS are discrete
  - DLCG used 120Hz chunks of F2
**Additional Issues with the F1xF2 Model**

- Should F1 and F2 be measured on the same scale?
  - F1 has less freedom for variation (space) than F2
  - F1 ~800Hz ; F2 ~ 1600Hz
  - jnd discrimination threshold = ~25Hz (Snodgrass, 1975)

- In F1, there are only about 800/25 or 32 possible distinct regions of perception; F2 = ~64
  - Yet we report values like 816Hz x 1507Hz...

- F1xF2 always co-varies for VOWELS in vowel-space
  - But not to the same extent for all vowels...
  - ...or for all speakers
F1xF2 Correlations for Different Vowels
F1xF2 Correlations for Different Vowels

Correlations of F1 & F2, females

Bar chart showing the correlations of F1 and F2 for different vowels.
**Are VOWELS Statistical?**

- Taleb, 2008 - Limits of Statistics

<table>
<thead>
<tr>
<th></th>
<th>Simple (Yes/No)</th>
<th>Complex (How much?)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thin-tailed, known distributions</td>
<td>Robust</td>
<td>Robust</td>
</tr>
<tr>
<td>Fat-tailed &amp; unknown distributions</td>
<td>Robust</td>
<td>FRAGILE! DO NOT STATISTIZE!</td>
</tr>
</tbody>
</table>

- What quadrant are VOWELS in?
THE POWER OF N IN SOCIOPHONETICS

- As N (sample size) increases, so does the chance of finding a significant difference.

- Is there an appropriate N for TOKENS or SPEAKERS?

- How many SPEAKERS, VOWELS, and TOKENS are enough? Is it possible to have too many?

- Comparing DRESS and TRAP using a basic $t$-Test.
- Note “Critical Difference”...
  ... below 25Hz is below jnd!
DIFFERENCE BETWEEN DRESS AND TRAP

The (statistical) difference between DRESS and TRAP

- Critical Difference
- Mean Difference
- Number of Speakers
- Number of Tokens
SUMMARY

- Reconsider our underlying principles
  - VOWELS are perceptual objects
  - Perceptual constraints must drive investigation

- Reconsider our methods
  - When can we use averages and when not? (GOOSE)
  - Do the same methods work for all VOWELS?
    - cf. diphthongs & monophthongs; front & back vowels
  - If we continue using F1xF2 we must establish a significance value (20Hz in F1; 40Hz in F2?)

- Reconsider our statistics
  - We need a linguistically-driven statistical method
  - Fewer speakers & tokens may be better than more
THANK YOU!

douglas.s.bigham@gmail.com