The (Null) Effect of Spectral Estimator on Estimates of Spectral Moments

Patrick Reidy *The Ohio State University, Dept. of Linguistics*

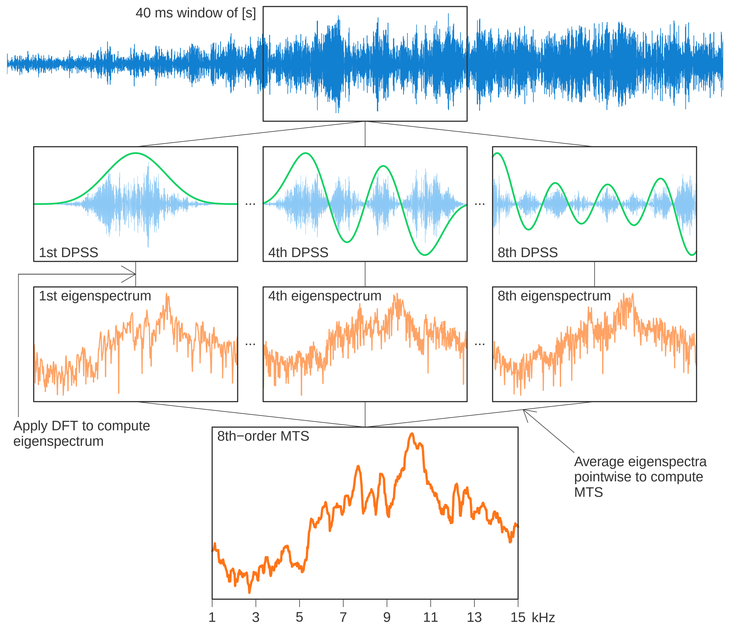
THE OHIO STATE UNIVERSITY

Purpose of Study

- The spectrum of a sibilant fricative is "noisy" and difficult to estimate accurately (Shadle, 2006).
- To improve the estimation of sibilant fricatives' spectra, recent work has argued for the adoption of "reduced-variance estimators" (e.g., Blacklock, 2004).
- However, **spectral estimation is not the endpoint of a linguisic 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 /ʃ/ (Reidy & 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.

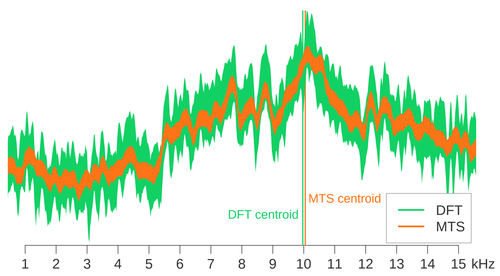
Background

- 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).



Variance properties of MTS & DFT

- Each ordinate of the MTS has $1/K^{\text{th}}$ the variance of the corresponding ordinate of the DFT (Percival and Walden, 1993).
- Hence, at each frequency, the magnitude of the spectrum is estimated more accurately with the MTS than with the DFT.

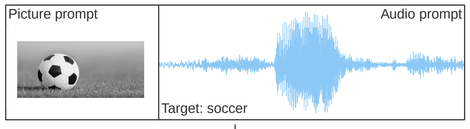


Experiment

Participants

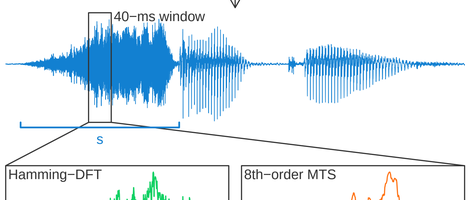
Age group	No. participants (males)	No. /s/ tokens (males)	No. /ʃ/ tokens (males)
Adults	20 (10)	297 (150)	300 (150)
5-year-olds	20 (9)	212 (90)	252 (111)
4-year-olds	21 (11)	191 (89)	265 (130)
3-year-olds	20 (10)	153 (84)	189 (96)
2-year-olds	19 (11)	120 (73)	82 (40)
Total	100 (51)	973 (486)	1088 (527)

Elicitation



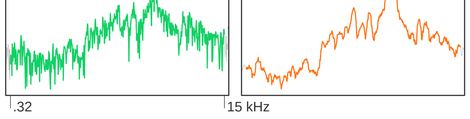
- Picture-prompted word-repetition task.
- Word-initial, pre-vocalic tokens of /s/ and /ʃ/.

Recording



- Recorded at 44.1 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.

Results and Analysis

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

Moment	$t(2060)$	μ(MTS − DFT)	p-value
Centroid	−1.772	−7.241	0.076
Variance	5.599	6.301×10^{4}	2.443×10^{-8}
Skewness	−1.597	−0.006	0.110
Kurtosis	−3.01	−0.068	0.002

- Caveat: when tested with all /s/ and /ʃ/ 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 $\alpha = .05$ level) for any of the moments.

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

- Place effect: $|\mu(/s/) - \mu(/ʃ/)|$, for MTS & DFT.
- Estimator effect: $\mu(|\text{MTS} - \text{DFT}|)$, averaged across all tokens.

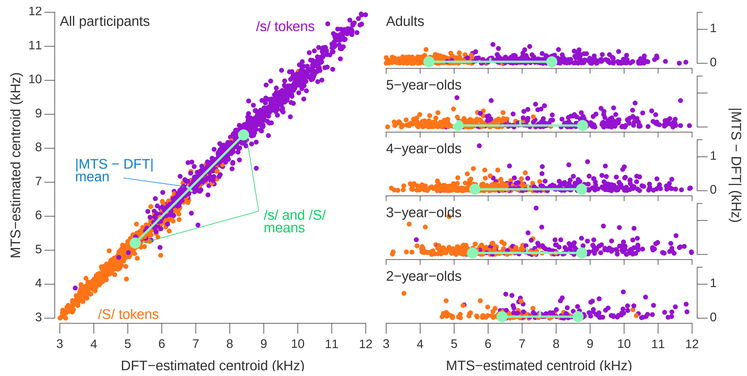
Moment	MTS place effect	DFT place effect	Estimator effect
Centroid	3.304×10^{3}	3.311×10^{3}	1.226×10^{2}
Variance	2.027×10^{6}	1.967×10^{6}	3.127×10^{5}
Skewness	1.175	1.178	0.130
Kurtosis	1.221	1.167	0.499

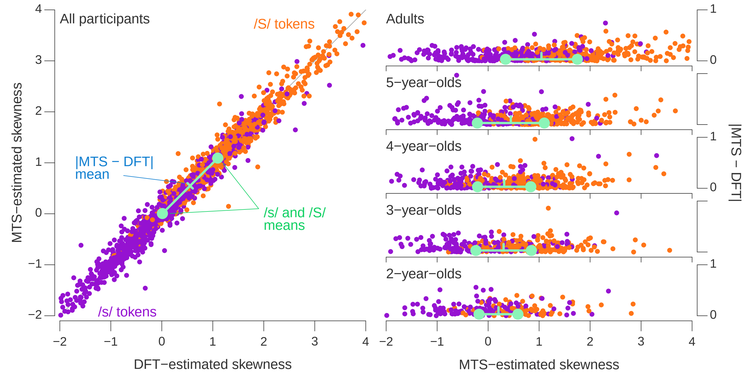
- Choice of spectral estimator does not seem to affect the ability to distinguish /s/ from /ʃ/ in terms of any of the spectral moments.
  - MTS and DFT give comparable estimates of the place effect for all moments.
  - For most moments, the estimator effect is an order of magnitude smaller than either place effect.

Results and Analysis

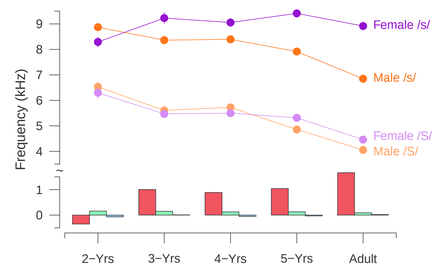
Comparison of estimator and place effects on centroid and skewness, by age group

- Across studies, centroid and skewness are the moments that most consistently differentiate /s/ from /ʃ/ (see Koenig, Shadle, Preston & Mooshammer, 2013).
- Two-year-olds produce /s/ and /ʃ/ relatively close together in terms of centroid and skewness; however, even for this group, the estimator effect is a fraction of the place effect.





Gender effects, when present, also dwarf the effect of spectral estimator.



- The /s/-/ʃ/ centroid contrast is used, from a young age, to signal gender identity (see also Fox & Nissen, 2005).
- This **gender effect** (i.e., difference in category means between females and males) seems insensitive to choice of spectral estimator given its size relative to the **estimator effect** and to the **difference between the MTS & DFT** gender effects.

Acknowledgements

- Data collection was supported by NIDCD grant R01 02932 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 Phonies discussion group at OSU.