# Effects of Hearing Aid Compression on Temporal Cues in Audiovisual Speech

Kaylah Lalonde, Ph.D.
Audiovisual Speech Processing Laboratory, Boys Town National research Hospital, Omaha, NE

## BACKGROUND

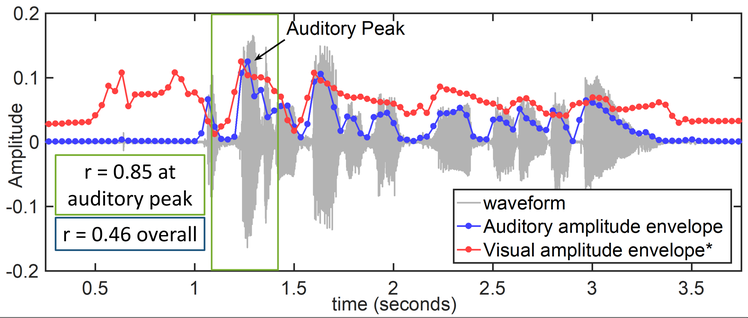
Hearing aids and visual speech (lip reading) are two strategies recommended by audiologists to enhance speech understanding in noise. [1]

### Visual Speech

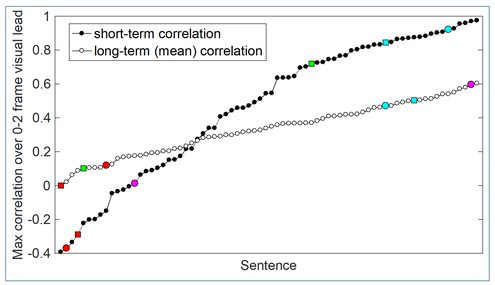
Auditory and visual speech streams are correlated in multiple ways. [2-3]

Neurophysiological evidence indicates that mouth movements help the auditory cortex to track the temporal amplitude envelope of auditory speech (the slow time-varying changes in signal energy). [4-7]

This help listeners predict the timing of peaks in the auditory signal and direct auditory analyses to the speech signal of interest, rather than surrounding background noise. [8, 9]




*area of mouth opening, scaled to peak of the auditory amplitude envelope




The natural correlations between auditory and visual amplitude envelopes vary across utterances.

One study showed higher AV benefit to speech detection for a sentence with higher correlations than a sentence with lower correlations. [8]

### Wide dynamic range compression (WDRC)

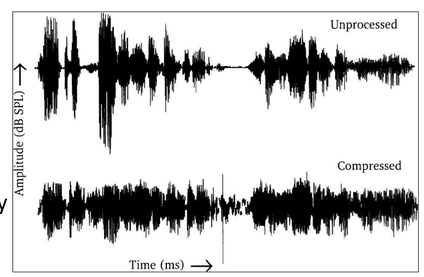
WDRC is a key feature in hearing aids and a primary means of improving auditory speech understanding in individuals with hearing loss. [10, 11]

Hearing aids with WDRC amplify low intensity parts of auditory signals more than the higher-intensity parts.

WDRC distorts the amplitude envelope of auditory speech. [12]

Distortions of the auditory amplitude envelope likely disrupt the natural temporal correspondence between auditory and visual speech.

Therefore, WDRC may affect the benefit derived from visual speech.




Souza, 2002. Unprocessed (top) and compressed (bottom) speech waveforms

## PURPOSE

**Long-term goal:** Understand the impact of wide dynamic range compression (WDRC), a key feature in hearing aids, on AV speech perception in individuals with hearing loss.

**Questions addressed in this proposal:**

- How does WDRC affect the AV temporal correspondences?
- How does WDRC affect AV speech perception benefits in individuals with normal hearing?

**Central hypothesis:** WDRC decreases the temporal correspondence between auditory and visual speech, which in turn affects listeners' ability to benefit of visual speech.

## AIMS AND HYPOTHESES

1. **Determine which AV temporal correlations are related to speech perception benefit.**

   Hypothesis: Correlations at the auditory peak are important for detecting speech in noise and correlations over the whole sentence are important for speech recognition.

2. **Determine the effect of WDRC on auditory-visual temporal correspondences and AV speech perception benefit in listeners with normal hearing.**

   Hypothesis: WDRC decreases AV temporal correlations, which decreases AV detection benefit.

   Competing hypotheses:

   a) Decreased correspondence between the auditory and visual envelopes may decrease AV recognition benefit
   b) Visual envelopes in the WDRC condition may serve as a complementary cue that helps to restore the original information about the degraded acoustic amplitude envelope.

## APPROACH

**Participants:** Young adults with normal hearing and normal or corrected-to-normal vision

**Stimuli:** 74 AV sentences spoken by 4 talkers

- Syntactically correct but semantically incorrect; e.g., "The fuzzy vests lead since then."

**Stimulus analysis:**

- Extracted the area of the mouth opening in each video frame using FaceScanner software. [12]
- Calculated the instantaneous correlation between the area of the mouth opening and the auditory amplitude envelope at each video frame over the duration of the auditory sentence, using a flat 334 ms moving window and 0-2 frames of visual lead.
- Calculated the **short-term** correlation over the 334 ms centered on the peak of the auditory envelope and the **mean** (long-term) correlation over the duration of the auditory sentence.

**Wide Dynamic Range Compression:**

- Apply WDRC to the auditory stimuli and examine changes to the mean and short-term correlations.

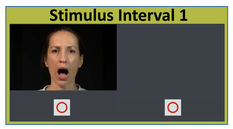
| Compression Settings | | | |
|---|---|---|---|
| Input:Output Ratio | 1:1 | 3:1 | 6:1 |
| Attack Time | <1 ms | 10 ms | 20 ms |
| Release Time | 12 ms | 100 ms | 800 ms |

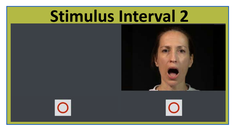
**Experimental Tasks:**

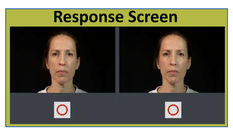
- Conducted with unprocessed and WDRC compressed signals.

***Adaptive Detection Threshold Measurement:***

- Participants select the interval that contains auditory speech.
- Same noise and visual signal (picture or video) presented in both intervals.



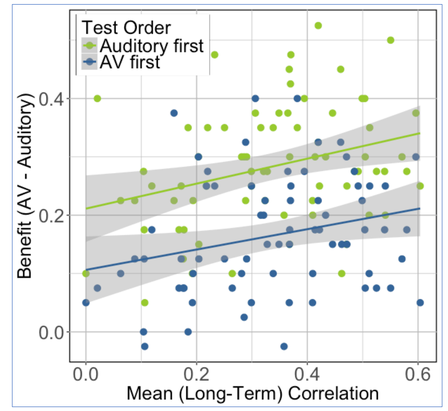




***Sentence Recognition:***

- 72 auditory and AV sentences presented in noise.
- Participants repeat the sentences; Experimenter scores accuracy of keyword identification.

## PRELIMINARY RESULTS

- 20 adults completed the sentence recognition task with unprocessed stimuli. Test order was counterbalanced across participants.
- Calculated mean benefit (AV accuracy – auditory accuracy) across subjects, for each sentence and test order.
- Mixed linear modeling was conducted on the benefit data with a fixed effect of test order, a random effect of mean (long-term) AV correlation, and a random intercept for sentence.
- Results showed an effect of test order, $t = 8.852$, $p < 0.0001$, and an effect of mean AV correlation, $t = 3.122$, $p = 0.0026$.
- Consistent with practice effects, benefit was greater for participants tested in the auditory condition first.
- Consistent with our hypothesis, benefit was greater for sentences with a higher auditory-visual correlation.




## NEXT STEPS / DELIVERABLES

- Assess the behavioral relevance of other AV signal correspondences.
- Assess the effects of WDRC on AV speech perception in individuals with hearing loss.
- Assess the development of children's ability to use these and other visual cues to better understand speech in noise.

## ACKNOWLEDGEMENTS

The project described was supported by the National Institute of General Medical Sciences, **1U54GM115458**. The content is solely the responsibility of the author and does not necessarily represent the official views of the NIH.

## References

1. Tye-Murray (2009). *Foundations of Aural Rehabilitation: Children, Adults, and Their Family Members* (3rd ed.). Clifton Park, NY: Delmar.
2. Chandrasekaran et al. (2009). The natural statistics of audiovisual speech. *PLoS Comput Biol, 5*(7), 1-18.
3. Yehia et al. (1998). Quantitative association of vocal-tract and facial behavior. *Speech Commun, 26*, 23-43.
4. ten Oever et al. (2014). Rhythmicity and cross-modal temporal cues facilitate detection. *Neuropsychol Rehab, 63*, 43-50.
5. Schroeder, & Foxe (2005). Multisensory contributions to low-level 'unisensory' processing. *Curr Opin Neurobiol, 15*, 454-458.
6. Schroeder et al. (2008). Neuronal oscillations and visual amplification of speech. *Trends Cog Sci, 12*, 106-113.
7. Zion-Golumbic et al. (2013). Visual input enhances selective speech envelope tracking in auditory cortex at a "cocktail party". *J Neurosci, 33*, 1417-1426.
8. Grant & Seitz (2000). The use of visible speech cues for improving auditory detection of spoken sentences. *J Acoust Soc Am, 108*(3), 1197-1208.
9. Summerfield (1987). Some preliminaries to a comprehensive account of audio-visual speech perception. In B. Dodd & R. Campbell (Eds.), *Hearing by eye: The psychology of lipreading* (pp. 3-51). London: Erlbaum Associates, Inc.
10. Wood (2004). Relative benefits of linear analogue and advanced digital hearing aids. *I J Audiol, 43*(3), 144-155.
11. Jenstad et al. (1999). Comparison of linear gain and wide dynamic range compression hearing aid circuits: aided speech perception measures. *Ear Hear, 20*(2), 117-126.
12. Souza (2002). Effects of compression on speech acoustics, intelligibility, and sound quality. *Trends Amplif, 6*(4), 131-165.
13. Saragih et al.. (2011). Deformable model fitting by regularized landmark mean-shift. *I J Comput Vis, 91*, 200-215.

## CONTACT INFORMATION

kaylah.lalonde@boystown.org (531) 355-5631