The development of a high-fidelity, multi-directional speech corpus Margaret K. Miller¹, Vahid Delaram², Rohit Ananthanarayana², Emily Buss³, Brian B. Monson², and G. Christopher Stecker¹



INTRODUCTION

Many standard speech testing materials (e.g., QuickSIN, BKB-SIN, NU-6, etc.) that are used both clinically and in hearing research were recorded decades ago using methodologies that limited the fidelity of the speech recordings.

Two **common limitations** of currently used speech corpora are that they were recorded:

1. With limited energy at frequencies higher than 8 kHz, also considered extended high frequencies (EHFs)

2. From the front of a talker only, resulting in speech tests that simulate target and maskers facing the listener.

Thus, we currently lack speech materials that account for broader aspects of real-world auditory scenes such as:

- 1. Speech directionality
- 2. Extended high frequency content

OBJECTIVE

The purpose of the current study was to develop a multi-directional, high-fidelity speech corpus using multi-channel anechoic recordings that can be used for future studies of speech perception in complex environments by diverse listeners.

PARTICIPANTS

Demographics:

- Fifteen male and 15 female talkers
- Age 21.3-60.5 years
- Native U.S. English-speaking

Screening for potential voice quality factors, e.g.:

- speech/hearing impairments
- vocal music training
- smoking
- respiratory issues

Spatial Hearing Lab, Boys Town National Research Hospital¹ University of Illinois Urbana-Champaign² Dept. of Otolaryngology/Head and Neck Surgery, University of North Carolina Chapel-Hill³

METHOD

- **Speech materials:**
- Four lists of BKB sentences (Bench, J., Kowal, A., & Bamford, J. (1979). The BKB (Bamford-Kowal-Bench) sentence lists for partiallyhearing children. British Journal of Audiology, 13(3), 108–112. https://doi. org/10.3109/03005367909078884)
- Digits 0-10
- Three-minute prompted narrative





Recording site:

- BTNRH Anechoic Chamber (SPACE) Lab
- 4x4x4 m interior dimension (ETS A100)

Microphones:

- Free-field condenser (B&K type 4189 or equivalent PCB Piezotronics 377B02)
- Seventeen azimuths 0 to 180° around the talker in 11.25° steps

Recordings (48 kHz and 24 bits) made using:

- Matlab (BKB sentences and digits)
- Audacity (extended narrative)

Post-processing included:

- best selection from three repeats of BKB sentences and digits
- removal of silent pauses (> 200 ms) from narratives

Speech acoustic analyses included average fundamental frequency, EHF spectral level, and directionality characteristics.



Resulting Corpus:

- Four BKB lists
- Digits 0-10
- Narratives produced by 30 talkers

Initial acoustic analyses indicate that EHF level does not change dramatically across speech materials (BKB vs. narratives) within talker, although EHF acoustic characteristics do vary across talkers and facing direction.

- speech corpus using standard speech materials.
- noise and speech maskers.
- content.

Acknowlegements

The authors would like to thank participating members of the Omaha community for lending us their voices and Brittany Williams, Ph.D.. This work was supported by the National Institute of Deafness and Other Communutication Disorders NIH R01-DC019745.



THE UNIVERSITY **f** NORTH CAROLINA at CHAPEL HILL

RESULTS

SCAN HERE

• Additional 15 BKB lists (21 total) produced by a subset of six talkers

CONCLUSIONS

• The goal of this study was to create an anechoic, high-fidelity, multi-directional

• We recorded more naturalistic narratives, useful for the creation of babble

• A large group of 30 talkers of both sexes permits researchers to select speech materials based on talker characteristics relevant to a specific task.

• The resulting speech corpus allows for more diverse and precise speech recognition testing, including testing effects of speech directionality and EHF