SELECTING A METRIC FOR QUANTIFYING PRESCHOOLERS' TONGUE SHAPE COMPLEXITY USING ULTRASOUND IMAGING

Heather Campbell, MS, CCC-SLP; Tara McAllister Byun, PhD, CCC-SLP
New York University, Department of Communicative Sciences and Disorders

INTRODUCTION

✧ Speech sound disorders can have a negative impact on academic, social, or psycho-emotional development.
✧ Speech sounds differ in tongue shape complexity.
✧ An objective measure for classifying productions into categories of lingual complexity serve as a predictor of persistence of errors and/ or treatment responsiveness.
✧ Ultrasound imaging may be more versatile and less invasive and costly than alternative visualization methods, such as electropalatography (Gibbon, 1999).
✧ **Goal of this study: Find the ultrasound measure that best correlates with lingual complexity categories identified in a previous study of adults (Dawson et al., 2016).**

BACKGROUND

✧ Dawson et al. (2016) compared a variety of specific metrics of lingual contours in adults in order to determine which was best at predicting complexity class and specific phoneme.

✧ Speech sounds were categorized as either low, medium, or high complexity according to a hypothesized standard.

- Low complexity /ɑ/, /æ/, /ɪ/, /ʌ/, /ɛ/ (unrounded vowels; single lingual constriction)
- Medium complexity /w/, /u/, /j/, /g/ (lip rounding; lateral bracing; dorsal constriction)
- High complexity: /d/, /l/, /ɹ/, /θ/, /ʒ/ (tongue tip or multiple lingual constrictions)

✧ Midsagittal images of speech sounds produced by 6 adults.

Compared three distinct techniques for transforming raw spatial coordinates:

✧ Modified curvature index (MCI)
Averaging technique, integrates curvature with length of the arc (Stolar and Gick, 2013); takes integral of half of the difference between two adjacent points until the difference is minimized.

✧ Procrustes Analysis
Minimizes sum of squared differences using translation, rotation, and scaling; computes difference between each frame compared with resting state (Goodall, 1991); then compared averaged image with next instance of that target sound to eventually compute the average shape for that sound.

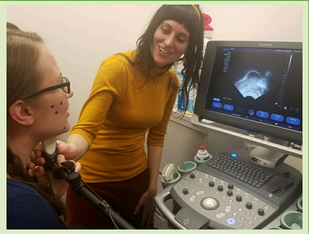
✧ Discrete Fourier Transform (DFT)
Measuring tangent angles for each tracked point along the tongue curve; transform into sine or cosine wave, yielding coefficients (C1, C2, C3, etc.) with real and imaginary components, C1 is largest.

Approach:
✧ LDA classification to distinguish complexity categories:
  ✧ Concluded that real & imaginary components of C1 were best combined predictors of complexity class.

The current study extends the experimental paradigm in Dawson et al. to preschool-aged children in order to determine which specific measure is predictive of tongue complexity in children.
✧ Metric must be sensitive to contrasts over the course of speech acquisition that may not be perceptually evident (Zharkova et al., 2016).

METHODS

✧ 8 preschool participants (5 females, 3 males) who were evaluated in an ultrasound study at Molloy College. The children ranged in age from 4;0 to 6;3 (mean = 4;10, sd = 10 months).



✧ Preliminary measurements from one participant are reported here:
  ✧ 6;1 year-old female with normal hearing and no history of speech or language impairment, although /r/ and /l/ sounds were emerging.

✧ Task: 48-word probe with randomized presentation of 3 elicitations of each word (78 measured phonemes, as some repeated).

Complexity Categories (as proposed in Dawson et al. (2016))

Low	Medium	High
(21 total):	(23 total):	(34 total):
✧ 15 /æ/ (cat, lamb, rat, yam)	✧ 5 /w/ (wake, wing)	✧ 9 /t/ (tape, tea, toe)
✧ 6 /ɪ/ (ring, wing)	✧ 3 /j/ (yam)	✧ 7 /l/ (lake, lamb)
	✧ 15 /k/ (cape, cat, coat, key)	✧ 18 /ɹ/ (rake, rat, ring, rope)

✧ Tagged sound files using .Textgrid file in Praat (Boersma & Weenick (2014), making separate tier for vowel intervals, approximant intervals, and plosives burst locations. Used Praat scripts to make stop interval.

✧ GetContours (Tiede, 2016) allows interaction between tagged intervals in.Textgrid and corresponding ultrasound images.
✧ 16 tagged data points summarized through a variety of algorithms in Matlab (MathWorks, 2000).

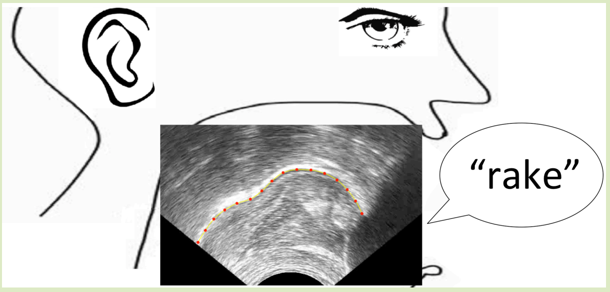


Figure 1. In GetContours, points were tagged along each lingual contour. Pictured in this image is an /ɹ/ production within the word "rake."

ANALYSES

✧ Python script (Dawson, 2016) runs MCI, Procrustes, and DFT analyses on raw coordinates extracted from Matlab.
✧ Scripts in R (RStudio Team, 2016)
  ✧ To find best individual classifier, rank-ordered boxplots each metric.
  ✧ To find best combined measure, linear discriminant analysis using the "lda" function in the MASS package (Ripley et al., 2013) in R.

RESULTS

All /l/ productions were removed from analysis based on 6/7 tokens being greater than 3SDs from the mean of at least one metric. Only 1/18 /ɹ/ productions was removed because it was greater than 3SDs from the mean of two metrics.

Individual measures
✧ Real component of C1 good at ranking medians, but there is minimal separation between the groups based on this individual metric.

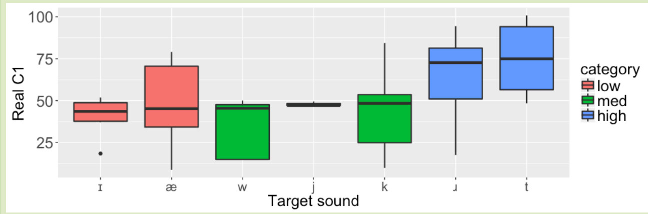


Figure 2. Rank ordering of medians of each target sound (colored by complexity category) for the real component of the first coefficient from DFT.

Combined measures
✧ LDA plots showing how individual sounds fit into proposed categories based on combined metrics.
✧ Best classifier is real component of C1 combined with Procrustes, with an error rate of 37.1%.

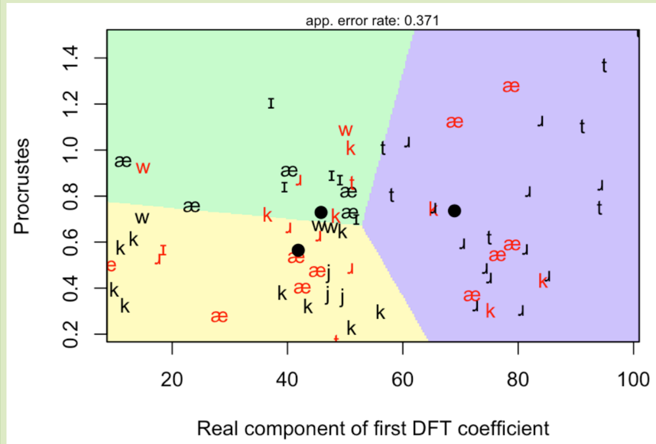


Figure 3. Linear discriminant analysis classifying target sounds into lingual complexity categories. Low, medium, and high complexity regions correspond with green, yellow, and purple, respectively. Target sounds in black were classified accurately, whereas those in red were classified inaccurately. The error rate is printed at the top.

CONCLUSIONS

✧ The real component of C1 is the best individual predictor of lingual complexity in this child.
✧ The best combination of metrics was Procrustes with either the real component of C1.
  ✧ Contrast with Dawson et al. finding for adults that real & imaginary components of C1 were best predictors.

Next steps:
✧ Measure contours of 7 remaining preschoolers to see if individual and combined data corroborate the combined metrics found to be the best classifiers in one child.
✧ Explore alternatives to the complexity categories proposed by Dawson et al.
✧ Account for incorrect productions:
  ✧ Improved classification may be found when looking only at correct productions, which may serve to predict perceptual accuracy.

The complexity metric found with full dataset will be applied to children (ages 9-15) enrolled in treatment for /ɹ/ misarticulation, which could help identify whether lingual complexity can predict treatment success.

REFERENCES

Boersma, P., & Weenink, D. (2014). Praat: doing phonetics by computer (Version 5.3.84) [Computer program]. Retrieved from http://www.fon.hum.uva.nl/praat/.
Dawson, K. M. (2016). tshape_analysis (Version 3) [Python script]. Retrieved from https://github.com/kdawson2/tshape_analysis.
Dawson, K. M., Tiede, M. K., & Whalen, D. (2016). Methods for quantifying tongue shape and complexity using ultrasound imaging. *Clinical Linguistics & Phonetics, 30*(3-5), 328-344.
Gibbon, F. E. (1999). Undifferentiated lingual gestures in children with articulation/phonological disorders. *Journal of speech, language, and hearing research: JSLHR, 42*(2), 382-397.
Goodall, C. (1991). Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society. Series B (Methodological)*, 285-339.
MathWorks Inc. (2000). MATLAB (Version 6.1) [Computer program]. Natick, MA. Retrieved from https://www.mathworks.com/products/matlab/.
Ripley, B., Venables, B., Bates, D. M., Hornik, K., Gebhardt, A., Firth, D., & Ripley, M. B. (2013). Package 'MASS' (Version 7.3-45): CRAN Repository. Retrieved from https://cran.r-project.org/web/packages/MASS/MASS.pdf.
RStudio Team. (2016). RStudio: integrated development for R (Version 0.99.879) [Computer program]. Boston, MA: RStudio, Inc. Retrieved from https://www.rstudio.com/products/rstudio/.
Stolar, S., & Gick, B. (2013). An index for quantifying tongue curvature. *Canadian Acoustics, 41*(1).
Tiede, M. (2016). GetContours (Version 1.0) [Matlab script]. Retrieved from https://github.com/mktiede/GetContours.
Zharkova, N., Gibbon, F. E., & Lee, A. (2016). Using ultrasound tongue imaging to identify covert contrasts in children's speech. *Clinical Linguistics & Phonetics*, 1-14.

This research was supported by an NIH NIDCD R01DC013668 grant.