IMPROVING CHILDREN'S ULTRASOUND-BASED TONGUE COMPLEXITY MEASURES BY QUANTIFYING MIDSAGITTAL ALIGNMENT

Heather Kabakoff, Tara McAllister, D.H. Whalen, Mark Tiede

NYU

CITY UNIVERSITY OF NEW YORK THE GRADUATE CENTER

Haskins Laboratories

Yale

Introduction

- Children acquire speech sounds following established developmental timelines. (Smit et al., 1990, Crowe & McLeod, 2020)
- Developmental error patterns reflect differences in:
  - Phonological knowledge
  - Skilled motor control of speech structures

Phoneme	Recommended age of acquisition (years:months)	
	Females	Males
/m/	3:0	3:0
/n/	3:6	3:0
/-ŋ/	7:0–9:0	7:0–9:0
/h-/	3:0	3:0
/w-/	3:0	3:0
/j-/	4:0	5:0
/p/	3:0	3:0
/b/	3:0	3:0
/t/	4:0	3:6
/d/	3:0	3:6
/k/	3:6	3:6
/g/	3:6	4:0
/f/ /f-/	3:6	3:6
/-f/	5:6	5:6
/v/	5:6	5:6
/θ/	6:0	8:0
/ð-/	4:6	7:0
/s/	7:0–9:0	7:0–9:0
/z/	7:0–9:0	7:0–9:0
/ʃ/	6:0	7:0
/tʃ/	6:0	7:0
/dʒ/	6:0	7:0
/l/ /l-/	5:0	6:0
/-l/	6:0	7:0
/r/ /r-/	8:0	8:0
/-ɚ/	8:0	8:0

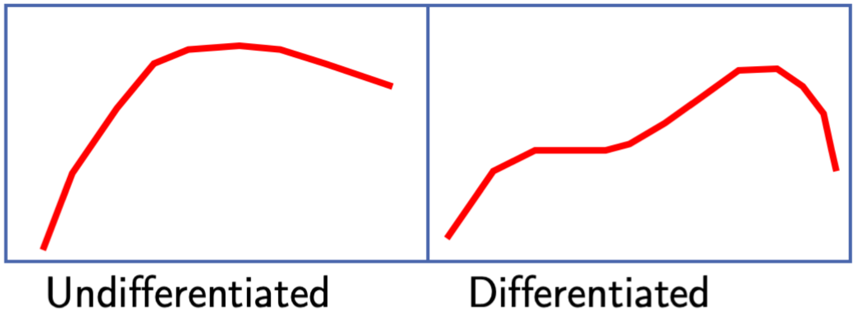
Timeline of speech sound acquisition

Developmental trajectory:
(Crowe & McLeod, 2020)

- **early 13** (2;0-3;11)
  - plosives, nasals, & glides
- **middle 7** (4;0-4;11)
  - /v,s,z,ʃ,ʧ,ʤ,l/
- **late 4** (5;0-6;11)
  - /ɹ,θ,ð,ʒ/

Lingual differentiation

- Reduced control of **anterior** versus **posterior** lingual regions may play a role in non-adult-like speech patterns:
  - **Typical development** (Fletcher, 1989)
    - Lingual differentiation increases with age for complex targets.
  - **Speech sound disorders** (Gibbon, 1999)
    - Lingual differentiation is reduced in children with atypical speech.



- Suggests that lingual differentiation reflects an **individual's capacity for skilled motor control** of the tongue.

Quantification of lingual differentiation could elucidate the relative contribution of motor control versus phonological factors in developmental and disordered speech.

Timeline of speech sound acquisition

Developmental trajectory:
(Crowe & McLeod, 2020)

- **early 13** (2;0-3;11)
  - plosives, nasals, & glides
- **middle 7** (4;0-4;11)
  - /v,s,z,ʃ,ʧ,ʤ,l/
- **late 4** (5;0-6;11)
  - /ɹ,θ,ð,ʒ/

Articulatory trajectory:
(e.g., Studdert-Kennedy & Goldstein, 2003)

- **Stage 1**
  - voiceless stops, nasals, glides, /h/
- **Stage 2**
  - voicing contrasts
- **Stage 3**
  - fricatives, affricates
- **Stage 4**
  - multiple lingual constrictions → /l, ɹ/

- Late-developing sounds tend to require complex tongue shapes.

Simplification strategies

Rhotics:

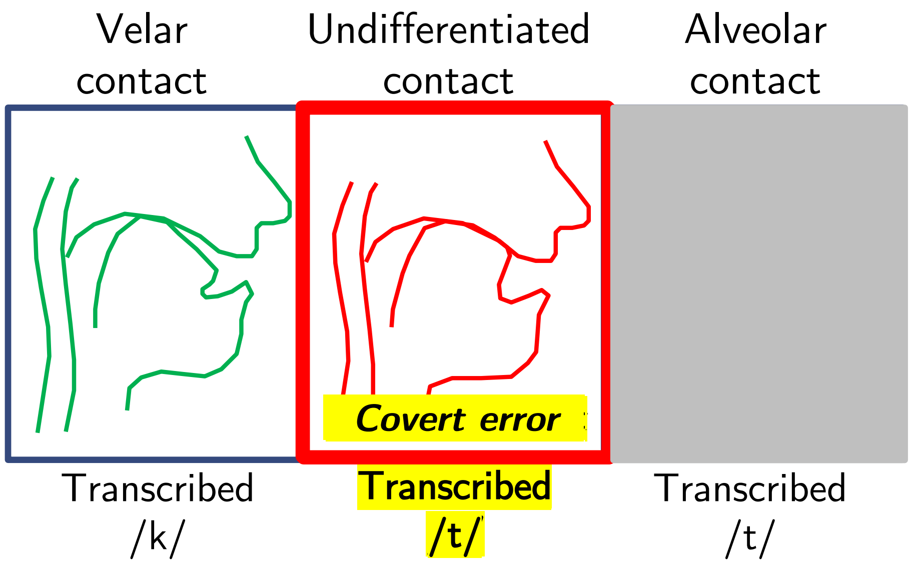
- Adults produce the English /ɹ/ with two lingual constrictions.
- Children omit a constriction or combine constrictions. (Gick et al., 2007)
- For children with speech sound disorder: (Klein et al., 2013)
  - Increased lingual differentiation associated with:
    - ↑ degree of perceived accuracy
    - ↑ over the course of treatment

Laterals:

- Adults produce light and dark /l/ with two lingual constrictions. (Lin & Demuth, 2015)
- Number of constrictions increases from age 3 to 7.
- Children produce /l/ with a single tongue constriction.
  - Onset /l/ perceived as accurate
  - Coda /ɫ/ perceived as inaccurate

Relationship with perceived accuracy

- Degree of lingual differentiation does not necessarily correspond with perceptually rated accuracy. (Gibbon, 1999)
  - "**Covert contrast**":
    - distinct gestures when output is perceptually neutralized. (Gibbon et al., 1993, McAllister Byun et al., 2016)
  - "**Covert error**":
    - atypical lingual configurations that are not detected perceptually. (Cleland et al., 2017; Klein et al., 2013)



- Lingual measures may reveal information about motor maturation not obtainable from perceptual ratings!

Clinical utility

- Measuring degree of lingual differentiation could be key to determining which treatments to recommend for children with speech sound disorder (SSD):
  - **Traditional motor-based treatment:**
    - Repetitive production practice in a context hierarchy (Van Riper, 1978) with incorporation of the principles of motor learning (Maas et al., 2008).
  - **Phonological-based treatment:**
    - Reorganize the phonological system (Hodson & Paden, 1983, McReynolds & Bennett, 1972, Barlow & Gierut, 2002)
- Different children *should* be matched with the best intervention approach based on strengths/weaknesses in motoric, phonological, and perceptual domains.
- **Need clinically viable way to quantify motor ability!**

Measuring tongue complexity via ultrasound

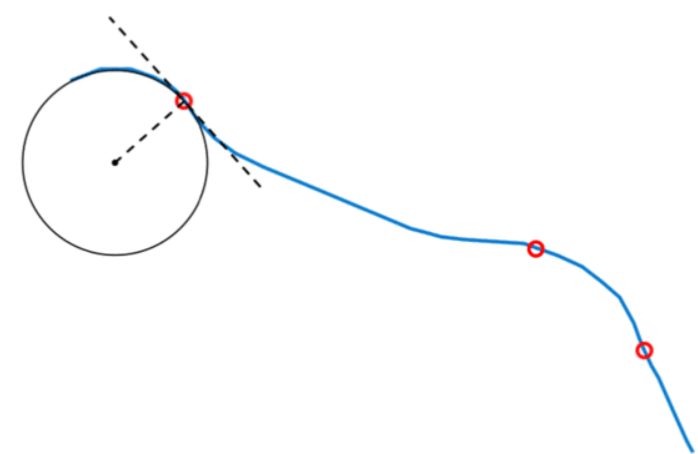
- Ultrasound is increasingly affordable, minimally invasive, and ready to use with small children (unlike EPG).
- Covert contrasts and covert errors can be revealed using ultrasound. (Cleland et al., 2017; McAllister Byun et al., 2016)
- Continuous tongue contour → "**tongue complexity**" (Dawson et al., 2016)
  - Can provide insight into a child's stage of motor development.

Metrics of tongue complexity

Modified Curvature Index

(MCI; Dawson, Tiede, & Whalen, 2016)

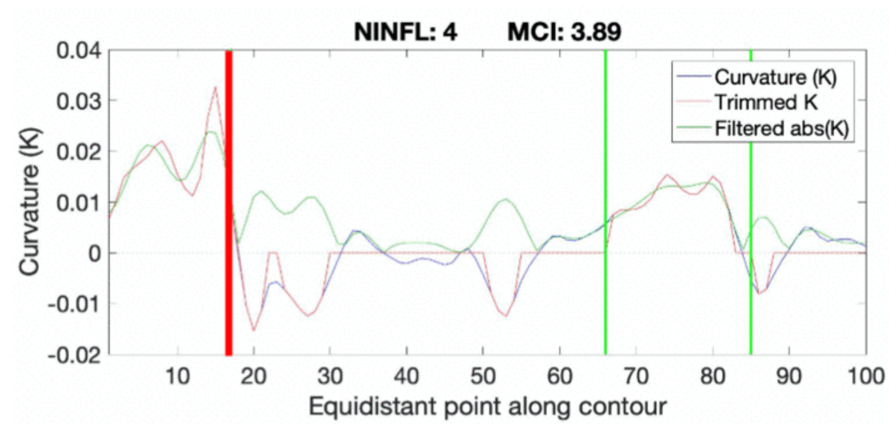
- Integral of absolute curvature (reciprocal of the tangent circle) at each of the equidistant points
- For adults producing various phonemes, higher MCI values were found in phonemes with multiple constrictions (including /ɹ/ and /l/) than those with a single constriction (e.g., /æ,ɪ/).



Number of INFLection points

(NINFL; Preston, McCabe, Tiede, & Whalen, 2019)

- Number of thresholded curvature sign changes along the arc of a given tongue contour (plus one for initial bias)
- For children producing /ɹ/, Preston et al. (2019) found NINFL patterns:
  - TD > SSD
  - Accurate > Inaccurate
  - Post-treatment > Pre-treatment



Questions and Hypotheses

1) **In typically developing (TD) children, does tongue complexity increase with age, and are there phoneme-specific patterns?**

   Early-developing: No relationship between age and tongue complexity

   Late-developing: Positive correlation between age and tongue complexity

2) **Do TD children use more complex tongue shapes than peers with speech sound disorder (SSD)?**

   This finding would suggest that ultrasound may have diagnostic value.

3) **Are perceptually incorrect /l/ and /ɹ/ sounds produced with less complex tongue shapes than correct productions?**

   Would extend finding from Preston et al. (2019) from school-aged children with speech sound disorder to younger children producing errors in the presence or absence of SSD.

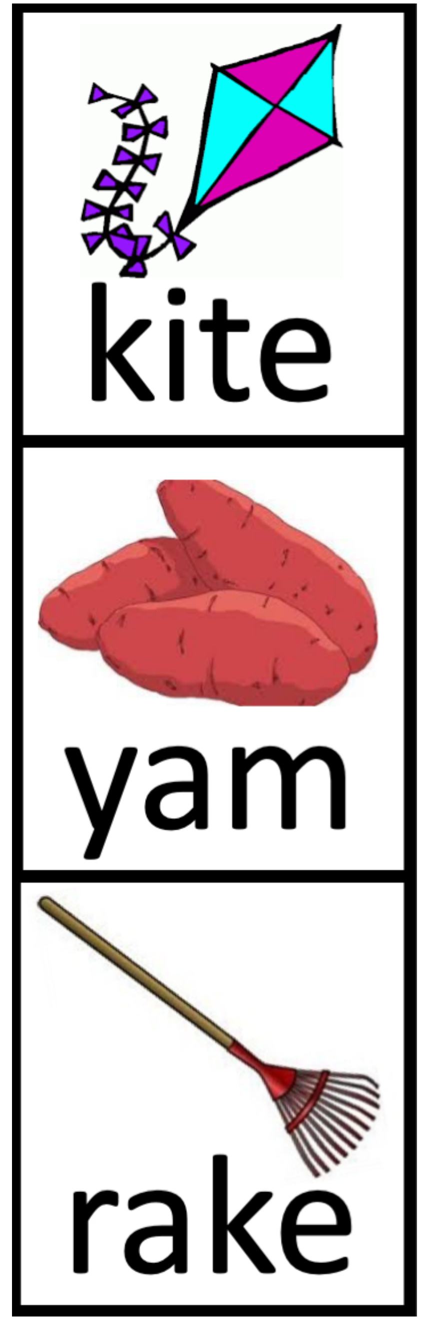
Methods: Participants

- 25 children identified as TD (17) or as having SSD (8) based on performance on HAPP-3. (Hodson, 2004)
- Evaluated at Haskins Labs (11), Molloy College (7), or Syracuse University (7)

	Gender	Age Range	% correct /l/	% correct /ɹ/
SSD	M (n = 6)	4;2-5;11	0, 0, 25, 63, 83, 89	0, 0, 0, 50, 73, 89
SSD	F (n = 2)	4;0; 5;0	0, 0	0, 77
TD	M (n = 6)	4;2-6;3	0, 67, 67, 75, 75, 100	0, 0, 0, 0, 67, 100
TD	F (n = 11)	5;11	60, 60, 100, 100, 100, 100, 100, 100, 100, 100, 100	0, 0, 0, 40, 50, 56, 91, 100, 100, 100, 100

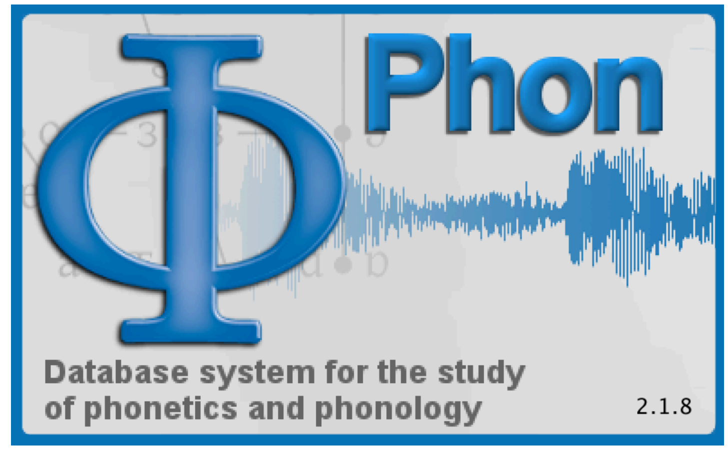
Methods: Task

- Each child produced the consonants /j,w,k,t,l,r/ in onset position with /æ,ɪ/ as the nucleus of CV/CVC words:
  - *yam, wake, wing, cape, cat, coat, key, tape, tea, toe, lake, lamb, rake, rat, ring, rope*
  - Elicited three times each in random order.
  - Children were initially familiarized with target words and their corresponding pictures.



Methods: Ratings of perceptual accuracy

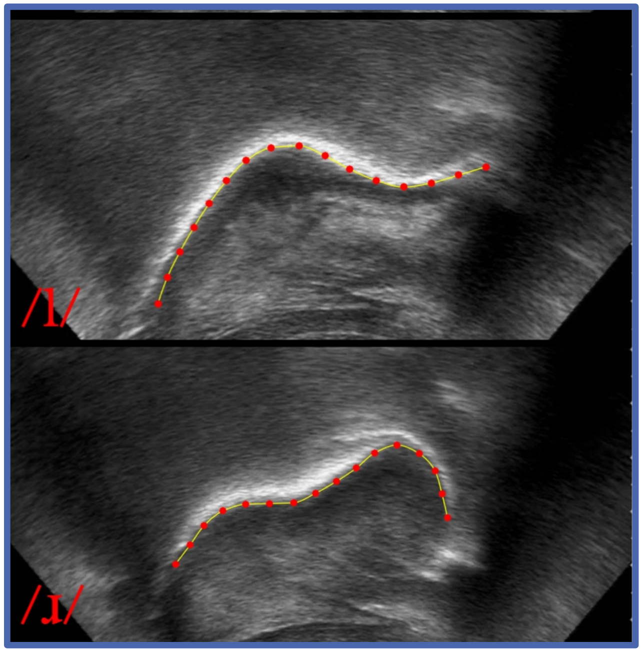
- Two trained and blinded students narrowly transcribed each token using Phon speech analysis software. (Hedlund & Rose, 2016)
- A third trained student transcriber resolved discrepancies.



- Transcriptions converted into ratings of perceptual accuracy:
  - Distortions were rated as incorrect.

Methods: Ultrasound data collection

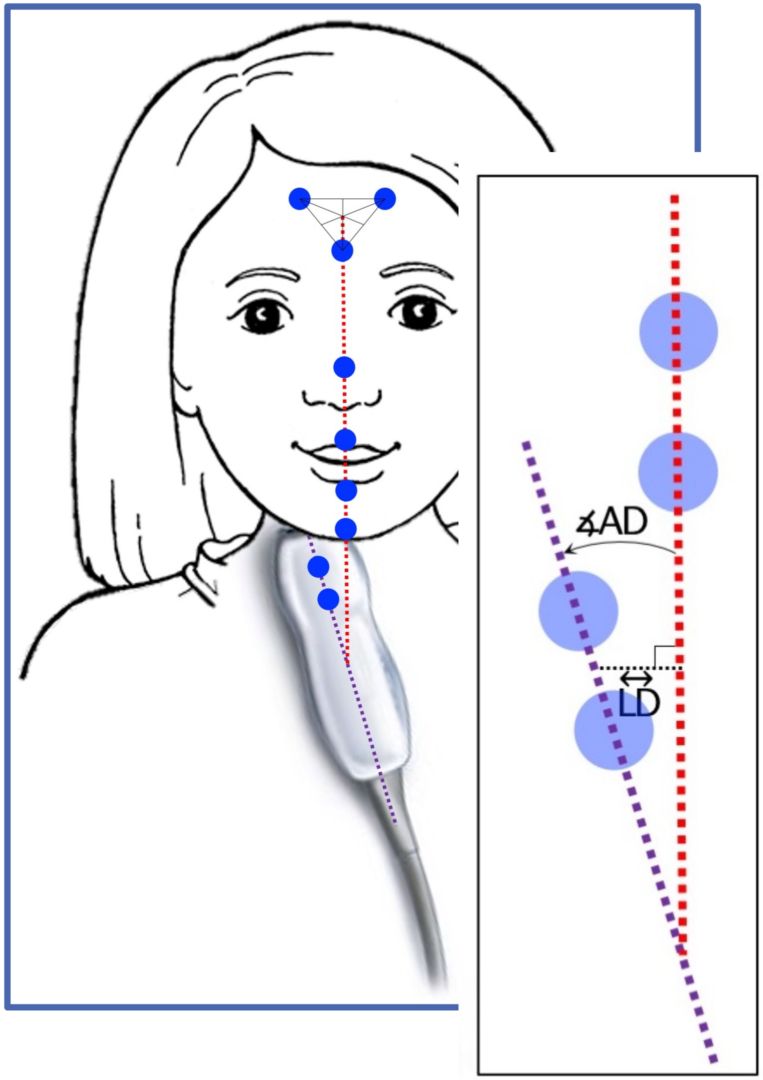
	Haskins Labs	Molloy College	Syracuse University
Machine	Siemens Acuson X300	Siemens Acuson X300	Telemed Echoblaster 128
Transducer	C6-2 wideband curved array	C8-5 wideband curved array	PV 6.5 wideband curved array
Frequency Range	1.8–6 MHz	3.1–8.8 MHz	5–8 MHz
Field of View	90 degree	109 degree	156 degree
Scanning Settings	36-37 fps 80 mm	43-49 fps 60-70 mm	11-25 fps 110 mm



- Probe placed in microphone stand and monitored throughout production.
- Ultrasound recorded through Debut/AverMedia video capture card.
- Representative tongue shapes nearest the midpoint of the acoustically-defined target intervals were traced using GetContours in MATLAB.
  (Tiede, 2016; MathWorks Inc., 2000)

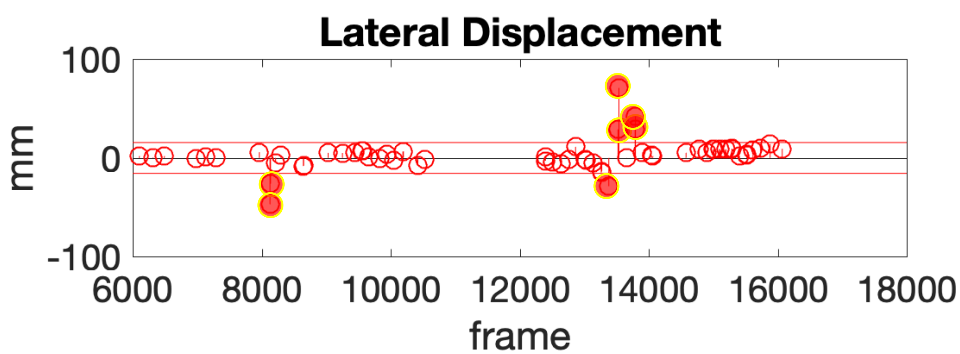
Methods: Quantifying probe alignment

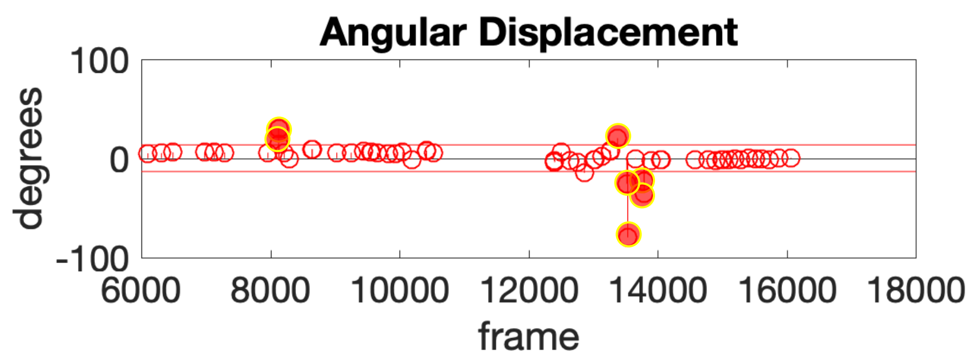
- We video-recorded position of blue dots placed along the vertical midline of the child's face and the probe.
- **Lateral displacement (LD):**
  - Perpendicular distance from head line to centroid of probe dots
- **Angular displacement (AD):**
  - Counter-clockwise angle from the head to the probe line
- Ultrasound and frontal video were temporally aligned using cross-correlation of their audio.



Methods: Discarding misaligned frames

- Frames with more than one standard deviation of lateral displacement (15.5 mm) or one standard deviation of angular displacement (13.3 degrees) across all files were discarded.
- Across all 25 children:
  - 15.2% (248/1634) of frames were flagged due to lateral displacement.
  - 15.6% (255/1634) of frames were flagged due to angular displacement.
  - A total of 1246 tokens remained for final analysis.





Analyses

MCI

- Token-level analysis performed with Python script 'tshape_analysis'
(Dawson, 2016; Python Software Foundation, 2016)

- Models analyzing MCI used linear mixed-effects regression with the 'lme4' package in R.
(Bates, et al., 2015; R Core Team, 2019)

NINFL

- Token-level analysis performed with MATLAB script 'ComputeCurvature'.
(Tiede, 2017; MathWorks Inc., 2000)

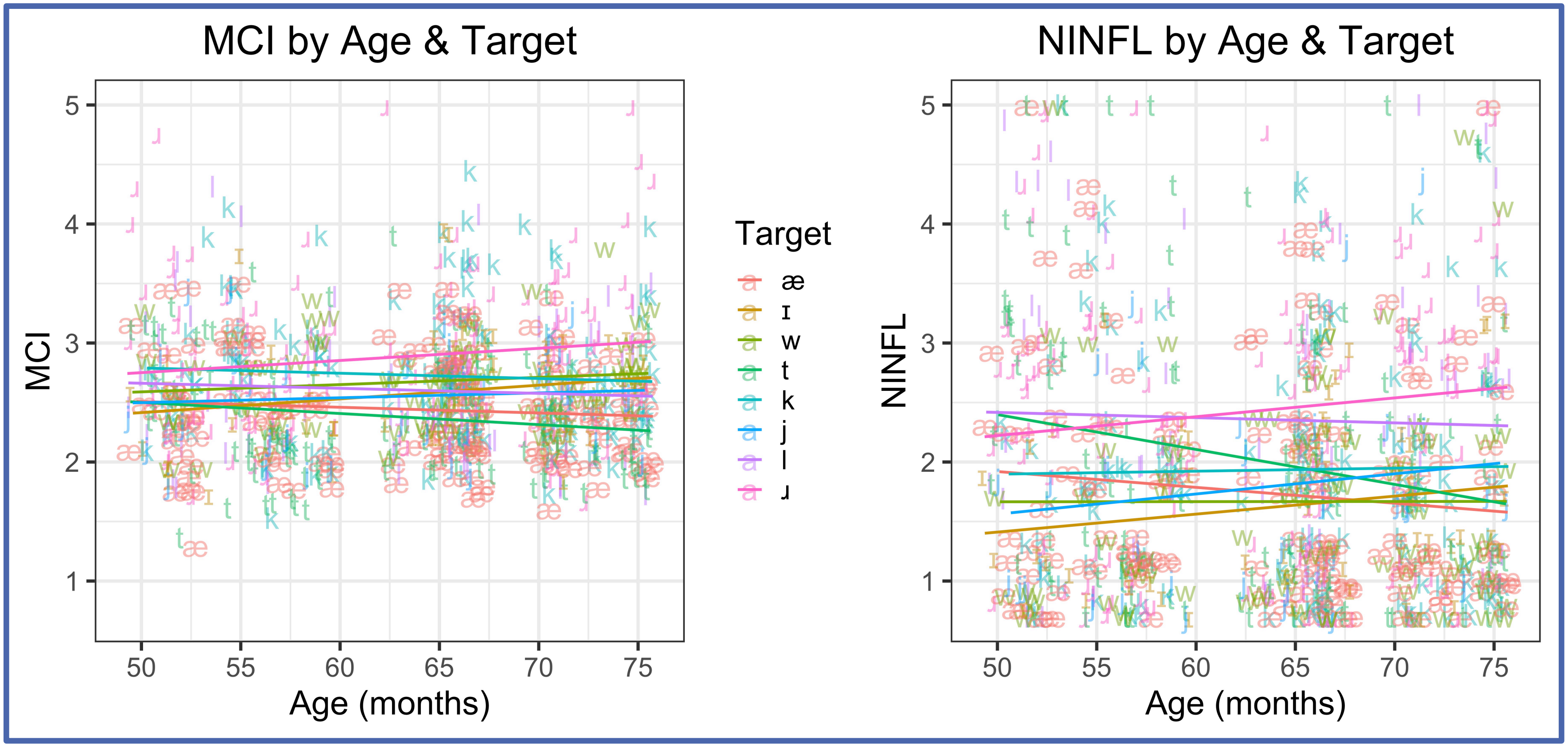
- Models analyzing NINFL used ordinal mixed effects regression with the 'clmm' package in R.
(Christensen, 2015)

- All models included random intercepts for word and child and by-child random slopes on target.

Results: Question 1

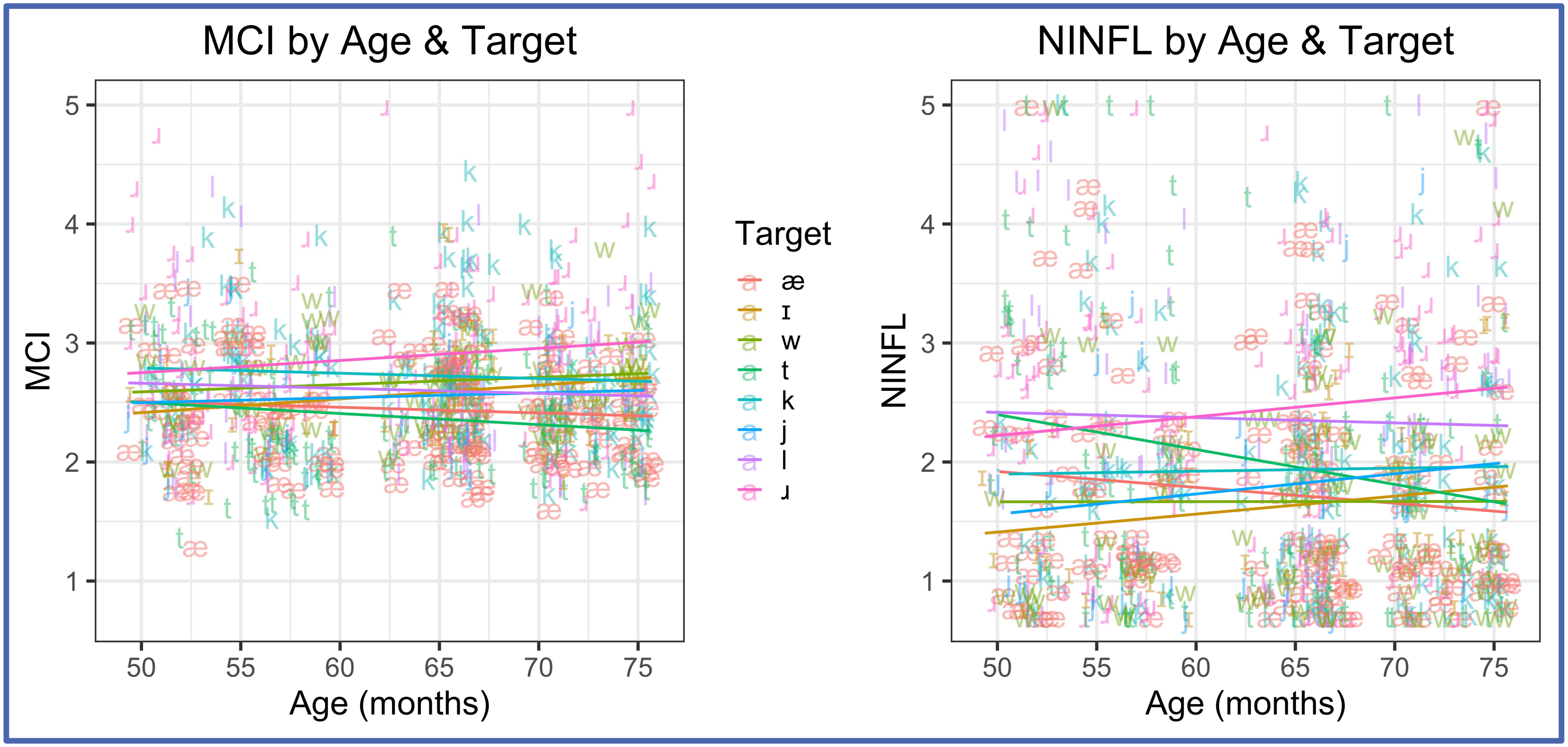
1) Is there an association between age and tongue complexity in TD children?

- MCI/NINFL ~ Age * Target



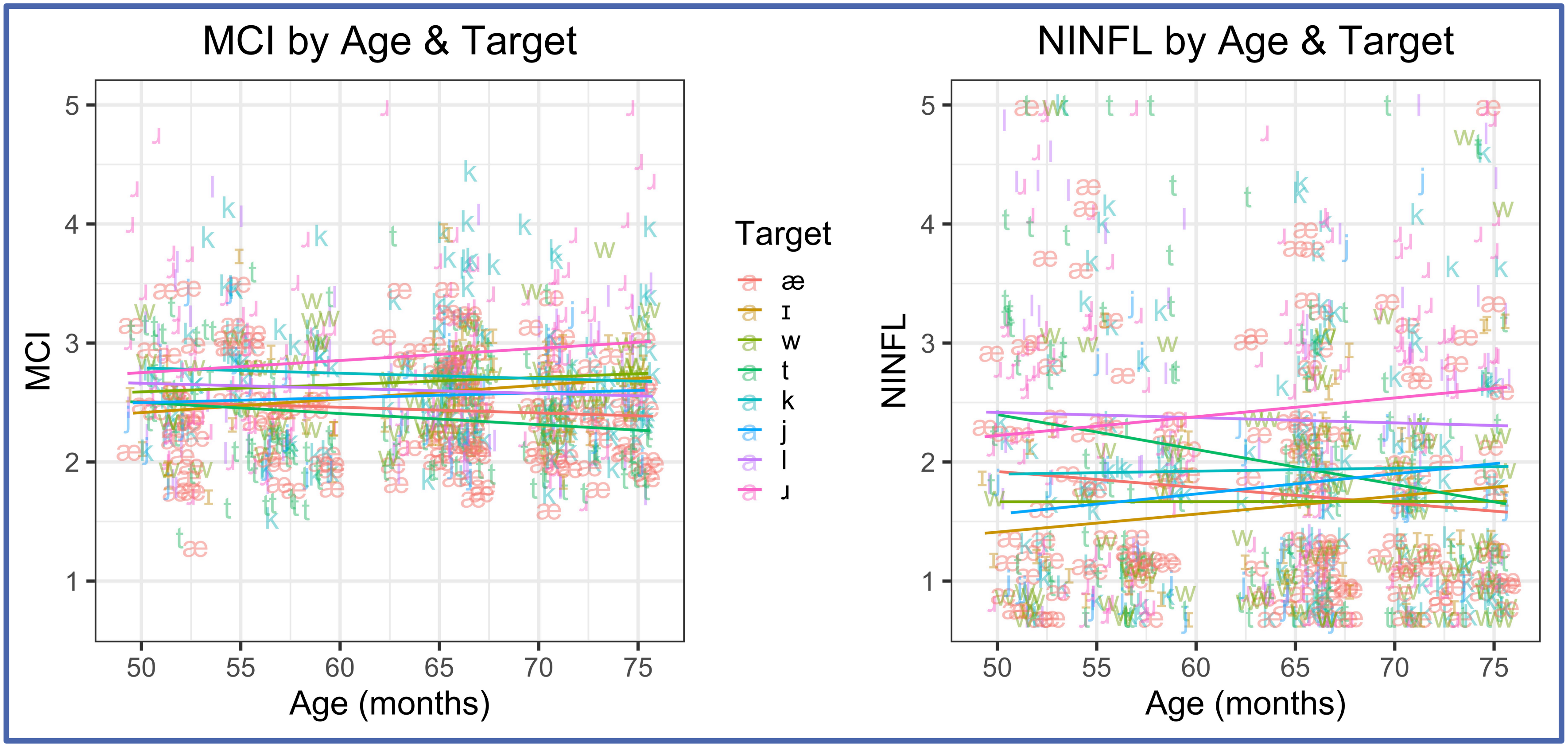
Results: Question 1

- No observed differences for children of different ages.
- No significant interactions between age and target.



Results: Question 1

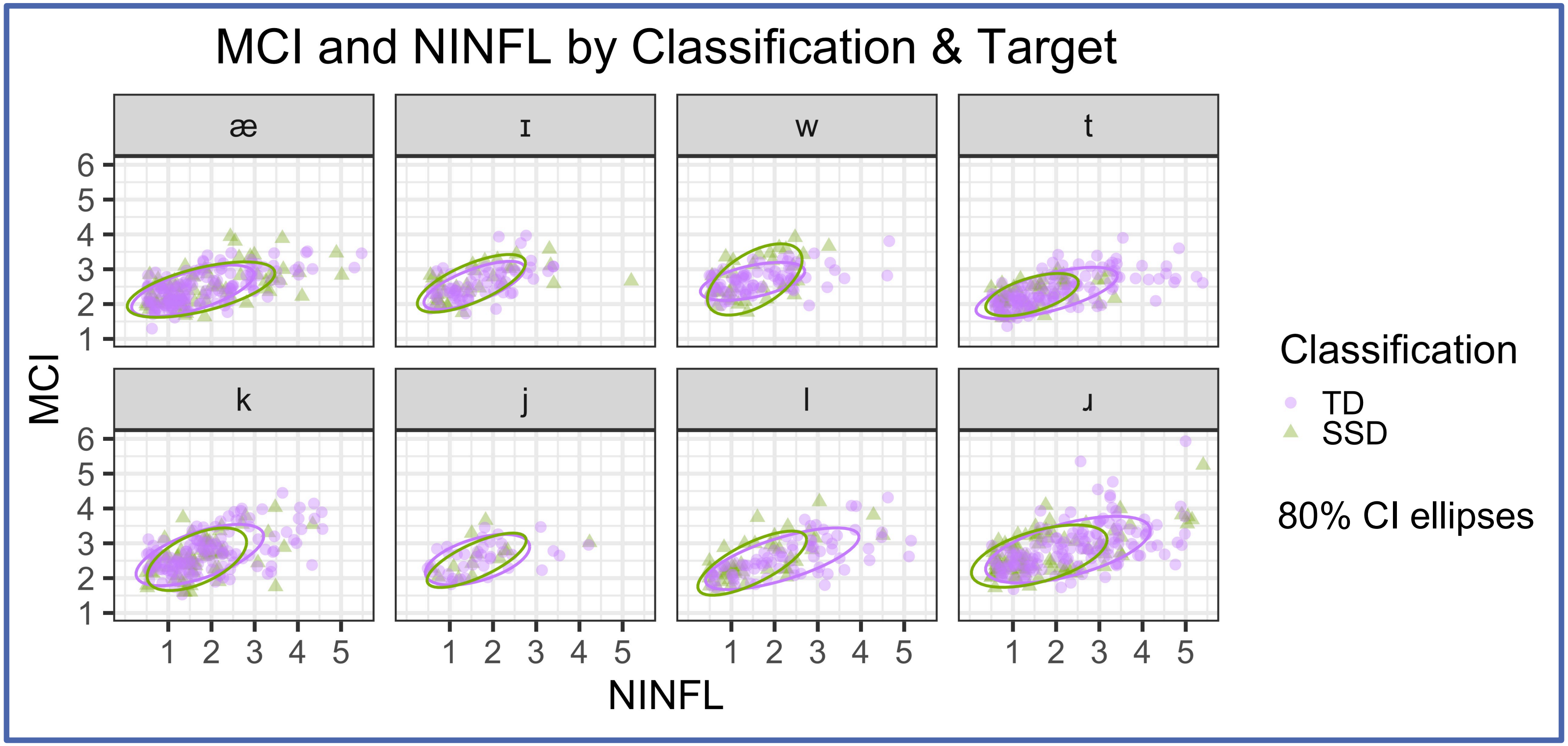
- Significant differences in tongue complexity across targets:
  - **MCI:** means for /w/, /k/, and /ɹ/ were greater than /æ/.
  - **NINFL:** means for /l/ and /ɹ/ were greater than/æ/.



Results: Question 2

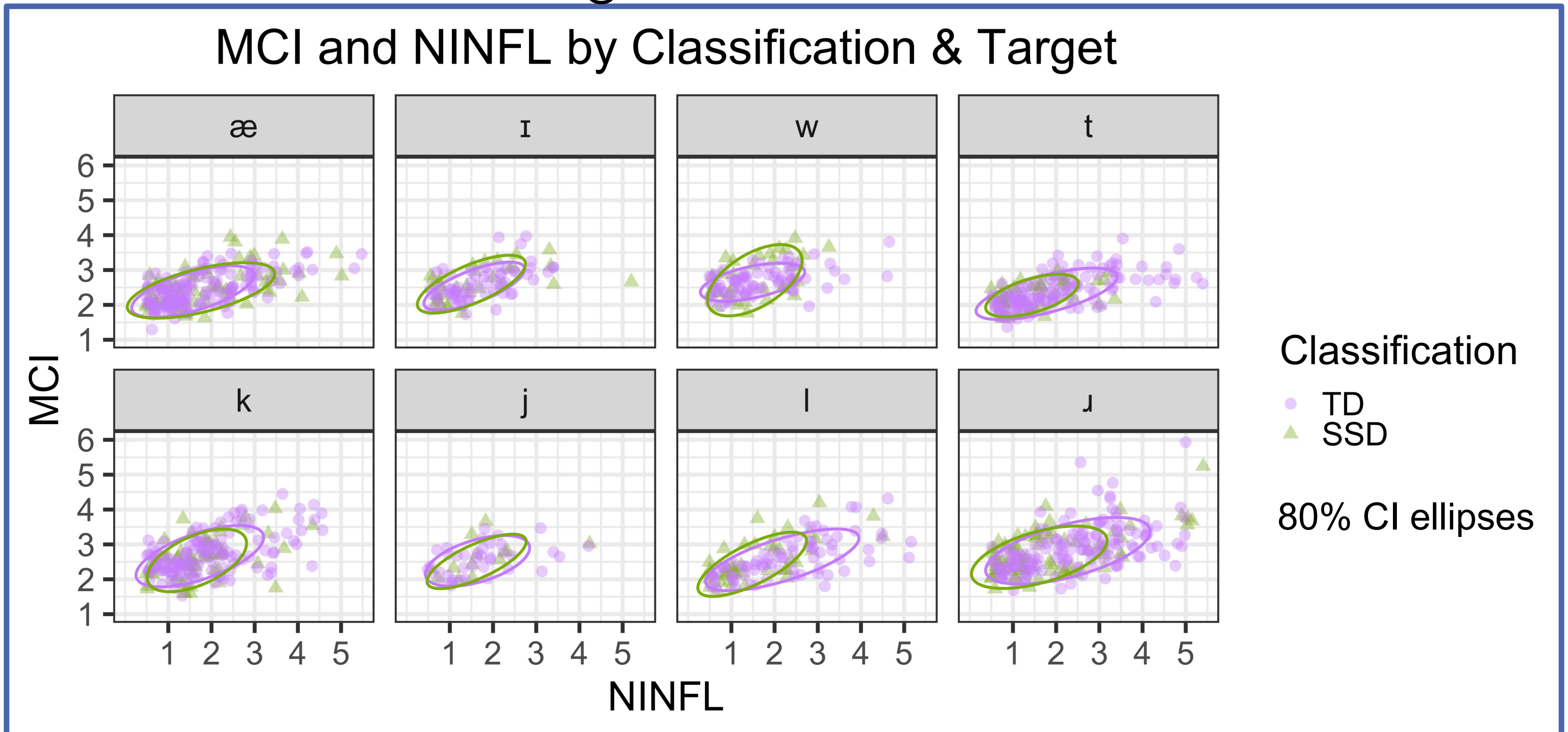
2) Do TD and SSD productions differ in tongue complexity?

- MCI/NINFL ~ Classification * Target



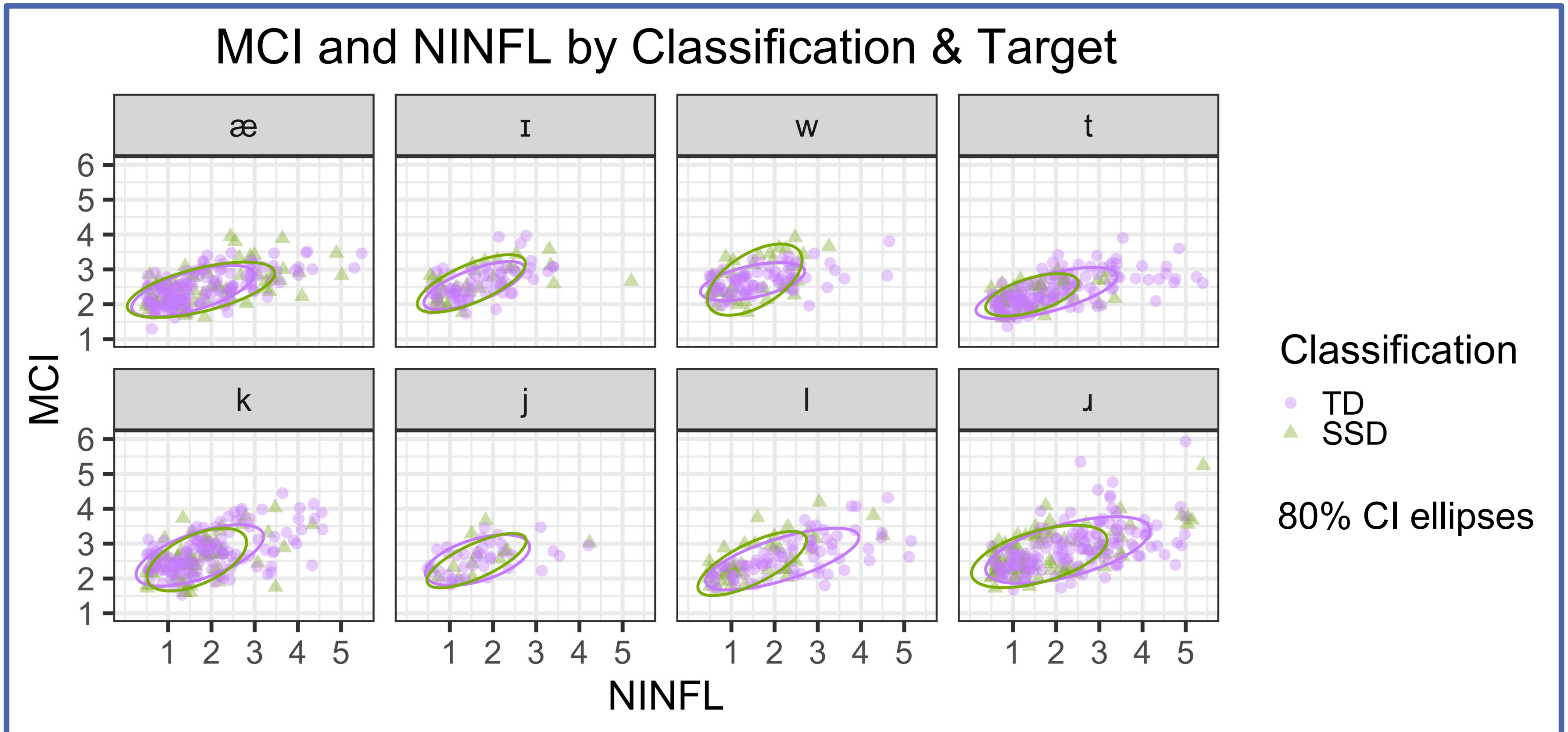
Results: Question 2

- No differences between children classified as TD versus SSD.
- The addition of children with SSD did not change the observed effects of target, as seen in Question 1.



Results: Question 2

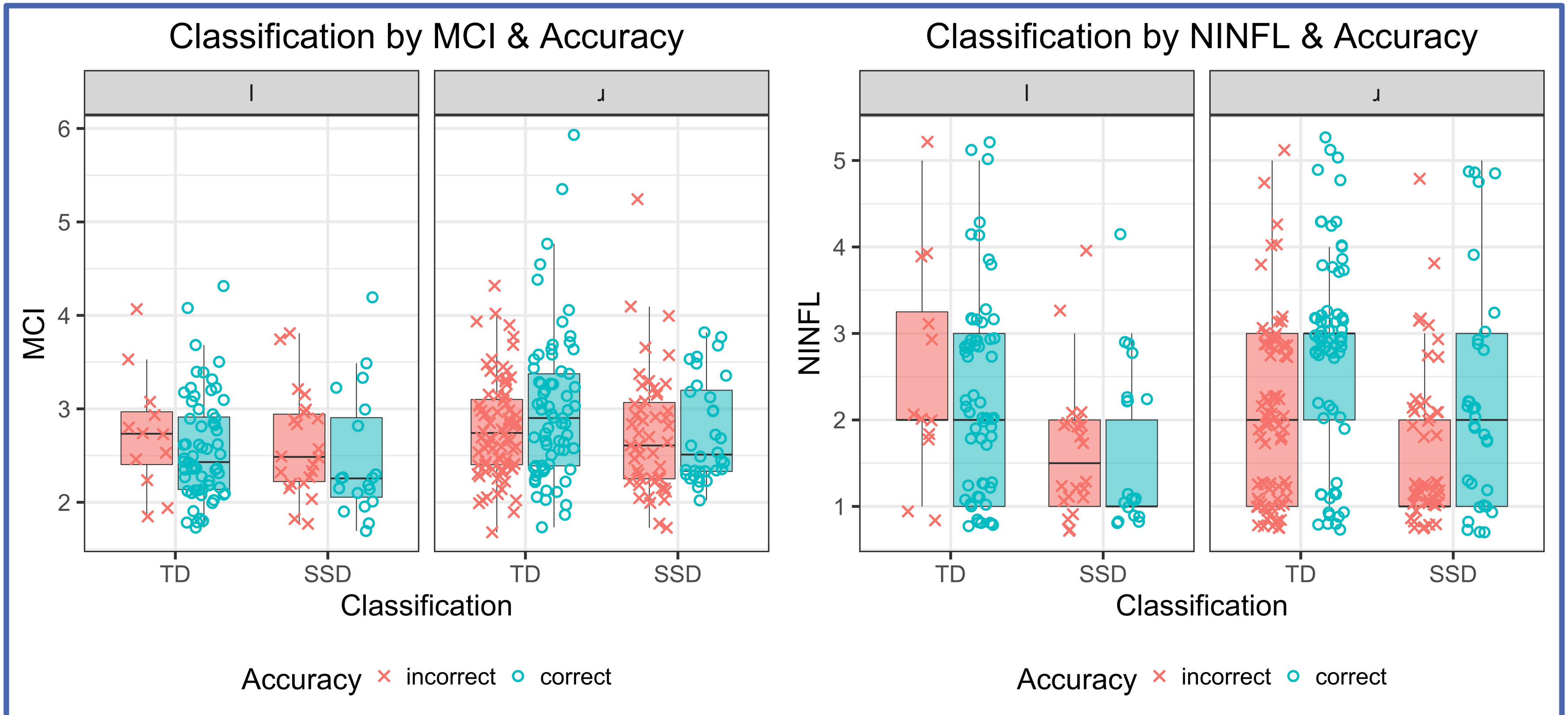
- Significant interaction between classification and target:
  - **NINFL**: greater means for /t/, /l/, and /ɹ/ in TD than SSD children
  - *No differences observed for MCI*



Results: Question 3

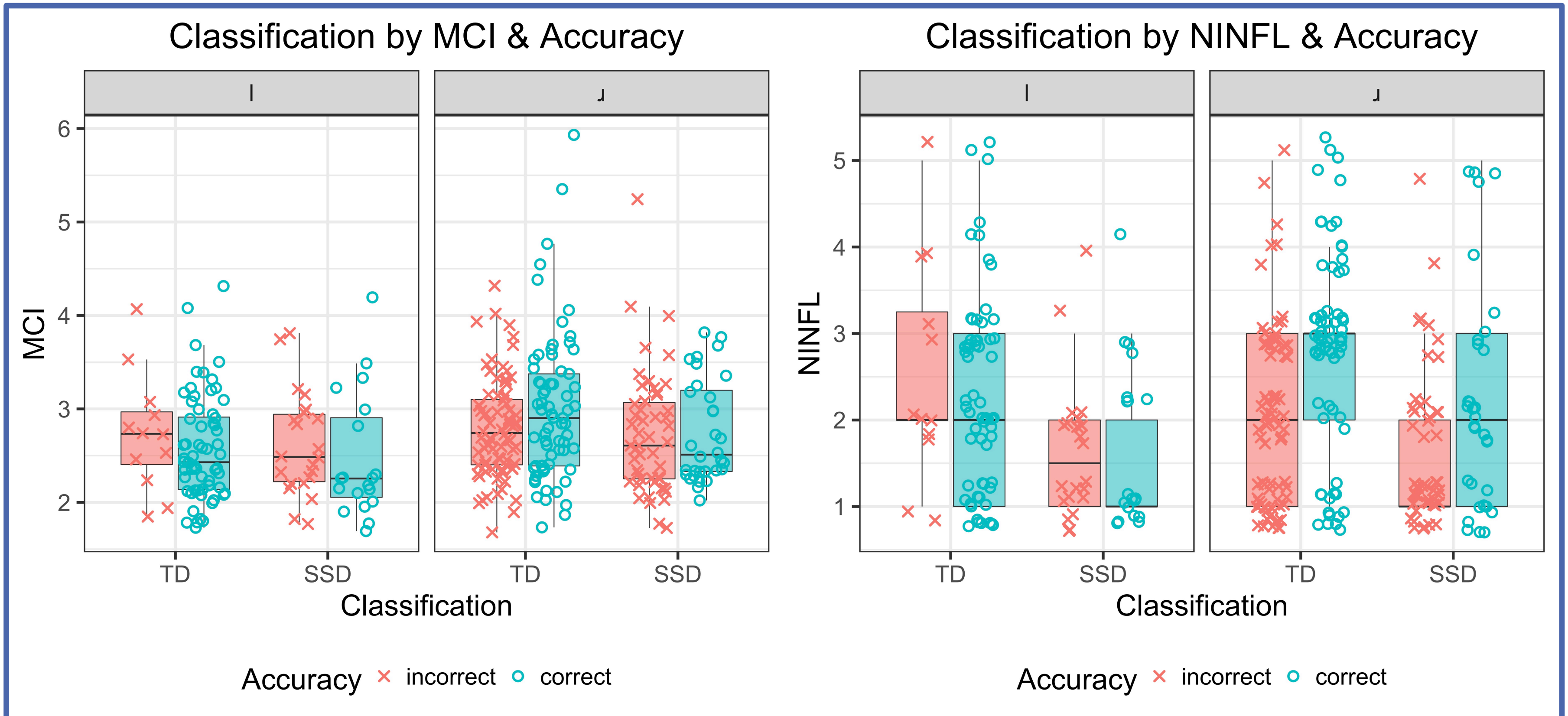
3) Is there a difference in tongue complexity between /l/ and /ɹ/ productions that differ in perceptually rated accuracy, pooled across TD and SSD groups?

- MCI/NINFL ~ Accuracy * Classification * Target



Results: Question 3

- Significant interaction between accuracy and target:
  - NINFL: correct /ɹ/ tokens > incorrect /ɹ/ tokens.
  - *Pattern not observed for /l/, or for MCI.*



Discussion: MCI versus NINFL

	MCI	NINFL
Q1 (AGE) & Q2 (CLASSIFICATION)	/w/, /k/, and /ɹ/ greater than /æ/	/l/ and /ɹ/ greater than /æ/
Q2 (CLASSIFICATION)	No differences	/t/, /l/, and /ɹ/ greater in TD than SSD
Q3 (ACCURACY)	No differences	correct /ɹ/ greater than incorrect /ɹ/

- Moderate correlation between MCI and NINFL: $r(1208) = 0.58$, $p < .0001$
- The discrepancies highlight their computational differences.
  - **MCI**: driven by curvature, such that MCI can be high with only one point of inflection as long as the local curvature is high.
  - **NINFL**: tongue contours only receive high NINFL values if regions of constriction are separated by anti-constriction.
- A combined metric may be optimal.
- NINFL aligns better with theory-driven expectations.

Discussion: Question 1

- Age was not significantly associated with tongue complexity in our TD child sample.
  - Differences may not be detectable during small age range.
  - Future analyses should compare the current sample with a group of older children producing same targets in the same task.
  - A follow-up longitudinal comparison could help elucidate the trajectory along which childhood tongue complexity becomes adult-like.

Discussion: Question 2

- TD children had more complex tongue shapes for /t/, /l/, and /ɹ/ than children with SSD.
  - **Expected**: /l/ and /ɹ/ are articulatorily complex, phonologically marked, late-developing, and commonly misarticulated.
  - **Unexpected**: /t/ is relatively early-developing and was generally transcribed as correct in the current sample.
- This pattern for /t/ can be considered **covert error** in which children with SSD produced perceptually accurate /t/ using different articulatory strategies than TD children.
  - Consistent with broad patterns found with EPG (Gibbon, 1999) and ultrasound (Cleland et al., 2017)

Discussion: Covert error clinical utility

- Covert errors suggest globally reduced tongue complexity, which may indicate presence of SSD. (Cleland et al., 2017; Gibbon, 1999)

SSD

Globally-reduced tongue complexity

- Motor-based treatment, such as ultrasound biofeedback (Sugden et al., 2019)

Typical degree of tongue complexity

- Phonological approach to intervention

- Future work should include more participants and track tongue complexity over the course of various treatments.

Covert error highlights how tongue complexity reveals information distinct from perceptual ratings of accuracy.

Discussion: Question 3

- **Expected:** For with/without SSD, correct /ɹ/ tokens had higher tongue complexity than incorrect tokens, as expected.
- **Unexpected**: No relationship between tongue complexity and /l/ accuracy, contrary to expectation based on reported developmental patterns (Gick et al., 2007; Lin and Demuth, 2015)

- Many exceptions such as accurate /ɹ,l/ productions with low tongue complexity and vice versa in both groups.

Such incongruencies highlight how tongue complexity reveals information distinct from perceptual ratings of accuracy.

Discussion: Limitations

- **Single midsagittal section provides limited insight**:
  - Ignores lateral bracing and asymmetries (Gick et al., 2017)
  - Measures based on 3D ultrasound (Lulich & Pearson, 2019) plus consideration of hard structures are needed to represent tongue complexity fully.
- **Delineation of "bunched" versus "retroflex" shapes could explain some variation in tongue complexity.** (Delattre & Freeman, 1968)
  - We avoid classifying tongue shapes because many perceptually accurate tokens do not match the classical categories. (Preston et al., 2019)
  - Future work may wish to address this question with a subset of tongue shapes that conform to the classical retroflex and bunched categories.

Conclusion

Overall, the results of the present study suggest that differences in tongue complexity can be found across multiple developmental dimensions, supporting the potential of these measures as a correlate of motor control.

- This research highlights the **value of ultrasound imaging as an increasingly affordable and child-friendly tool**.
  - Ongoing technological advances can be expected to expand the accessibility of the clinically valuable measures examined here.

References 1

Barlow, J. A., & Gierut, J. A. (2002). Minimal pair approaches to phonological remediation. *Seminars in Speech and Language, 23*(01), 57-68.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using 'lme4'. *Journal of Statistical Software, 67*(1), 1-48.

Christensen, R. H. B. (2015). 'ordinal' package in R: Regression Models for Ordinal Data via Cumulative Link (Mixed) Models (Version 2015.1-21). Retrieved from https://cran.r-project.org/web/packages/ordinal/index.html

Cleland, J., Scobbie, J. M., Heyde, C., Roxburgh, Z., & Wrench, A. A. (2017). Covert contrast and covert errors in persistent velar fronting. *Clinical Linguistics & Phonetics, 31*(1), 35-55.

Crowe, K., & McLeod, S. (2020). Children's English consonant acquisition in the United States: A review. *American Journal of Speech-Language Pathology*, 1-15.

Dawson, K. M. (2016). tshape_analysis (Version 3) [Python script]. Retrieved from https://github.com/kdawson2/tshape_analysis

Dawson, K. M., Tiede, M., & Whalen, D. H. (2016). Methods for quantifying tongue shape and complexity using ultrasound imaging. *Clinical Linguistics & Phonetics, 30*(3-5), 328-344.

Delattre, P., & Freeman, D. C. (1968). A dialect study of American r's by x-ray motion picture. *Linguistics, 6*(44), 29-68.

Fletcher, S. G. (1989). Palatometric specification of stop, affricate, and sibilant sounds. *Journal of Speech, Language, and Hearing Research, 32*(4), 736-748.

Gibbon, F. E. (1999). Undifferentiated lingual gestures in children with articulation/phonological disorders. *Journal of Speech, Language, and Hearing Research, 42*(2), 382-397.

Gick, B., Allen, B., Roewer-Després, F., & Stavness, I. (2017). Speaking tongues are actively braced. *Journal of Speech, Language, and Hearing Research, 60*(3), 494-506.

Gick, B., Bacsfalvi, P., Bernhardt, B. M., Oh, S., Stolar, S., & Wilson, I. (2007). A motor differentiation model for liquid substitutions in children's speech. *Proceedings of Meetings on Acoustics, 1*, 060003.

Hedlund, G., & Rose, Y. (2016). Phon (Version 3.0.0b2) [Computer Software]. Retrieved from https://phon.ca

Hodson, B. W. (2004). *Hodson Assessment of Phonological Patterns, 3rd ed.* Austin, TX: Pro-Ed.

Hodson, B. W., & Paden, E. P. (1983). *Targeting intelligible speech: A phonological approach to remediation*. San Diego, CA: College-Hill Press.

Klein, H. B., McAllister Byun, T., Davidson, L., & Grigos, M. I. (2013). A multidimensional investigation of children's /r/ productions: perceptual, ultrasound, and acoustic measures. *American Journal of Speech-Language Pathology, 22*(3), 540-553.

References 2

Lin, S., & Demuth, K. (2015). Children's acquisition of English onset and coda /l/: Articulatory evidence. *Journal of Speech, Language, and Hearing Research, 58*(1), 13-27.

Lulich, S. M., & Pearson, W. G., Jr. (2019). Three-/four-dimensional ultrasound technology in speech research. *Perspectives of the ASHA Special Interest Groups, 4*, 733-747.

Maas, E., Robin, D. A., Hula, S. N. A., Freedman, S. E., Wulf, G., Ballard, K. J., & Schmidt, R. A. (2008). Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech-Language Pathology, 17*(3), 277-298.

MathWorks Inc. (2000). MATLAB (Version 6.1) [Computer program]. Natick, MA. Retrieved from https://www.mathworks.com/products/matlab/

McAllister Byun, T., Buchwald, A., & Mizoguchi, A. (2016). Covert contrast in velar fronting: An acoustic and ultrasound study. *Clinical Linguistics & Phonetics, 30*(3-5), 249-276.

McReynolds, L. V., & Bennett, S. (1972). Distinctive feature generalization in articulation training. *Journal of Speech and Hearing Disorders, 37*(4), 462-470.

Preston, J. L., McCabe, P., Tiede, M., & Whalen, D. H. (2019). Tongue shapes for rhotics in school-age children with and without residual speech errors. *Clinical Linguistics & Phonetics, 33*(4), 334-348.

Python Software Foundation. (2016). Python programming language (Version 3.5.2) [Computer program]. Retrieved from https://www.python.org/downloads/

R Core Team. (2019). R: A language and environment for statistical computing [Programming language]: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Smit, A. B., Hand, L., Freilinger, J. J., Bernthal, J. E., & Bird, A. (1990). The Iowa articulation norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders, 55*(4), 779-798. Retrieved from http://jshd.pubs.asha.org/article.aspx?articleid=1775656

Studdert-Kennedy, M., & Goldstein, L. (2003). Launching language: The gestural origin of discrete infinity. In M. H. Christiansen & S. Kirby (Eds.), *Language Evolution* (pp. 235-254). New York, NY: Oxford University Press, Inc.

Sugden, E., Lloyd, S., Lam, J., & Cleland, J. (2019). Systematic review of ultrasound visual biofeedback in intervention for speech sound disorders. *International Journal of Language and Communication Disorders, 54*(5), 705-728.

Tiede, M. (2016). GetContours (Version 1.3) [MATLAB script]. Retrieved from https://github.com/mktiede/GetContours

Tiede, M. (2017). ComputeCurvature (Version 1.0) [MATLAB script].

Van Riper, C. (1978). *Speech correction: Principles and methods (6th edition)*. Englewood Cliffs, NJ: Prentice-Hall.

Thank you! Questions?

heather.kabakoff@nyu.edu

- Acknowledgments: Special thanks to members of the Biofeedback Intervention Technology for Speech (BITS) Lab at NYU for assistance with data collection and processing.

- This work is supported by the National Institutes of Health (NIDCD Grant F31DC018197, PI: Kabakoff; NIDCD Grant R01DC013668, PI: Whalen).
- Additional support provided by the Acoustical Society of America and the ASHFoundation.