On the Importance of Information-Bearing Acoustic Changes for Understanding Speech in Simulated Electrical-Acoustic Stimulation

Christian E. Stilp, University of Louisville
Gail Donaldson, University of South Florida
Soohee Oh, University of South Florida
Ying-Yee Kong, Northeastern University

1aPPb2

INTRODUCTION

- Information-bearing acoustic changes in the speech signal are important for understanding speech. This has been demonstrated using cochlea-scaled entropy in simulated cochlear implant speech (CSE_{CI}), which measures short-term spectral changes calculated across all channels of noise-vocoded speech (Stilp *et al.*, 2013; Stilp, 2014; Stilp & Goupell, 2015).
  - Poorer sentence intelligibility occurred when high-CSE_{CI} intervals were replaced by noise as compared to replacing low-CSE_{CI} intervals.
- Speech intelligibility for CI users is often aided by preserved low-frequency acoustic cues in the non-implanted ear (Electric-Acoustic Stimulation, or EAS; Kong *et al.*, 2005; Kong & Carlyon, 2007).
- While listeners are known to benefit from the addition of low-frequency intact speech to vocoded speech (hybrid hearing), it is unclear how they rely on information-bearing acoustic changes from the two different sources (intact low frequencies, vocoded higher frequencies).
- Here, we investigated the perceptual importance attributed to information-bearing acoustic changes in vocoded versus hybrid speech with relatively few spectral channels (conditions which were not systematically studied previously).

METHODS

Participants

- 37 native English speakers ages 18-35; normal hearing confirmed via hearing screening.

Stimuli

- IEEE sentences spoken by female talker.
  - Spectrum divided into 3-6 channels using Greenwood (1990) spacing.
  - Amplitude envelopes extracted and assigned to white Gaussian noise.
  - Vocoded stimuli: noise-channel vocoding from 80 to 8000 Hz.
  - Hybrid stimuli: low-pass-filtered intact speech (<500 Hz) combined with high-pass-filtered vocoded speech (>500 Hz).
- CSE_{CI} calculation
  - CSE_{CI} = Euclidean distances between RMS-amplitude profiles of successive 16-ms sentences slices, summed into 80-ms intervals.
  - Same calculations performed for vocoded and hybrid sentences.
  - In each sentence, four 80-ms intervals with high- or low-CSE_{CI} were replaced with speech-shaped noise; continuous sentences had no noise replacement.

Procedure

- 1. Training
  - Familiarization (passive listening) followed by training (testing with feedback).
  - Baseline testing: criterion of ≥27% on 3-channel vocoded sentences or ≥31% on 4-channel vocoded sentences.
    - 13 listeners failed to meet criteria, did not participate in testing.
- 2. Testing
  - CSE_{CI} levels tested at two levels of spectral resolution (3/5 or 4/6, n=12 in each); EAS (vocoded/hybrid) counterbalanced across sessions.
  - Mixed design employed due to limited number of IEEE sentence lists:
    - Within-subjects factors: CSE_{CI}, EAS
    - Between-subjects factor: higher (4/6 channels) vs. lower (3/5) numbers of vocoder channels
- Stimuli presented diotically at 70 dB SPL.

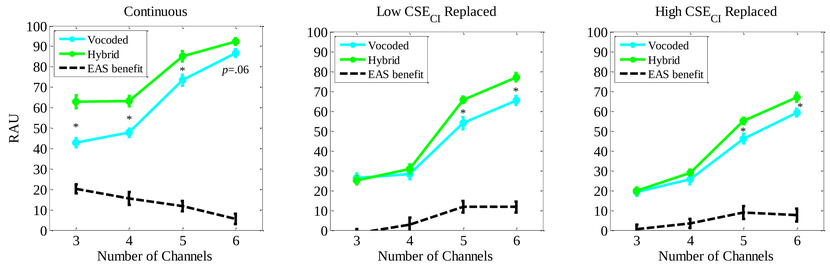
Main Findings

Results were arcsine-transformed (Studebaker, 1985) for analysis. Analyses used within-subjects ANOVA (comparing performance at 3 vs. 5 channels, or 4 vs. 6 channels) or mixed ANOVA (comparisons across listener groups) and corresponding paired-samples or independent-samples *t*-tests. Statistically significant differences are indicated by *.

In general, performance was impaired more when high-CSE_{CI} intervals were replaced by noise than when low-CSE_{CI} intervals were replaced, consistent with Stilp *et al.* (2013), Stilp (2014), and Stilp and Goupell (2015).

EAS benefit (RAU improvement for hybrid condition relative to vocoded condition) varies with stimulus condition:

- For continuous sentences, EAS benefit decreases with more spectral channels.
- For noise-interrupted sentences, EAS benefit increases with more spectral channels.
  - No benefit at 3-4 channels, comparable significant benefits across CSE conditions at 5-6 channels.
- EAS benefit does not interact with CSE conditions.



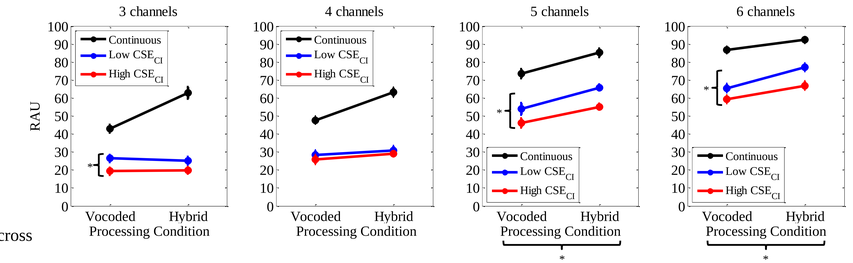
Number of Spectral Channels

	3	4	5	6
EAS			*	*
CSE_{CI}	*		*	*

Contrasts Between Different Numbers of Spectral Channels

	3/4	3/5	3/6	4/5	4/6	5/6
EAS		*	*		*	
CSE_{CI}	*	*		*	*	

* = statistically significant difference for this factor across the given pair of spectral channels at $p < .05$
(only low-CSE_{CI} and high-CSE_{CI} conditions analyzed)
No EAS-by-CSE_{CI} interactions were significant



Perceptual Importance of CSE_{CI}

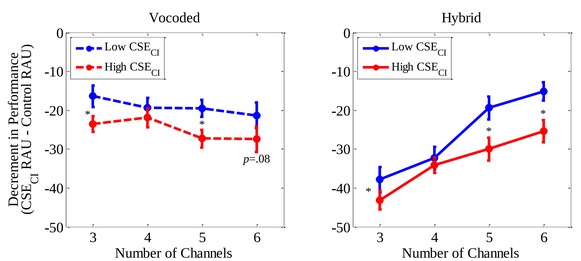
- Decrements from control performance reveal the perceptual importance of the speech in sentence intervals that were replaced by noise.
  - Larger decrements indicate greater importance of the replaced speech

Vocoded

- Decrements are fairly flat across 3-6 spectral channels.
  - Stilp and Goupell (2015) reported similar results for 4 channels but far larger decrements for high-CSE_{CI} sentences with 6 channels.

Hybrid

- Largest decrements with few spectral channels.
  - May be due to abolishing EAS benefit rather than high CSE_{CI} importance.
- Smaller decrements with more channels.
  - Ceiling effects for continuous conditions may be contributing to smaller decrements.



DISCUSSION

Listeners are able to use information-bearing acoustic changes to understand hybrid speech.

- For sentences with 5 and 6 channels, performance suffers more when high-CSE_{CI} intervals are replaced by noise than when low-CSE_{CI} intervals are replaced.
- A similar pattern was observed for vocoded speech, consistent with Stilp and Goupell (2015).

For sentences with very few (3-4) spectral channels, listening mode and information-bearing acoustic changes had little effect on performance.

- Large EAS benefit when sentences are intact; near-zero EAS benefit when sentences are interrupted by noise.
- Performance only slightly worse when high-CSE_{CI} intervals were replaced as compared to replacing low-CSE_{CI} intervals.
- Similar to Stilp and Goupell (2015), listeners utilized information-bearing acoustic changes to facilitate recognition of 6-channel sentences, but no differences were observed for 4-channel sentences.
  - This is despite differences in materials (IEEE vs. TIMIT sentences) and procedure (training with feedback versus virtually no practice).

The present data provide no evidence that listeners rely on CSE_{CI} differently in vocoded versus hybrid speech.

- Information-bearing acoustic changes produced different patterns of decrements, but this could be due to abolishing the EAS benefit for 3-4-channel sentences. It made little difference whether low-CSE_{CI} or high-CSE_{CI} intervals were replaced.
- CSE_{CI} is a measure of broadband spectral change, and this often captures changes in the speech envelope. Envelope information is relatively consistent across full-spectrum and vocoded speech (Stilp *et al.*, 2013). This congruence may explain the lack of interactions between CSE_{CI} and vocoder / hybrid conditions, as information-bearing acoustic changes do not differ very much across these processing conditions.
- The intact low-frequency channel (<500 Hz) was not weighted differently from noise-vocoded channels in CSE_{CI} calculations. Future research will consider frequency-specific weighting in CSE_{CI} calculations to reflect EAS benefit.

REFERENCES

Kong, Y.-Y. & Carlyon, R.P. (2007) *JASA, 121(6)*, 3717-3727.
Kong, Y.-Y., Stickney, G.S., & Zeng, F.-G. (2005) *JASA, 117(3)*, 1351-1361.
Greenwood, D.D. (1990) *JASA, 87(6)*, 2592-2606.
Stilp, C.E. (2014) *JASA, 135(3)*, 1518-1529.
Stilp, C.E. & Goupell, M.J. (2015) *JASA, 137(2)*, 844-855.
Stilp, C.E., Goupell, M.J., & Kluender, K.R. (2013) *JASA, 133(2)*, EL136-EL141.
Studebaker, G.A. (1985) *JSLHR, 28(3)*, 455-462.