INTRODUCTION

Perception enhances spectral differences between a preceding acoustic context with a reliable spectral property (e.g., amplified frequency region, long-term spectral slope) and a subsequent target vowel sound, producing spectral contrast effects.

- Shifting F0 frequencies downward in a preceding sentence (swounding more [i]-like) increased the number of [i]-high F0 responses and vice versa (Ladefoged & Broadbent, 1957).
- Sentence filtered to emphasize lower frequencies (spectrum of [i]-minus-[e]) produced more [e] responses and vice versa (Watkins, 1991).

These contrast effects are very robust, influencing identification of vowels as well as stop consonants (Laing et al., 2012), fricatives (Watkins & Makin, 1998), and musical instruments (Stilp et al., 2010).

However, conditions conducive to contrast effects are still poorly understood:

- Description calibrates to narrowband (100 Hz) spectral regularities that are reliable across both context and target (Kietze & Kluender, 2008; Alexander & Kluender, 2010; Alexander & Stilp, 2014), are narrowband spectral peaks sufficient to elicit contrast effects as well?
- Do contrast effects require shifting formant ranges up or down to simulate different talkers, or is mere amplification of formant ranges sufficient?
- When are reliable spectral properties too weak (e.g., insufficient amplification) to influence speech perception?
- Finally, what acoustic properties of these spectral regularities predict not only the presence of a contrast effect, but its magnitude as well?

Spectral contrast effects in vowel identification were explored using precursors filtered to have reliable narrowband (100 Hz) or broadband (300 Hz) peaks, or the difference between vowel spectral envelopes. Results reveal the efficacy of different spectral regularities in influencing speech perception.

METHODS

Precursor
- “Please say what this vowel ic.” said by CS (2174 ms)
- Vowels
  - Natural vowels linearly interpolated from [i] to [e] using PRAAT (246 mc)
  - [i] endpoint: F0 = 100 Hz, F1 = 400–940 Hz, F2 = 2000–1800 Hz
  - [e] endpoint: F0 = 100 Hz, F1 = 580–950 Hz, F2 = 1800–1700 Hz
- Filters
  - Narrowband (NB): 100 Hz bandwidth (250–550 Hz)
  - Broadband (BB): 300 Hz bandwidth (100–850 Hz)
  - Spectral Envelope Difference (SED): spectral envelopes for [i] and [e] endpoints derived via FFT then subtracted from one another
  - Filter gains set to .5 / 10 / 15 / 20 dB (NB, BB) or 25% / 50% / 75% / 100% of total filter power (SED)
- Participants
  - All native English speakers with normal hearing
  - n = 11 (NB20, BB20, SED100%), n = 14 (NB15, NB10, NB5), n = 11 (BB15, BB10, BB5), n = 11 (SED75%, SED50%, SED25%)
- Procedure
  - All precursors – vowel pairs presented dichotically at 70 dB SPL via circumaural headphones in sound isolating booths
  - Filter types were blocked and tested in random orders

RESULTS

- Logistic regressions were fit to each listener’s identification curves. Midpoints were calculated from these regression functions.
- Shifts in midpoints across filtering conditions (i.e., translation along the abscissa) measures the magnitude of the contrast effect.
- Midpoint shifts were analyzed using paired-sample t-tests (Bonferroni correction for multiple analyses within each participant group; n = .05 / .3 = .047).
- Mean responses are plotted below, with filters that processed the precursor depicted in inset. * indicates statistically significant midpoint shifts; error bars depict ± SEM.

NARROWBAND FILTERS

<table>
<thead>
<tr>
<th>Vowel Bandwidth (kHz)</th>
<th>Midpoint Shift</th>
<th>Peak Amplification (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NB20</td>
<td>1.07 steps*</td>
<td>0.644, p &lt; .001</td>
</tr>
<tr>
<td>NB15</td>
<td>0.80 steps*</td>
<td>0.596, p &lt; .001</td>
</tr>
<tr>
<td>NB10</td>
<td>0.53 steps*</td>
<td>0.444, p &lt; .01</td>
</tr>
<tr>
<td>NB5</td>
<td>-0.17 steps*</td>
<td>0.081, p &lt; .43</td>
</tr>
</tbody>
</table>

BROADBAND FILTERS

<table>
<thead>
<tr>
<th>Vowel Bandwidth (kHz)</th>
<th>Midpoint Shift</th>
<th>Peak Amplification (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BB20</td>
<td>1.67 steps*</td>
<td>0.653, p &lt; .001</td>
</tr>
<tr>
<td>BB15</td>
<td>1.90 steps*</td>
<td>0.653, p &lt; .001</td>
</tr>
<tr>
<td>BB10</td>
<td>1.10 steps*</td>
<td>0.635, p &lt; .0001</td>
</tr>
<tr>
<td>BB5</td>
<td>0.77 steps*</td>
<td>0.465, p &lt; .01</td>
</tr>
</tbody>
</table>

SPECTRAL ENVELOPE DIFFERENCES

<table>
<thead>
<tr>
<th>Vowel Bandwidth (kHz)</th>
<th>Midpoint Shift</th>
<th>Peak Amplification (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SED100%</td>
<td>1.09 steps*</td>
<td>0.599, p &lt; .001</td>
</tr>
<tr>
<td>SED75%</td>
<td>0.93 steps*</td>
<td>0.482, p &lt; .001</td>
</tr>
<tr>
<td>SED50%</td>
<td>0.50 steps*</td>
<td>0.286, p &lt; .017</td>
</tr>
<tr>
<td>SED25%</td>
<td>0.63 steps*</td>
<td>0.317, p &lt; .01</td>
</tr>
</tbody>
</table>

CONCLUSIONS

- Spectral contrast effects in vowel identification are demonstrated for very modest peaks in the precursor spectrum (as little as 5 dB) or very narrowband frequency regions (100 Hz-wide).
- Reveals acute sensitivity to a very broad range of spectral properties of a listening context.
- Contrast effects may be more pervasive in speech perception than previously thought.
- Total filter power is the most complete predictor of contrast effect magnitude.
- Explains tradeoff between bandwidth and peak amplification for NB and BB regularities (NB20 results = BB10 results; NB10 = BB5).
- Contrasts effects are observed for spectral regularities in speech that are not directly derived from speech or vocal tracts (NB, BB).
- Rejects classic findings by Ladefoged and Brandwein (1957), consistent with general auditory account of speech perception (e.g., Laing et al., 2012).
- Generalizability of results across speech and nonspeech targets, speech and nonspeech precursors, and a wide variety of reliable spectral properties reveals optimizing sensitivity to change is a fundamental operating characteristic of the auditory system.

REFERENCES

Alexander & Kluender (2010), JASA
Anderson & Stilp (2014), JASA
Kietze & Kluender (2008), JASA
Ladefoged & Brandwein (1957), JASA
Stilp, Alexander, Kietze, & Kluender (2010), JASA
Watkins (1991), JASA
Watkins & Makin (1998), JASA