The (Null) Effect of Spectral Estimator on Estimates of Spectral Moments
Patrick Reidy  The Ohio State University, Dept. of Linguistics

Purpose of Study

- The spectrum of a sibilant fricative is "noisy" and difficult to estimate accurately (Skade, 2004).
- To improve the estimation of sibilant fricatives' spectra, recent work has argued for the adoption of "reduced-variance estimators" (e.g., Buishand, 2004).
- However, spectral estimation is not the endpoint of a linguistic analysis, as the spectral estimate is almost always reduced to a small number of measures that describe its shape properties, such as spectral moments (e.g., Jongman, Wayland & Wong, 2000).
- Previous work has found no significant effect of spectral estimator on estimates of peak and centroid frequency for adult productions of English /s/ and // (Beckman & Beckman, 2012).
- Current study: Investigates the effect of spectral estimator on estimates of the first four spectral moments from adults' and children's productions of English sibilants.

Background

- Two commonly used spectral estimators are the discrete Fourier transform (DFT) and the multitaper spectrum (MTS).
- The MTS is equal to the pointwise average of K DFTs computed from data that have been windowed by discrete prolate spheroidal sequences (Thomson, 1982).

Participants

<table>
<thead>
<tr>
<th>Age group</th>
<th>No. participants (males)</th>
<th>No. /s/ tokens (males)</th>
<th>No. // tokens (males)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adults</td>
<td>20 (10)</td>
<td>297 (155)</td>
<td>300 (155)</td>
</tr>
<tr>
<td>5-year-olds</td>
<td>20 (9)</td>
<td>212 (90)</td>
<td>252 (111)</td>
</tr>
<tr>
<td>6-year-olds</td>
<td>21 (11)</td>
<td>191 (89)</td>
<td>265 (130)</td>
</tr>
<tr>
<td>7-year-olds</td>
<td>20 (10)</td>
<td>153 (84)</td>
<td>189 (96)</td>
</tr>
<tr>
<td>24-year-olds</td>
<td>19 (11)</td>
<td>120 (73)</td>
<td>82 (46)</td>
</tr>
<tr>
<td>Total</td>
<td>100 (51)</td>
<td>973 (488)</td>
<td>1088 (527)</td>
</tr>
</tbody>
</table>

Elicitation

- Picture-prompted word-repetition task.
- Word-initial, pre-vocalic tokens of /s/ and //.

Recording

- Recorded at 44 kHz.
- Frication onset and offset marked by hand.
- Phonemically transcribed; excluded if incorrect.

Spectral analysis

- DFT & MTS estimated from central 40 ms.
- Centroid, variance, skewness & kurtosis computed within the 32−15 kHz band.

Comparison of estimator and place effects on centroid and skewness, by age group

- Across studies, centroid and skewness are the moments that most consistently differenten /s/ from // (see Koenig, Skade, Pisoni & Moonen, 2011).
- Two-year-olds produce /s/ and // relatively close together in terms of centroid and skewness; however, even for this group, the estimator effect is a fraction of the place effect.

Results and Analysis

Paired t-tests revealed an effect of estimator on the even, but not the odd moments.

- Caveat: when tested with all /s/ and // tokens included and the moments computed across the entire frequency range (0−22.05 kHz), the estimator effect did not reach significance (at the Bonferroni corrected α = .05 level) for any of the moments.

Magnitude of estimator effect is dwarfed by place-of-articulation effect for all moments.

- Place effect: μ( /s/) − μ(//) for MTS & DFT.
- Estimator effect: μ(ESTIMATOR−DFT) averaged across all tokens.
- Choice of spectral estimator does not seem to affect the ability to distinguish /s/ from // in terms of any of the spectral moments.

- MTS and DFT give comparable estimates of the place effect for all moments.
- For most moments, the estimator effect is an order of magnitude smaller than either place effect.

Acknowledgements

- Data collection was supported by NIDCD grant R01 120932 to Jan Edwards (learningtotalk.org).
- This work was supported by funding from the OSU Center for Cognitive Science.
- Special thanks are due to Mary Beckman, whose advice greatly improved this material, and to the Phonics discussion group at OSU.