Narrow Phonetic Transcription of Voicing Ambiguity in Stroke Survivors

Morgan McGowan, Jordan Jarrett, and Katarina Haley

Purpose

The overall objective of this research is to understand voicing ambiguity in stroke survivors with aphasia and coexisting phonemic paraphasia (APP) or apraxia of speech (AOS). In the present study, the objective was to develop an effective method of training to improve reliability of narrow phonetic transcription of voicing ambiguity.

Background

APP and AOS

- APP is reported to impact "phonological-linguistic retrieval or assembly," leading to more instances of substitutions. A substitution is reported when the phoneme produced by the speaker is entirely different from the intended phoneme.[1]
- AOS is classified as a "phonetic-motor disorder of speech production," leading to more instances of distortions. Distortions are sounds that are altered, but still with the boundaries of the target phoneme.[1]

Transcription Methods

- Broad phonetic transcription describes speech on the phonemic level, using IPA symbols. Narrow phonetic transcription allows a more sensitive description of speech samples that do not cross phonemic boundaries, allowing for better notation of distortions.[2]

Cunningham and colleagues (2016)

- Researchers used narrow phonetic transcription and transcriber training that consisted of reviewing IPA symbols, establishing operational definitions of distortions, and practicing the coding of motor speech evaluations. Their coding system consisted of 35 diacritics, with two separate diacritics for voicing and devoicing. For voicing distortions, they observed an intraclass correlation of 0.63 for stop consonants and 0.57 for fricatives and affricates.

Voicing

- Many studies have reported distorted voicing control patterns in speech of stroke survivors with APP or AOS which corresponds to how voicing distortions were the second most commonly noted distortion in Cunningham's study.[5,1]
- The use of a single diacritic for voicing ambiguity simplifies the coding system and the perceptual task of the transcribers. This corresponds to the idea that "the most efficient and reliable coding system is one that is closely matched to the capacity of auditory discriminations and auditory memory."[4]
- Focusing on voicing ambiguity enables the transcriber to subdue the tendency to categorize sounds as substitutions based on language biases rather than hearing the distortions of the sound.[3]
- The frequent occurrences of voicing ambiguity heard in everyday speech, such as the word-final stop in "gag" and word-medial stop in "zipper," were discussed to emphasize that not all instances of voicing ambiguity are an example of distorted speech.

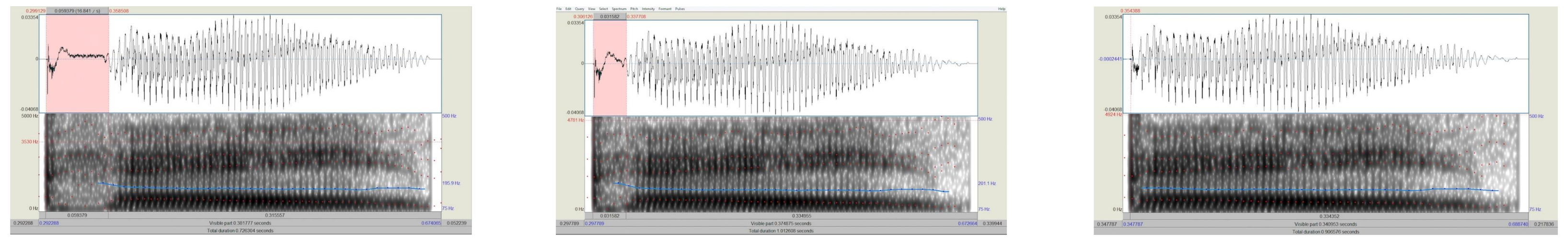

Figure 1. Voice Onset Time for Speech Output
Above are three examples from the voicing continuum practice that comprised a portion of our transcriber training. The image on the far left displays a typical VOT for a completely voiceless sound and the far right image demonstrates a completely voiced sound. The middle image, however, displays the VOT of an ambiguously voiced sound.

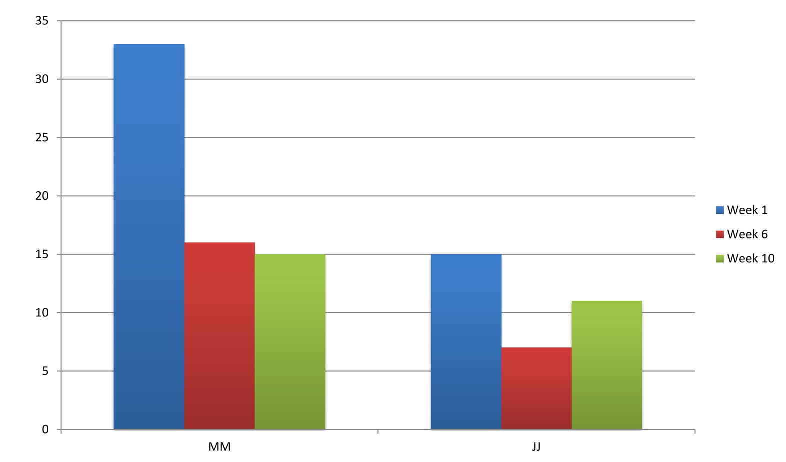

Figure 2. Frequency of Perceived Voicing Distortions
The general decrease in number of voicing distortion errors over time reflects how the transcriber training helped us distinguish between different errors and led us to a better understanding of what is defined as a voicing ambiguity error.

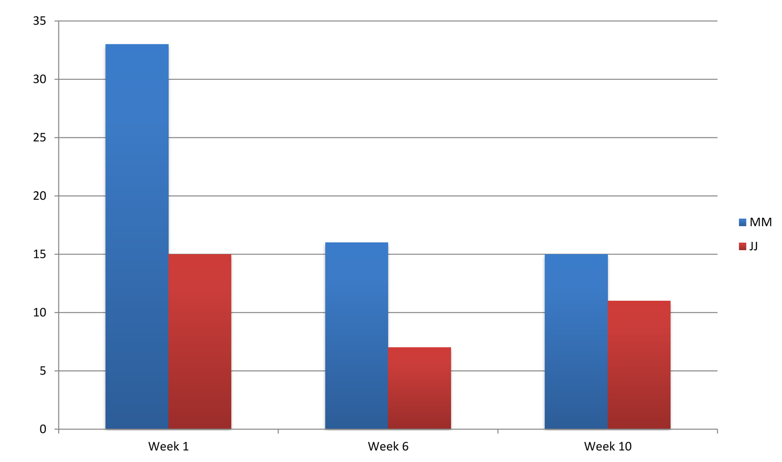

Figure 3. Frequency of Perceived Voicing Distortions
As the training progressed, the difference in perceived distortion errors between transcribers decreased. This reflects how the training in place helped improve reliability over time.

Method

- Four speech evaluations were transcribed independently and compared for agreement of voicing distortions.
- Point-to-point agreement was defined as the number of sound segments agreed upon as having a voicing ambiguity divided by the total number of segments (agreements + disagreements). We discussed coding discrepancies, re-listened jointly to any speech samples that we did not agree on, and either came to a consensus after listening to it twice, or settled on a disagreement and noted potential explanations and remediation strategies.

Results

	Stops	Fricatives
Speaker 1	67%	50%
Speaker 2	89%	75%
Speaker 3	88%	60%
Speaker 4	56%	75%
Overall	76%	67%

Figure 4. Inter-observer Reliability

	Word Initial	Word Medial	Word Final
Speaker 1	100%	50%	100%
Speaker 2	100%	60%	86%
Speaker 3	60%	75%	100%
Speaker 4	71%	83%	33%
Overall	73%	68%	80%

Figure 5. Inter-observer Reliability

Training

Coding System

The coding system used consisted of 11 diacritics. Voicing ambiguity, covering both devoicing and partially voiced sounds, is represented by one diacritic.

Training Procedure

- Initially listening to continuous speech and thinking of sounds as distortions, substitutions, distorted substitutions, omissions, and additions
- Followed by 40 hours of transcription practice, where training transcribers were assigned 4 motor speech evaluations which were coded using Praat. Weekly discussions between transcribers focused on the 1-2 new diacritics that were learned each week. For each diacritic, we collectively developed operational definitions, analyzing how that distortion could be applied to vowels and consonants and whether certain instances fit into different distortion criteria.
- We listened to clear examples of different distortions that were identified by an experienced transcriber. We then compiled our own lists of clear examples of certain distortions to use and compare.

Voicing Training

- Voicing ambiguity was the first diacritic discussed in training.
- During that week of training we analyzed computer generated word initial and word final stops, ranking them on a continuum scale from 1-9, 1 representing completely voiced and 9 representing completely devoiced. We listened to examples from motor speech evaluations using the same ranking system.
- To better attune our perception of errors, we practiced producing stops with voicing ambiguity.

Suggestions for Future Training

- Fricative production practice and fricative voicing continuum analysis to help increase agreement for sub-phonemic voicing variations in fricatives.
- Targeted training for the word-medial consonants.
- Ongoing reliability calculations to help transcribers acknowledge the differences in perception among individuals and to gain confidence in defending sounds they hear as ambiguously voiced.

Conclusions

Our strong percent agreement suggests that the transcription training protocol was effective in identifying voicing ambiguity in stroke survivors with AOS or APP. The training and reliability estimation experiences indicate that targeted exercises and subjective input from transcribers may be used constructively to shape future training methods, so that researchers and clinicians accurately document salient sub-phonemic speech properties.

Acknowledgements

We would like to express our extreme gratitude to our co-author Dr. Haley. She provided us with a great research opportunity, allowing us to learn so much about narrow phonetic transcription and distortions with her continued guidance and support.

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

References

1. Cunningham, K., Haley, K.L., & Jacks, A. (2016). Speech sound distortions in aphasia and apraxia of speech: Reliability and diagnostic significance. Aphasiology, 30(4), 396–413
2. Haley, K. L., Bays, G. L., & Ohde, R. N. (2001). Phonetic properties of aphasic-apraxic speech: A modified narrow transcription analysis. Aphasiology, 15(12), 1125–1142.
3. Haley, K. L. (2004). Vowel duration as a cue to postvocalic stop voicing in aphasia and apraxia of speech, Aphasiology, 18, 443-456.
4. Kent, R. D. (1996). Hearing and Believing: Some Limits to the Auditory-Perceptual Assessment of Speech and Voice Disorders. American Journal of Speech-Language Pathology, 5(3), 7.
5. Nespoulous, J. L., Baque, L., Rosas, A., Marczyk, A., Estrada, M., (2013) Aphasia, phonological and phonetic voicing within the consonantal system: preservation of phonological oppositions and compensatory strategies. Language Sciences, 39, 117-125.