

**5aPPa20**  
**Contribution of consonant versus vowel information to sentence intelligibility by normal and hearing-impaired listeners**



T. Zachary Burkle\*  
Diane Kewley-Port  
Larry Humes  
Jae Hee Lee



Department of Speech & Hearing Sciences  
Indiana University  
Bloomington, IN 47405  
\*e-mail: tzachbur@indiana.edu

1

## Abstract

The purpose of this study was to examine the contribution of information provided by vowels versus consonants to sentence intelligibility in young normal-hearing (YNH) and elderly hearing-impaired (EHI) listeners. Sentences were presented in three conditions, with either the vowels or the consonants replaced with speech shaped noise, or unaltered. Sentences from male and female talkers in the TIMIT database were selected. EHI subjects listened at 95 dB SPL, and YNH subjects at both 95 and 70 dB SPL. Subjects listened to each sentence twice and were asked to repeat the entire sentence after each presentation. Words were scored correct if identified exactly. Average performance for unaltered sentences was greater than 94%. Vowel-present conditions were always significantly more intelligible than consonant-present conditions, similar to data reported by Cole and colleagues [Proceedings of ICASSP, 1996]. Across groups, performance in the vowel-present conditions exceeded that in the consonant-present conditions by 14 to 40%, although EHI subjects performed more poorly than YNH subjects. In contrast to written English, vowels in spoken language carry more information about sentences than consonants for both normal and hearing-impaired listeners. [Research supported by NIHDCD-2229]

2

## Introduction

The objective of this work was to examine the contributions of vowels and consonants to sentence intelligibility in both young normal-hearing and elderly hearing-impaired listeners.

If audibility is ensured, most researchers and clinicians hypothesize that consonants provide more critical information than vowels to sentence intelligibility.

...Is this true?

3

## Background

Early experiments by Fletcher (1929) focused on articulations of vowels and consonants over a transmission line. Since consonants routinely had more errors it was concluded that they would contribute more towards the overall intelligibility than vowels.

Cole et al. (1996) used a noise replacement paradigm, in which either vowels or consonants were replaced by noise in sentences, and listeners were asked to repeat what they heard. Cole et al. found that young normal-hearing listeners identified twice as many words when vowels were present than when consonants were present.

4

## Purpose

The purpose of the present study was to compare the contribution of information provided by vowels versus consonants to sentence intelligibility when hearing is compromised in the elderly with typical hearing loss.

In addition, the purpose was to investigate vowel versus consonant information in relation to standard clinical evaluations of hearing loss.

5

## Methods

A noise replacement paradigm was used with sentence from the TIMIT database, North Midland dialect.

There were a total of three conditions:

1. **"Full"** had full phonetic information, i.e. unaltered sentences.
2. **"Vin"** had vowels only, with consonants replaced by speech-shaped noise.
3. **"Cin"** had consonants only, with vowels replaced by speech-shaped noise.

Speech-shaped noise was 10 dB louder for replacing vowels than for consonants, similar to natural speech.

6

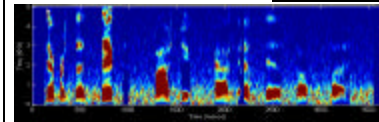
## Classification of TIMIT phonetic symbols into consonants and vowels

Group	Phonetic symbols	Total
Consonant	b d g p t k dx q jh ch s sh z zh f th v dh hh hv l r y w el eng nx m n ng em en	32
Vowel	iy ih eh ey ae aa aw ay ah ao oy ow uh uw ux er ax ix axr ax-h	20

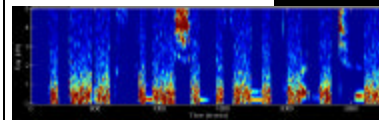
Phonetic boundaries in TIMIT sentences were scrutinized and usually accepted.

7

Example Vin (Vowel only) :  
"There was a thick, squashy  
crack of fist on flesh."



Example Cin (Consonant only) :  
"But that explanation  
is only partly true."



8

### Distribution of sentences over gender and conditions

42 test sentences were divided randomly into the three phonetic conditions.

	Number of Sentences (Male)	Number of Sentences (Female)	Number of Words per Condition
Full	7	7	115
Vin	7	7	114
Cin	7	7	114
Total	21	21	343

9

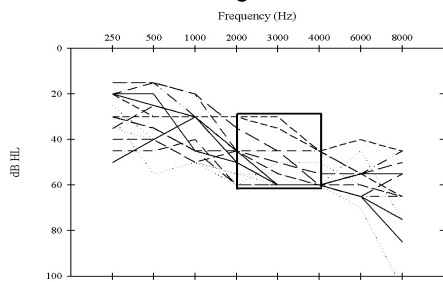
### Listener groups

1. YNH95 - 16 young normal hearing listeners (95 dB SPL)
2. YNH70 - 8 young normal hearing listeners (70 dB SPL)  
Age ranging from 20-35 years for YNH
3. EHI95 - 16 elderly hearing-impaired listeners (95 dB SPL)  
Age ranging from 65-80 years for EHI

All listeners were American-English natives from the North Midland dialect region and were screened for normal cognitive function.

10

### Audiograms



EHI95 hearing thresholds: The box indicates the hearing criteria for this study such that most consonants were audible

11

### Procedures

#### Presentation:

Sentences were low-pass filtered at 4400Hz and presented to the right ear at either 95 or 70 dB SPL. Each sentence was presented twice. After each presentation, listeners repeated any words they thought they heard.

#### Training:

6 practice sentences presented with feedback.

#### Testing:

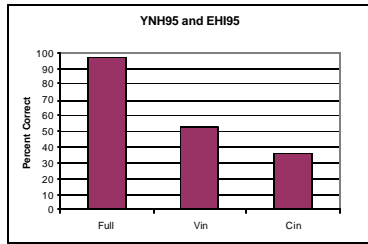
42 sentences randomly presented with no feedback.

#### Scoring:

The experimenter noted the words that were correctly repeated. The total words correct were counted as the score

12

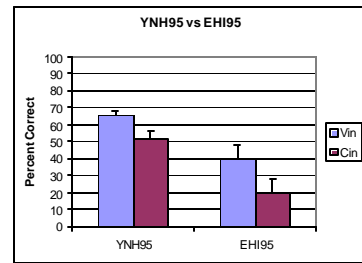
### Results (conditions)



Words correct in each condition averaged over YNH95 and EHI95 groups. Identification of Full sentences was equally accurate for both groups.

13

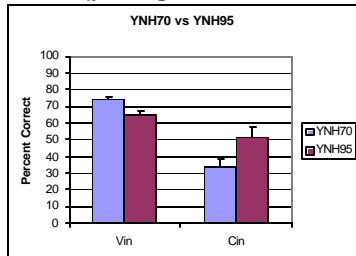
### Results (group)



Words correct by YNH95 and EHI95 in Vin and Cin conditions. Vin condition was better than Cin condition for both groups with no interaction.

14

### Results (young normal hearing)



Words correct by YNH70 and YNH95 in Vin and Cin conditions. Cin performance improved with higher signal level.

15

### Discussion

#### Relation to Previous Research:

2:1 better performance for Vin over Cin conditions by YNH70 and EHI95 replicates Cole et al. (1996)

#### Relation to Clinical Evaluations (Linear Regression):

Consonant only performance by EHI predicted best by pure-tone audiogram ( $R^2=0.93$ , Variables = 2K, 3K, 250, 500Hz)

Vowel only performance by EHI predicted best by speech intelligibility tests ( $R^2=0.70$ , Variables= word recognition, SPIN low)

16

## Conclusions

In noise replacement tasks, EHI performance

1. was poorer by 20% compared to YNH for both Cin and Vin conditions
2. was predicted accurately for Cin from audiogram
3. was predicted moderately for Vin by clinical tests of speech intelligibility.

Vowels in spoken language carry twice as much information about sentences than consonants, for both normal and hearing-impaired listeners.

17

## Acknowledgments

This research was supported by NIHDCD-2229.

## References

- Cole, R., Yan, Y., Mak, B., Fanty, M., and Bailey, T. (1996). Journal Proceedings of the 1996 International Conference on Acoustics, Speech, and Signal Processing. Atlanta, Ga.
- Fletcher, H. (1929). Speech and Hearing. New York: Van Nostrand
- Garofolo, J., Lamel, L., Fisher, W., Fiscus, J., Pallett, D., and Dahlgren, N. (1993). DARPA TIMIT: Acoustic-Phonetic Continuous Speech Corpus.

18