PSCR 2021

THE DIGITAL EXPERIENCE

#PSCR2021 • PSCR.GOV

NIST

PSCR

Designing Remote Listening Experiments for the Partially Muted Word Impairment

Jaden Pieper PSCR MCV

Jesse Frey PSCR MCV

NIST

#PSCR2021

PSCR

DISCLAIMER

Certain commercial entities, equipment, or materials may be identified in this document in order to describe an experimental procedure or concept adequately.

Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose.

RPO 2 – Exempt Human Subjects Research: The National Institute of Standards and Technology Research Protections Office reviewed the protocol for this project and determined it meets the criteria for 6 “exempt human subjects research” as defined in 15 CFR 27, the Common Rule for the Protection of Human Subjects.

*** Please note, unless mentioned in reference to a NIST Publication, all information and data presented is preliminary/in-progress and subject to change**

OVERVIEW

- Background
  - Intelligibility
  - Access Time Impairments
  - Motivation
- Remote Listening Experiments Design
- Current Experiments

QUALITY OF EXPERIENCE (QoE) MEASUREMENTS

- QoE KPIs for Mission Critical Voice (MCV)
  - Mouth-to-Ear Latency
  - Access Time
  - Voice Quality/Intelligibility
  - Probability of Successful Delivery
- Intelligibility
  - Describes whether information was actually received

INTELLIGIBILITY

- Modified Rhyme Test (MRT)
  - Used to test intelligibility of SCBA masks[1]
  - Batches of six words
    - *went, sent, bent, dent, tent, rent*
    - Words: consonant-vowel-consonant
    - Each batch: Either leading or trailing consonant varies
  - MRT Trial
    - Carrier phrase + word
    - "Please select the word *went*"
    - Success (identified) or Failure (mis-identified)
  - Over lots of trials, scores are generated
    - Score is value between 0 and 1
    - Corrected for guessing
  - High time cost

1: NFPA 1981 Standard on open-circuit self-contained breathing apparatus (SCBA) for emergency services (2007)

INTELLIGIBILITY

MRT: Standard Use Case Examples

Batch: *fun, sun, bun, gun, run, nun*
Distortion: Background noise + system

- Extreme
- Moderate
- Mild
- None

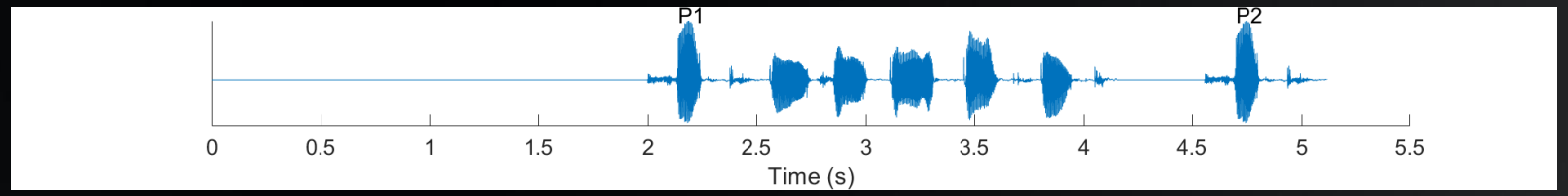
INTELLIGIBILITY

- Articulation Band Correlated Modified Rhyme Test (ABC-MRT)
  - Objective algorithm to provide estimates of true MRT scores
  - Most recent version is ABC-MRT16
  - Relies on temporal correlations within articulation index bands
    - Break speech into a “musical score”
    - Representation of speech in time and frequency
  - Costs: Cheaper and faster
    - Relatively low time cost
      - Can get estimated MRT scores “on demand”
    - Much lower infrastructure cost

ACCESS TIME MEASUREMENT CONCEPT

Audio Clips

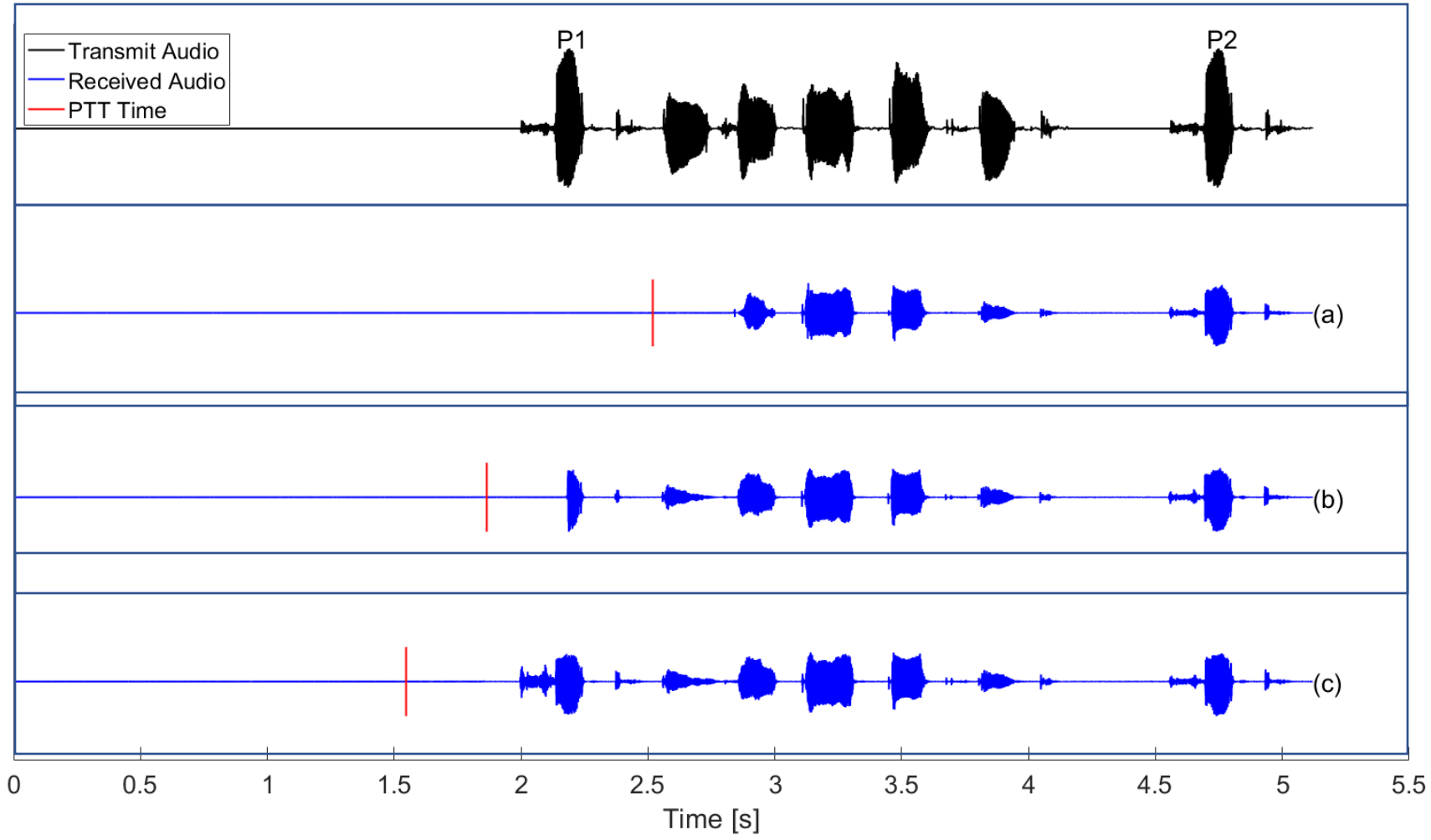
- Select single word from ABC-MRT16 database[1]
  - Use only words from batches where leading consonant varies
    - E.g. *went, sent, bent, dent, tent, rent*
    - Places majority of intelligibility emphasis on beginning of word
- Vary when PTT pressed in relation to start of this keyword
  - P1 – first play
  - P2 – second play
- P2 baselines word intelligibility
- P1 relates PTT time with intelligibility



1: Voran SD (2017) A multiple bandwidth objective speech intelligibility estimator based on articulation index band correlations and attention. 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp 5100–5104.doi: 10.1109/ICASSP.2017.7953128

ACCESS TIME MEASUREMENT CONCEPT

Intelligibility Examples: *hook*



MOTIVATION

- Partially muted words
  - An impairment ABC-MRT was not designed around
  - No human truth data to reference
- Existing audio
  - Hard to identify when MRT keyword is spoken within full MRT phrase
    - “Please select the word west”
    - “d-west”
  - Co-articulation in our audio
- Start over with new speech
  - Record carrier phrase and keywords separately

GOAL

- Design studies centered on partially muted word impairment
- Use audio with no co-articulation
  - Know exactly when a keyword starts

EXAMPLE RESULT

Batch: *thaw, law, raw, paw, jaw, saw*

Distortion: Partially muting of beginning of keyword

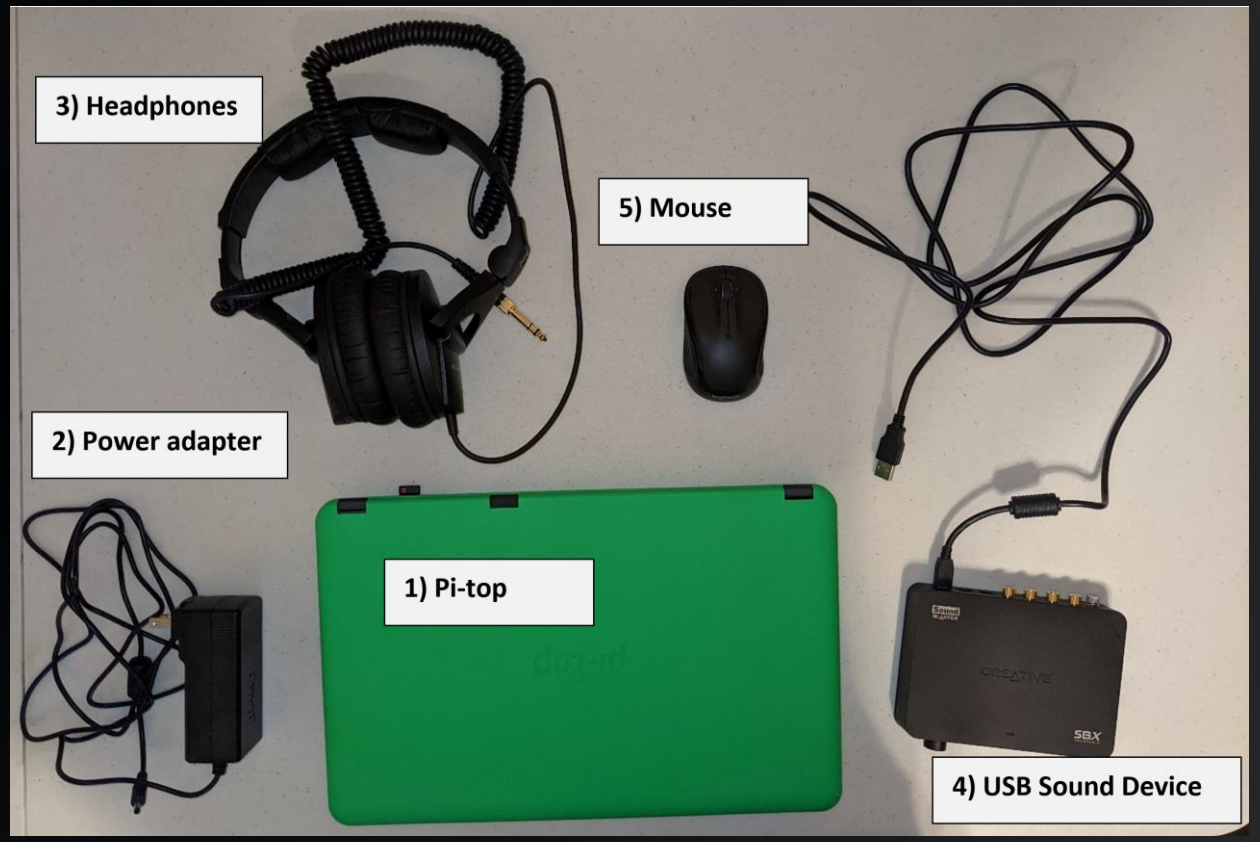
- Extreme
- Moderate
- Mild
- None

LISTENING EXPERIMENTS

- Normal controls
  - Environment ✗
  - Equipment ✓
- How to do that remotely?
  - Balance data quality vs time cost
  - Online Tool
    - Pro – Easier to run
    - Con – Variability in data quality
  - Dispersed Hardware Setup
    - Pro – Control data quality
    - Con – Time to run

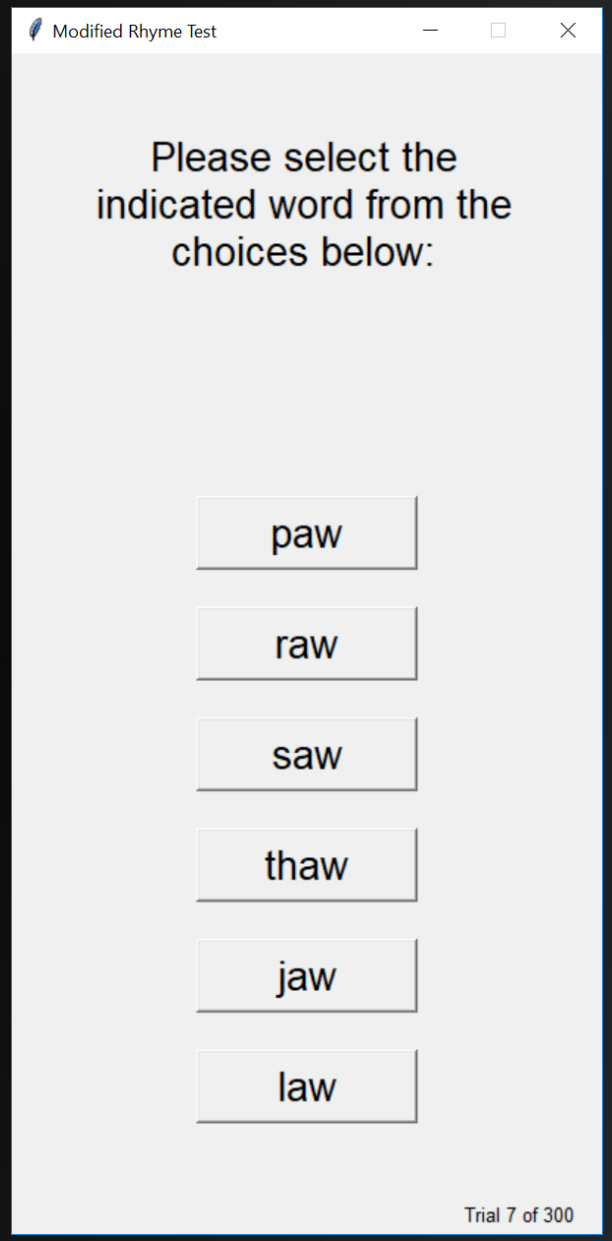
DISPERSED SETUP

- Hardware
  1) Pi-top – simple laptop enclosure for Raspberry Pi
  2) Power Adapter
  3) Headphones
  4) USB Sound Device
  5) Mouse
- Software
  - MRT GUI
  - Kiosk Design



MRT SOFTWARE

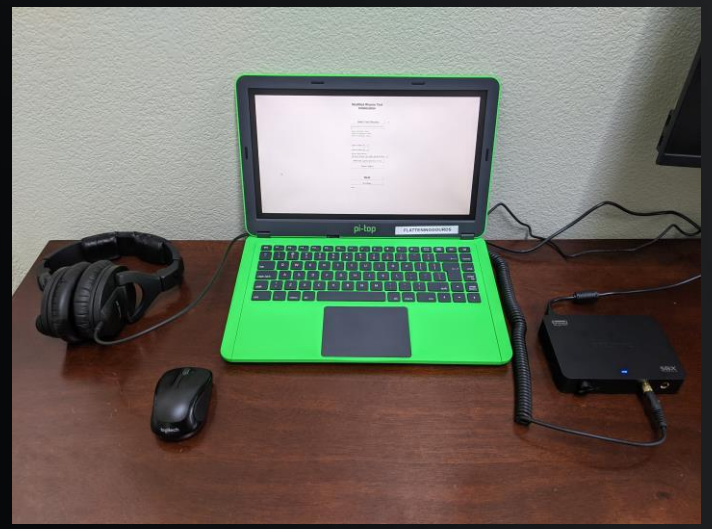
- Python MRT GUI
- In development
- Publicly available
  - https://github.com/usnistgov/mrt
  - Still in development
  - Has setup information for kiosk design



KIOSK DESIGN

- Want it to work, no matter the participant's background
- Does one thing: MRTs
  - Nothing more, nothing less
- Built on Arch Linux
  - Stripped down, easy configuration
- Uses openbox for desktop environment
  - Simple and configurable
- Locked down configuration
  - Auto start MRT GUI
    - Restart on close (test complete)
  - Force GUI on top
  - Disable keyboard shortcuts
    - Alt+F4 (window close)
    - Alt+tab (window switch)
    - Etc…
  - Disable desktop menu

SETUPS IN THE WILD



FUTURE WORK

- Three Stages of MRTs
- PTT Gate Critical Region Boundaries
- Cross-technology Critical Region Boundaries
- Critical Region Study

CONCLUSION

- Created dispersed MRT setup design
  - MRT setups useful in person as well
- Running dispersed MRTs
  - Capture data on access time specific impairments

RELATED PSCR 2021 SESSIONS

- Mission Critical Voice Quality of Experience Measurement Methods Overview
- Introducing a Start of Word Correction for Access Delay Measurements
- Measuring the Probability of Successful Delivery: a QoE Based Approach
- Optimal Transmit Volume Conditions for MCV QoE Measurement Systems
- QoE Software and Hardware Packaging
- QUARC: Quality Under Adjustable Realistic Conditions for Communication Systems
- Lab from Home: Distributed QoE Testing for Mission Critical Voice

QUESTIONS

- Looking forward to questions at our live Q&A session!

THANK YOU

#PSCR2021 • PSCR.GOV