

# Children need coherence masking protection

Eric Tarr and Susan Nittrouer

Department of Otolaryngology, Ohio State University

## BACKGROUND

For much of the history of speech research the role of acoustic cues in phonetic decisions was studied, but in 1981 Best et al. showed that there is more to perception than just the extraction of acoustic cues: Signal components must cohere.

Perceptual coherence refers to the fusing of signal components to support recovery of the sensory-generating object.

Nittrouer and Crowther (2001) found that children were less able to segregate separate components of speech than adults. That led to the question of whether children need to have signals they can perceptually fuse more than adults.



### Coherence Masking Protection (CMP)

Gordon (1997; 2000) presented F1 for 2 vowels in noise to listeners, alone and with a constant F2/F3 cosignal. Recognition thresholds were at poorer signal-to-noise ratios (SNRs) with the cosignal, even though it contributed no new information (i.e., masking release).

CONCLUSION: The stable cosignal made it easier to fuse signal components to recover the sound-generating object, and so speech was recognized at poorer SNRs.

Gordon showed that the effect was not due to F1 and cosignal sharing harmonic structure, but was due to the signal being heard as speech.

### QUESTION

Will children benefit more from the cosignal than adults?

## STIMULI

Speech sounds were steady-state vowels, 60 ms long

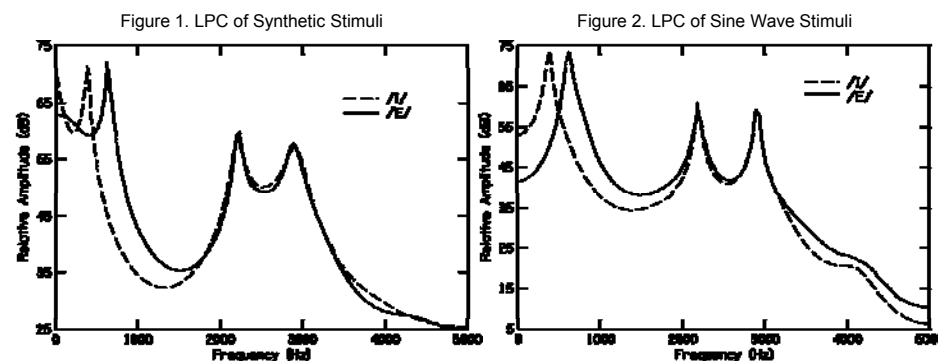
### Synthetic Stimuli

f0 = 125 Hz  
F1 frequency = 375 Hz (/i/) or 625 Hz (/ε/), bandwidth 50 Hz  
F1 amplitude = 74-48 dB SPL in 2-dB increments  
F2 = 2200 Hz, bandwidth of 110 Hz,  
F3 = 2900 Hz, bandwidth of 170 Hz  
Identical F2/F3 for both vowels, -12 dB from F1

### Sine Wave Stimuli

Individual sine waves substituted at formant peaks

**Masking Noise:** White noise low-pass filtered below 1000 Hz at 62 dB to mask F1, 600 ms in duration. Stimuli presented 420 ms after the onset of noise.



## METHOD

### Participants

Adults (24), 8-year-olds (24), 5-year-olds (14).

### Criterion for inclusion

Training was provided with feedback using, in order: 1. natural words, 2. synthetic or sine wave words, and 3. synthetic or sine wave 60-ms stimuli (F1 or F1 + cosignal). Participants had to correctly label 9 of 10 consecutive stimuli in quiet to move to testing.

### Task

Participants labeled stimuli by saying the word *bit* or *bet* and pointing to a picture associated with the word. An adaptive procedure was performed following the '3-down, 1-up' procedure to obtain an 80% correct labeling threshold (Levitt, 1971). Masking noise remained constant at 62 dB SPL, speech stimuli varied from 74 dB to 48 dB. Step size of 8 dB was used for the first two runs, 4 dB for the next 2 runs, and 2 dB for the last 12 runs. Mean of the last 8 runs was the threshold.

4 conditions: 1. synthetic, F1; 2. synthetic, F1 + cosignal;  
3. sine wave F1; 4. sine wave, F1 + cosignal

## RESULTS

	Synthetic Stimuli			Sine Wave Stimuli		
	F1-Only	Cosignal	$\mu$ difference	F1-Only	Cosignal	$\mu$ difference
Adults	-0.8 (3.4)	-4.1 (1.3)	-3.3 (2.2)	-5.6 (1.5)	-4.2 (1.7)	1.4 (1.2)
8-year-olds	2.5 (4.0)	-3.4 (1.0)	-5.9 (3.2)	-3.0 (4.7)	-1.8 (3.7)	1.2 (3.7)
5-year-olds	8.0 (3.3)	-0.4 (3.6)	-8.4 (3.2)	3.5 (4.8)	3.4 (4.3)	-0.1 (4.4)

Table 1. Mean SNR for each condition. Standard deviations are in parentheses.

### Synthetic Stimuli

(Table 1, left, Figure 3 below)

All age groups showed improvement in the Cosignal condition compared to the F1-Only condition.

This effect increased with decreasing age.

### Sine Wave Stimuli

(Table 1, right, Figure 4 below)

Adults performed better in the F1-Only condition

8- & 5-year-olds performed similarly in the two conditions

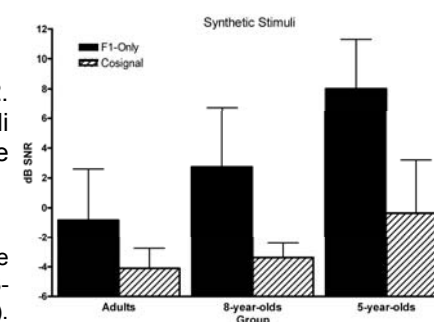


Figure 3. Mean thresholds for synthetic stimuli.

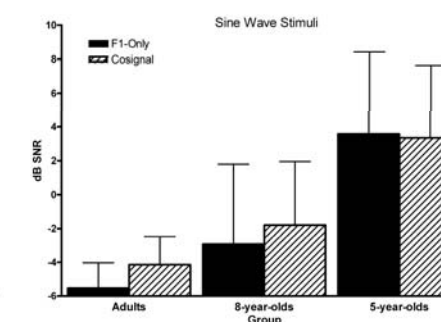


Figure 4. Mean thresholds for sine wave stimuli.

Synthetic F1-Only stimuli had the highest (worst) thresholds of all conditions for all ages.

Children had greater masking release than adults for the synthetic Cosignal condition.

Adults had their lowest (best) threshold for the sine wave F1-Only condition, but they did not hear these signals as "speech."

No masking release was found for sine wave stimuli, which were not heard as speech.

## CONCLUSIONS

Spectrally coherent speech signals facilitate recognition under degraded conditions.

Coherence masking protection is obtained only when stimuli are recognized as phonetically relevant objects.

**Children need rich and coherent speech signals even more than adults.**

## REFERENCES

- Best, C. et al. (1989). Discovering phonetic coherence in acoustic patterns. *Percept. & Psychophys.* 45, 237-250.
- Gordon, P. (1997). Coherence masking protection in speech sounds: The role of formant synchrony. *Percept. & Psychophys.* 59, 232-242.
- Gordon, P. (2000). Masking Protection in the perception of auditory objects. *Speech Communication* 30, 197-206.
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *JASA* 49, 467-477.
- Nittrouer, S. & Crowther, C. (2001). Coherence in children's speech perception. *JASA* 110, 2129-2140.

## ACKNOWLEDGEMENT

Work supported by NIDCD Grant No. DC-00633