Mission Critical Voice Speech-Based Access Time Measurement

Jaden Pieper, PSCR

#PSCR2019

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.

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

Overview

- User Experience for Push-to-talk (PTT) Communications
- Access Definitions in Other Standards
- Speech Intelligibility
- Measurement Concept
- Measurement Results
- Future Work

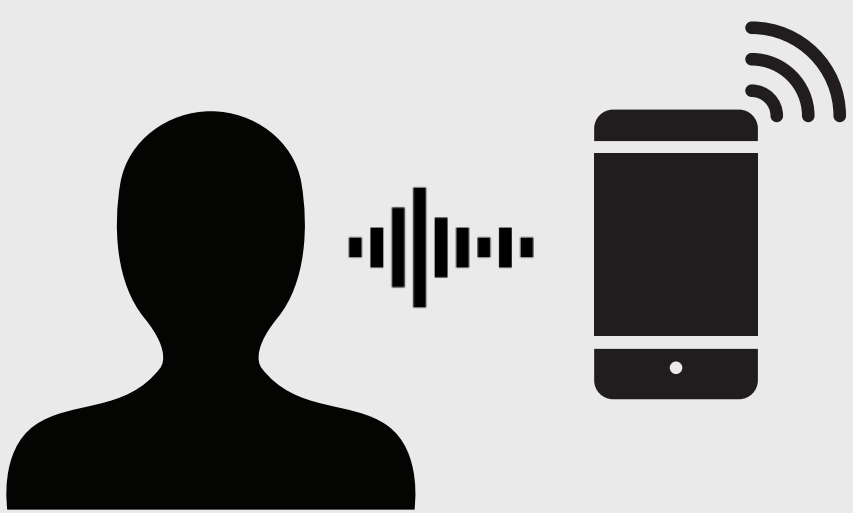
The User Experience: PTT Communications

- **Press PTT and speak into a device**
- **Listening to speech output from a device**
- **It's all about speech**

- **Goal - Create an access time measurement system that is:**
  - **Based upon the user experience -- speech**
  - **Comparable and fair across technologies**
- **This is not:**
  - **Diagnosing internal components of specific systems**

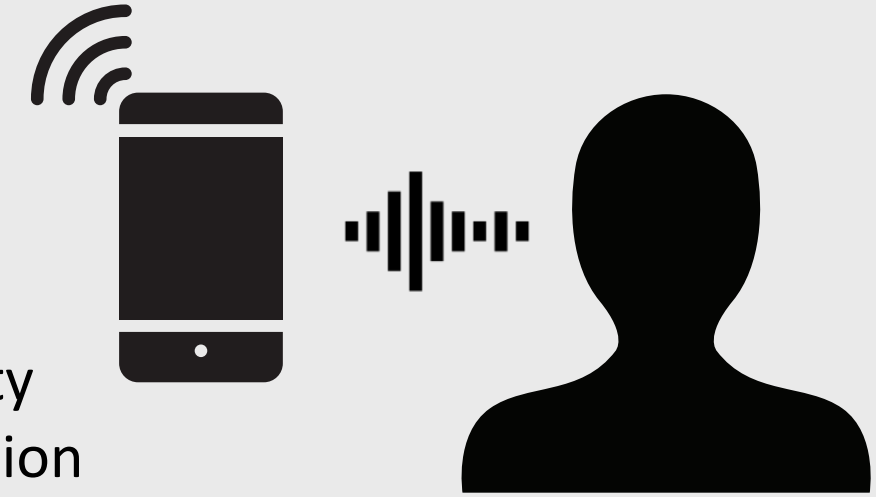
Technology Agnostic Measurements

Communications System



KPIs:

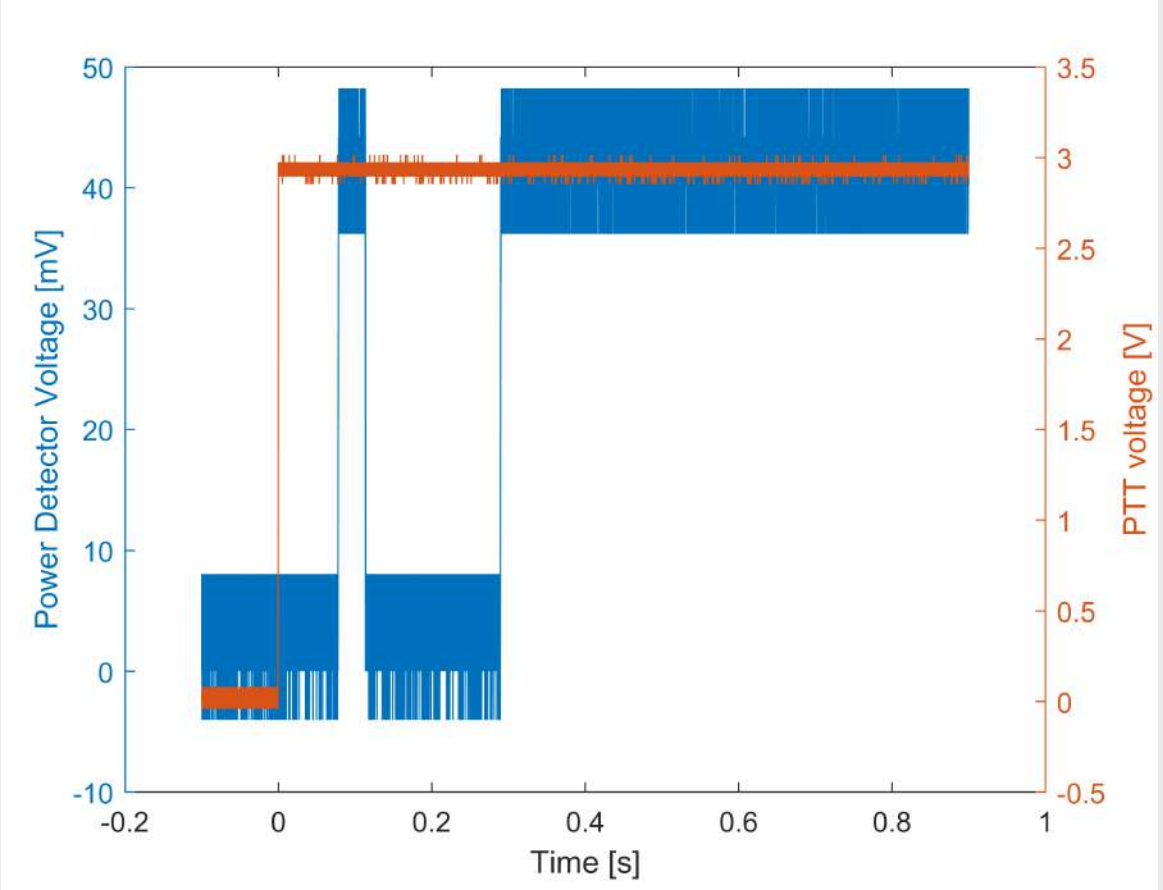
- Mouth-to-ear Latency
- End to End Access Time
- Audio Quality/Intelligibility
- Access Probability/Retention



Access Time in Other Standards

TIA-102 P25 Standard

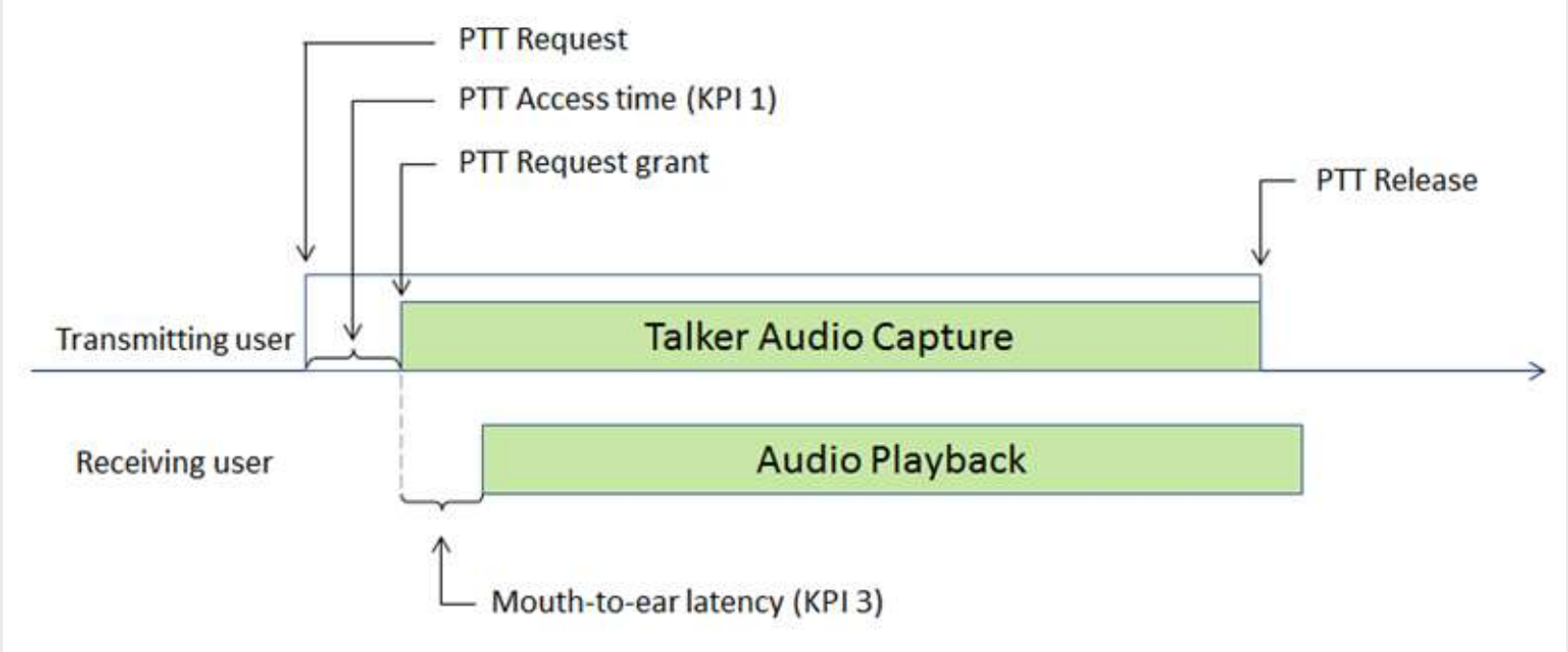
- **TIA-102 P25 Access Time**
  - **Completely focused on transmitting user**
  - **No reference to receiving user**
  - **No reference to speech**



Access Time in Other Standards

3GPP LTE PTT

- **3GPP M2E Latency and Access Time[1]**
  - Primary definition focused on transmit user
  - End-to-end **definition incorporates acknowledgement by receiving user**
  - **No guarantees on what receiving user experience is**



1: 3GPP (2017) Mission Critical Push to Talk (MCPTT). 3rd Generation Partner-ship Project (3GPP), Technical Specification (TS) 22.179. Version 16.0.0 URLhttps://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=623.

Access Time in Other Standards

Compatibility Across Technologies

- **TIA Access**
  - **Analog FM LMR**
    - **Technically compatible**
    - **Not meaningful**
  - **P25 Direct**
    - **Technically compatible**
    - **Not meaningful**
  - **LTE**
    - **Not compatible**
- **3GPP Access**
  - **Analog FM LMR**
    - **Not applicable**
  - **P25 Direct**
    - **Not applicable**
  - **P25 Trunked**
    - **End-to-end definition is compatible and meaningful**

User Driven Access Definition

- **End-to-end Access Time**
  - ***The total amount of time from when a transmitting user first presses PTT until a receiving user hears intelligible audio.***
- **Two Components:**
  - **Mouth-to-ear Latency**
    - ***The time between speech being input into one device and its output through another***
  - **Access Delay**
    - ***The minimum length of time a user must wait between pressing a PTT button and starting to speak to ensure that the start of the message is not lost***

Formalizing Access Delay

- **Access Delay**
  - **All about if a message is lost or not**
  - **Intelligibility is the key to the measurement**
- **Formal Definition:**
  - ***The minimum length of time a user must wait between pressing a PTT button and starting to speak to ensure that the first word of the message has an average intelligibility that is no lower than*** $\alpha \cdot I_0$
  - $0 < \alpha < 1$, **defines acceptable intelligibility level**
  - I_0 **is the baseline intelligibility of that word through the communications system**
    - **No system is perfectly intelligible**
    - **Some level of degradation almost always present**

Intelligibility

Modified Rhyme Test (MRT)

- **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**
    - **e.g. "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

Intended 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)

- **Articulation Band Correlated Modified Rhyme Test (ABC-MRT)**
  - **Objective algorithm to provide estimates of true MRT scores**
  - **Developed by Stephen Voran and DJ Atkinson, ITS-NTIA**
  - **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**

Measurement Concept

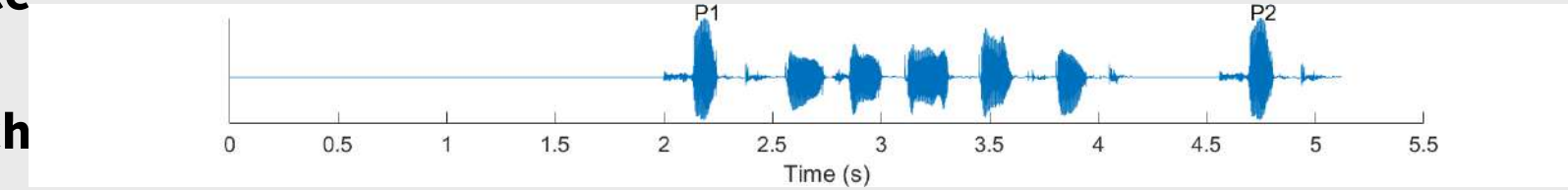
Overview

- **Access Delay definition:**
  - ***The minimum length of time a user must wait between pressing a PTT button and starting to speak to ensure that the first word of the message has an average intelligibility that is no lower than*** $\alpha \cdot I_0$
- **Repeatedly send pre-defined audio clips through communications system**
- **Vary where in the clip PTT is triggered**
- **Measure relationship between PTT time and intelligibility of the first word in the clip**
  - **No more carrier phrase**

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**
- **Structure:**
  - **T seconds of silence**
  - **Play word, P_1**
  - **T seconds of speech**
  - **Play word again, P_2**

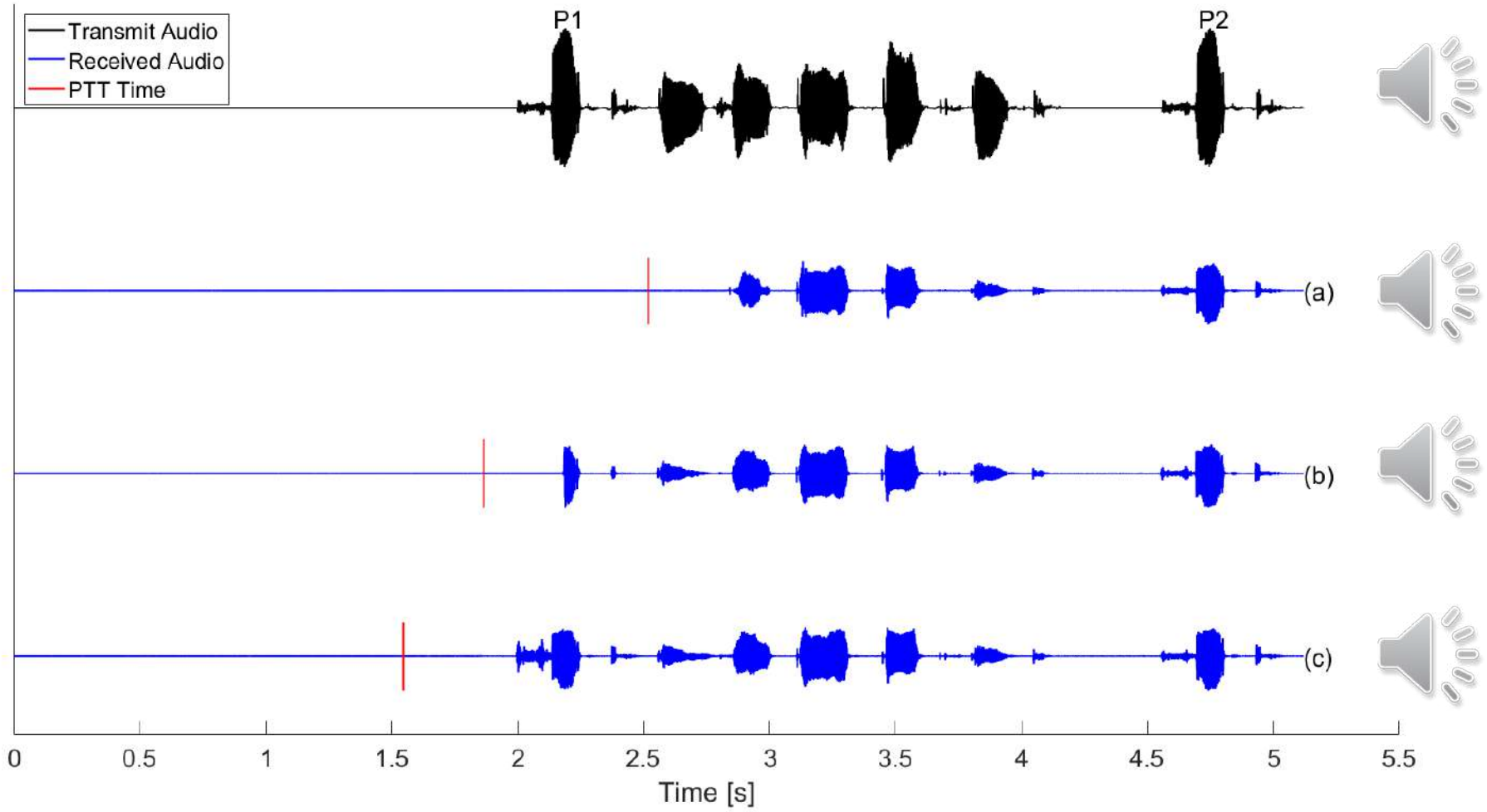


- **T chosen so that system access time is less than T seconds**
- **Intelligibility of P_2 describes the asymptotic intelligibility, I_0**
- **Intelligibility of P_1 relates PTT time with intelligibility**

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

Measurement Concept

Intelligibility Examples: *hook*



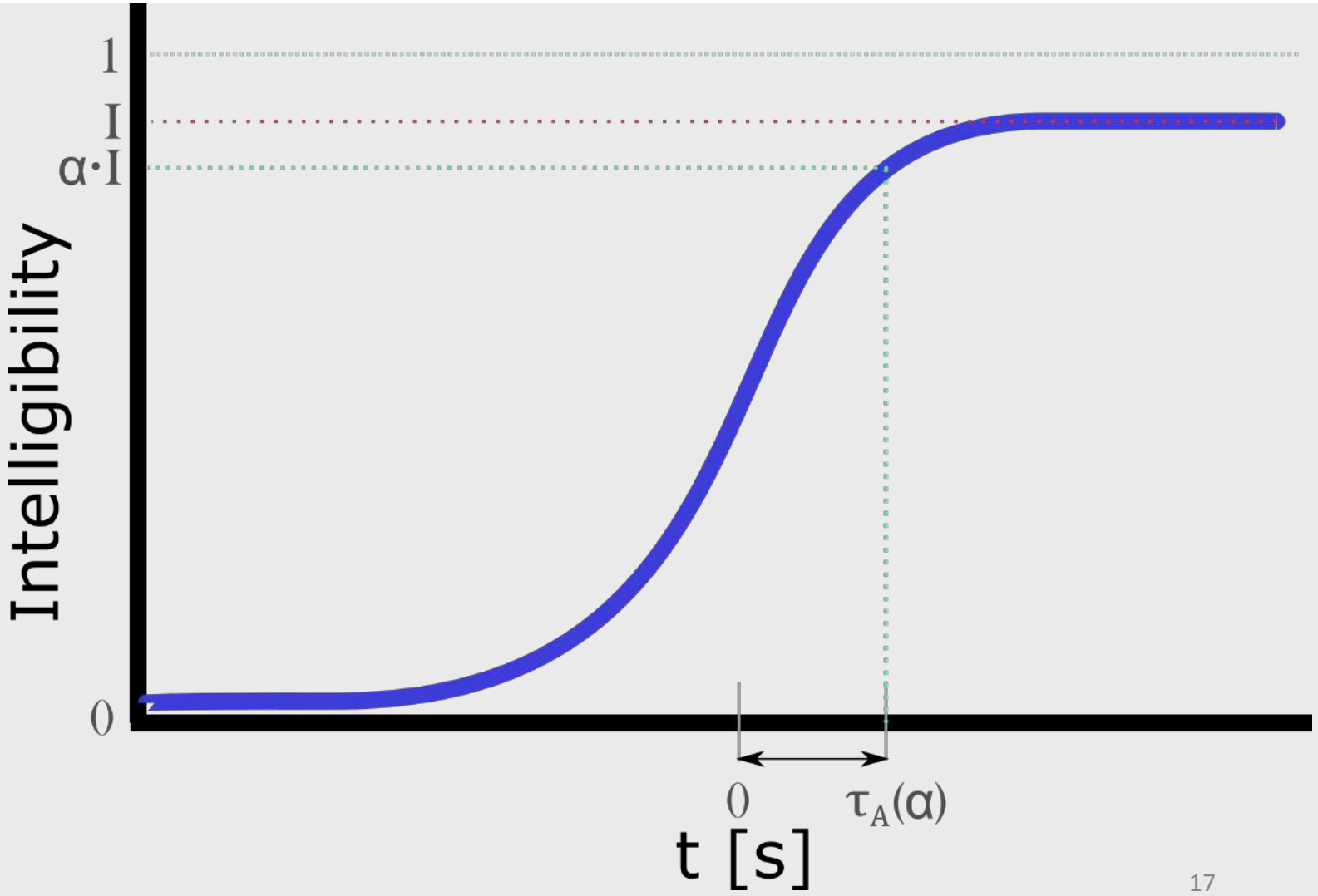
Measurement Concept

Intelligibility Curve

- **Fit a curve to data**
- **Logistic curve has properties we want**

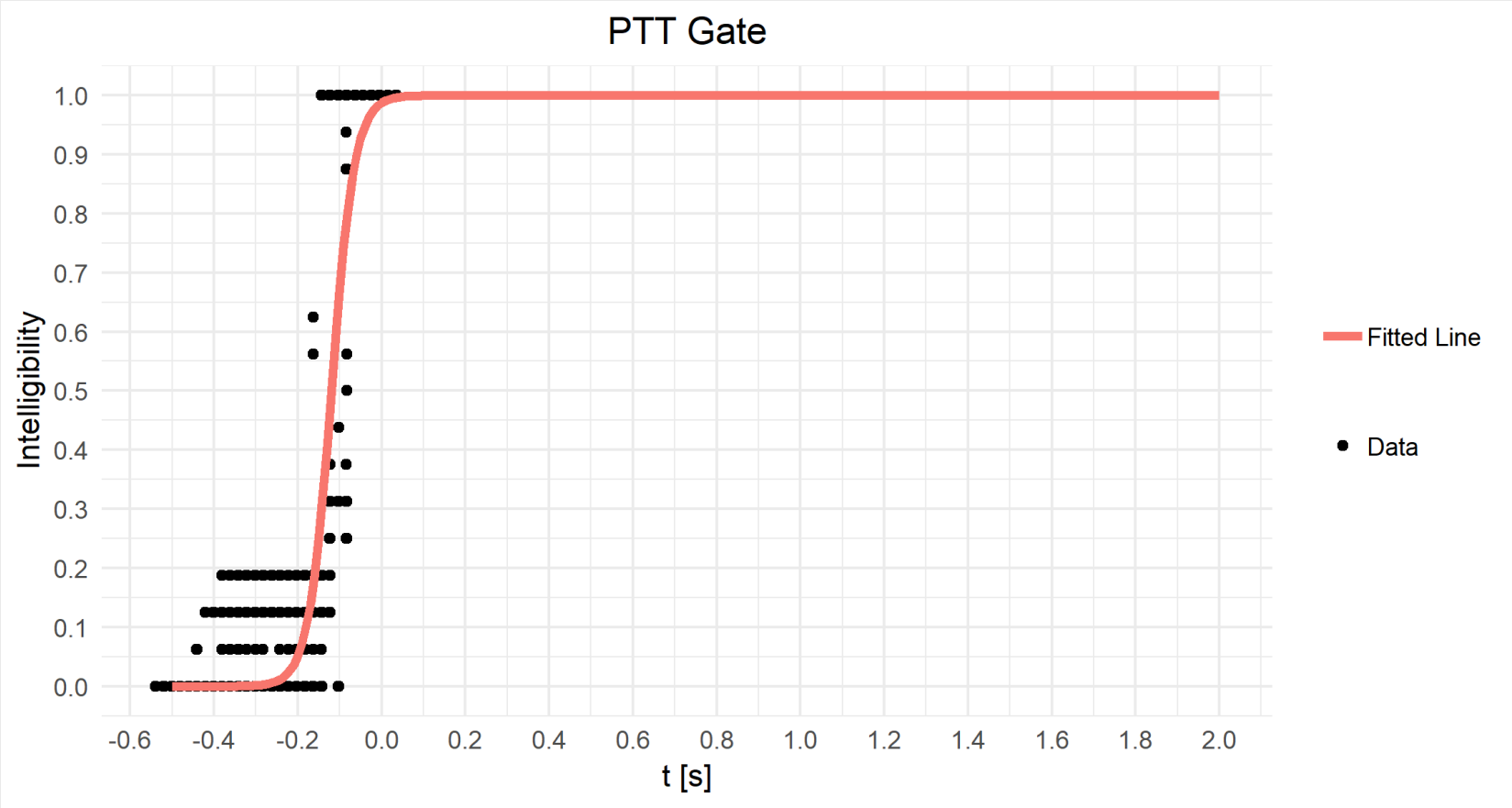
$$I(t) = \frac{I_0}{1 + e^{(t - t_0)/\lambda}}$$

- λ: **Steepness of intelligibility transition**
- t_0: **50% intelligibility point**

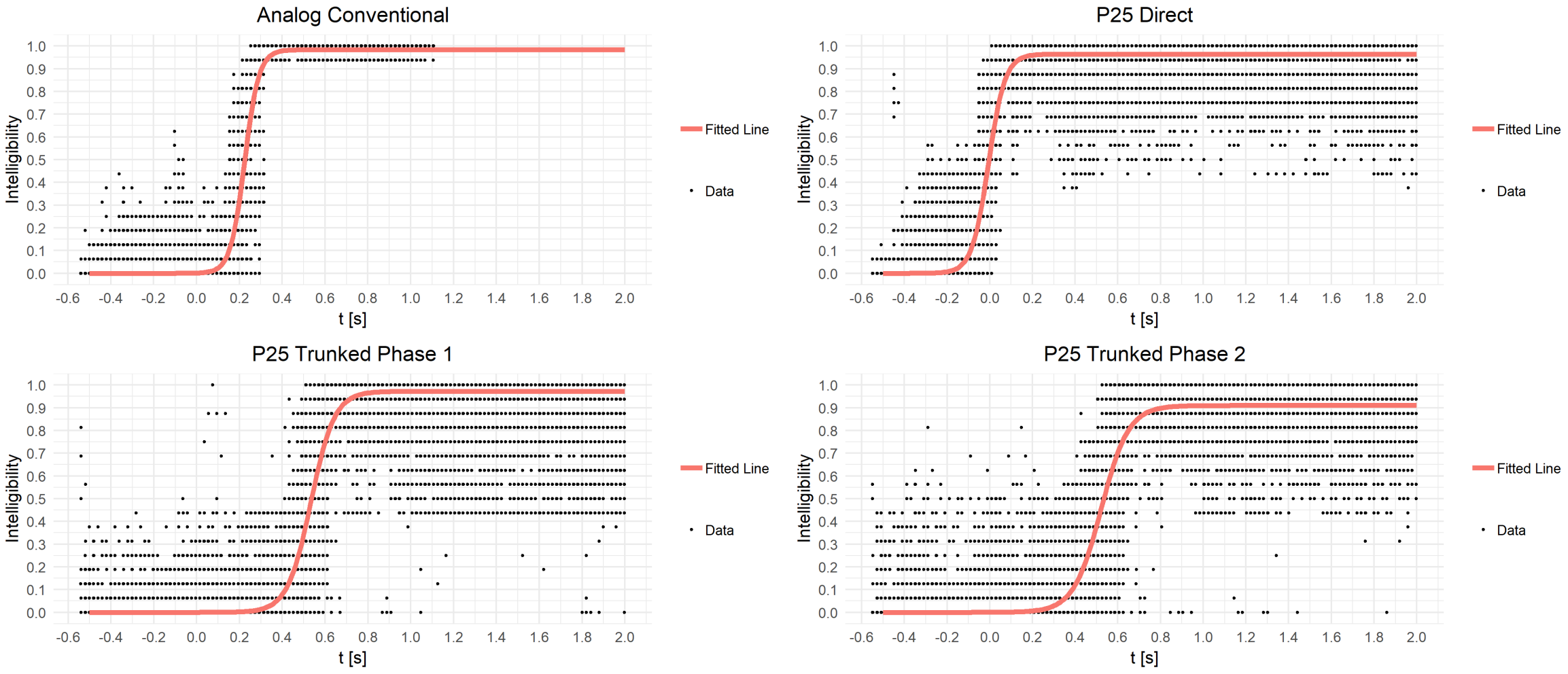


Measurement Concept

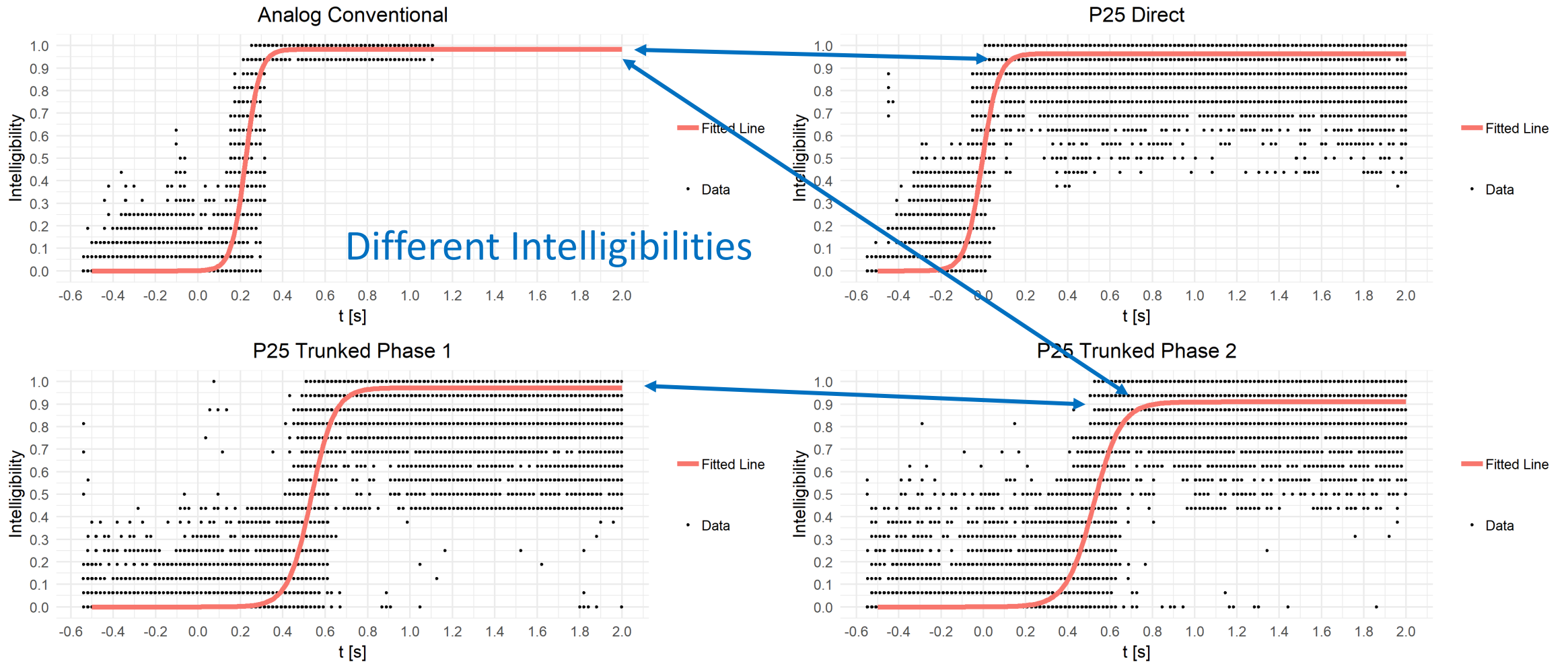
Simple Radio Replacement



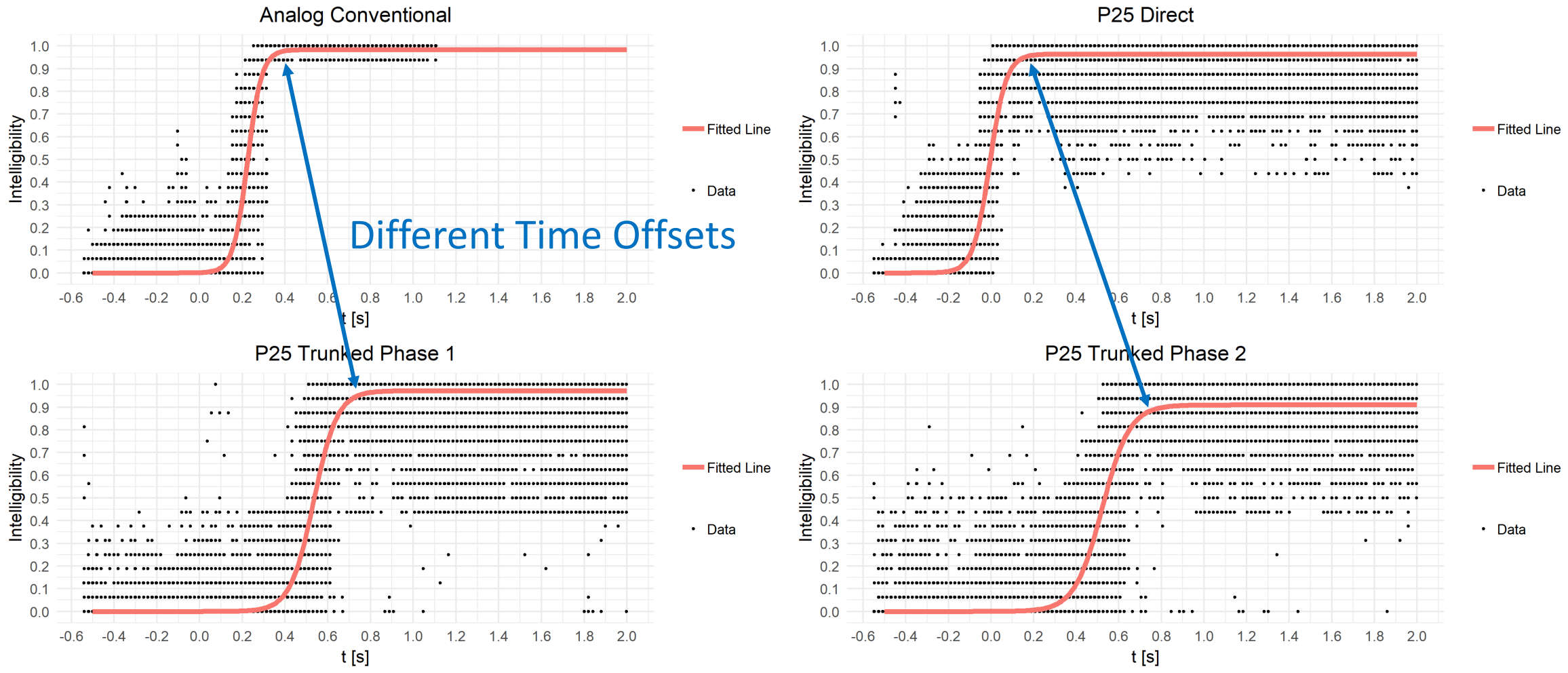
Intelligibility Curves for PTT Technologies



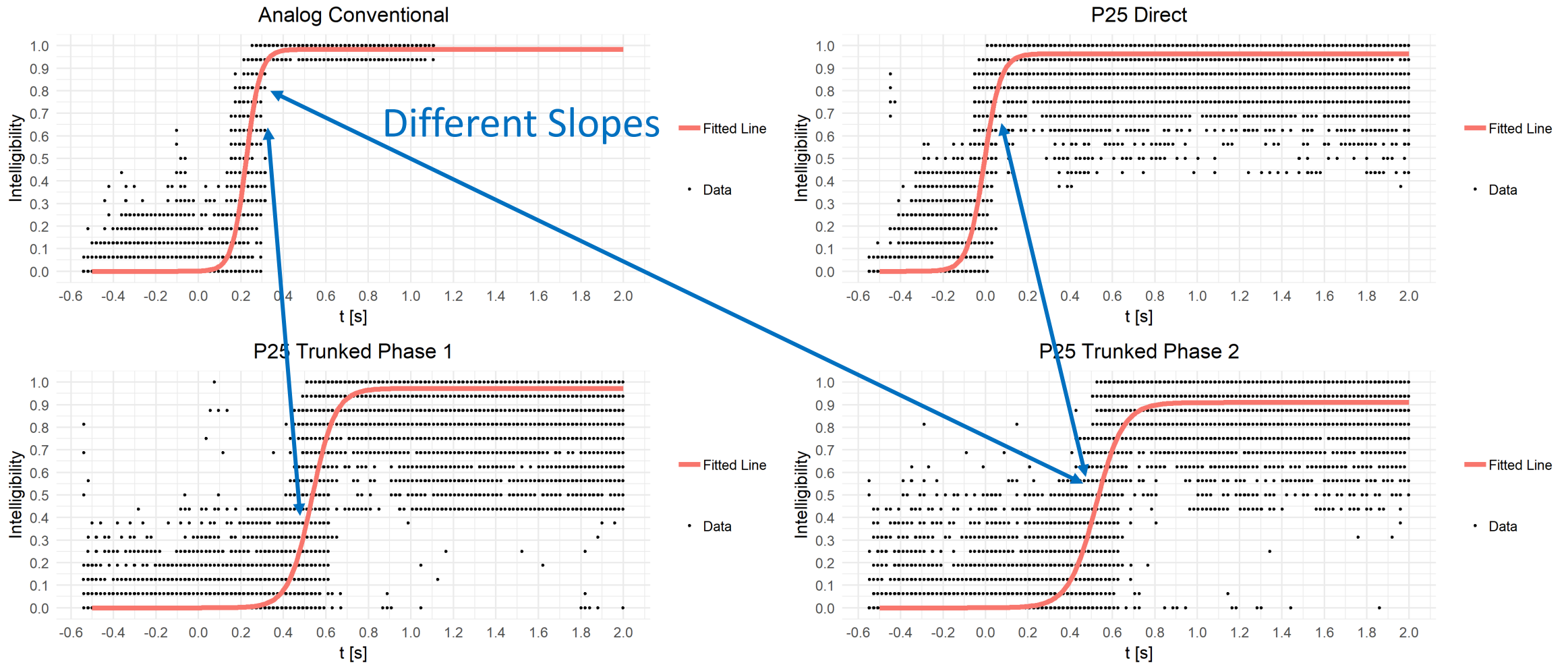
Intelligibility Curves for PTT Technologies



Intelligibility Curves for PTT Technologies



Intelligibility Curves for PTT Technologies



Measurement Concept

Access Delay Function

$$I(t) = \frac{I_0}{1 + e^{(t-t_0)/\lambda}}$$

Given $0 < \alpha < 1$, **an intelligibility of** $\alpha \cdot I_0$ **can be achieved with** $t = I^{-1}(\alpha \cdot I_0)$.

Access Delay defined as:

$$\tau_A(\alpha) = \lambda \cdot \ln\left(\frac{1-\alpha}{\alpha}\right) + t_0$$

Measurement Concept

Uncertainty of Access Delay

For an access delay estimate, $\hat{t}$, for some choice of α:

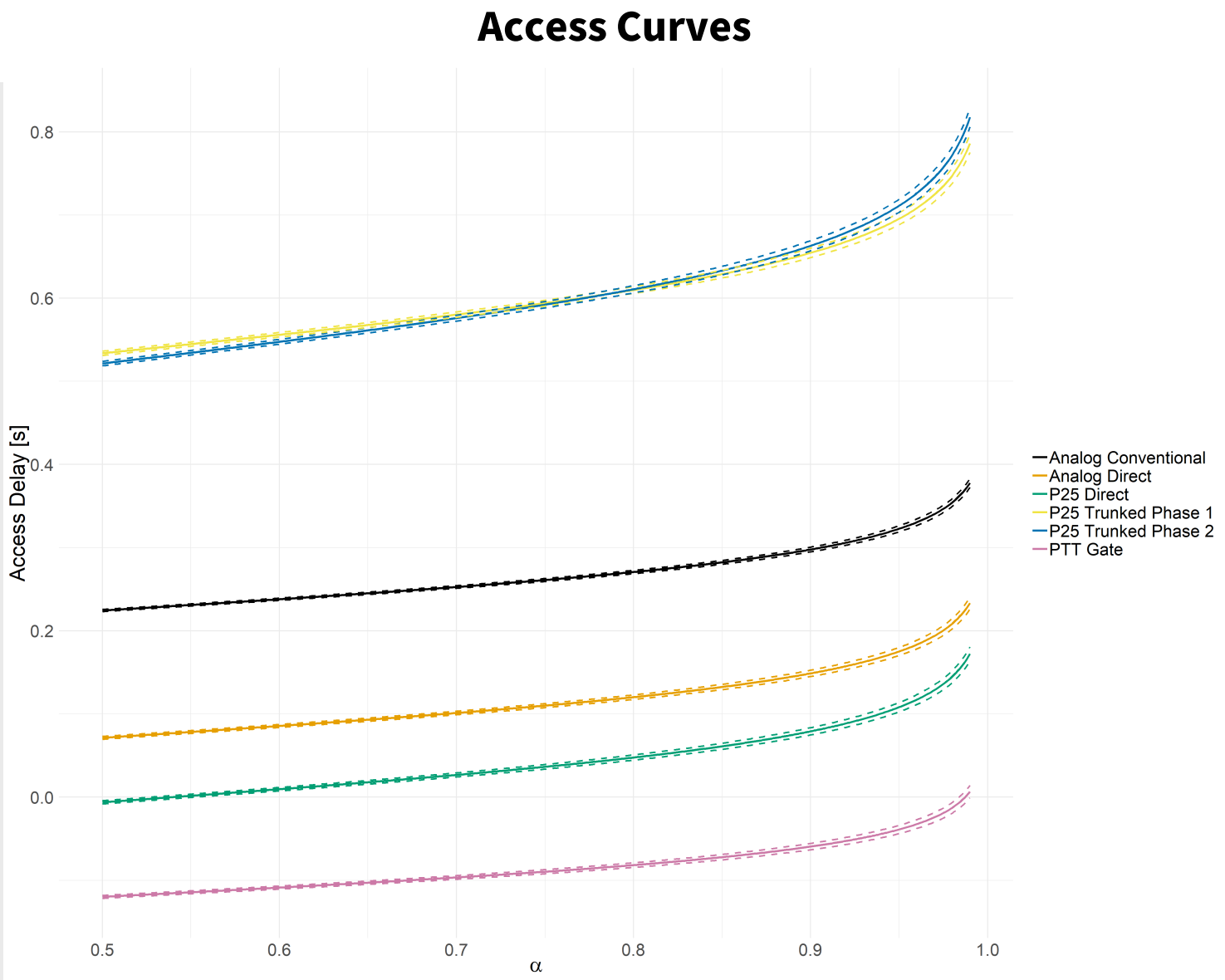
$$\hat{t} = \hat{\lambda} \cdot \ln\left(\frac{1-\alpha}{\alpha}\right) + \widehat{t_0}$$

Let $C = \ln\left(\frac{1-\alpha}{\alpha}\right)$

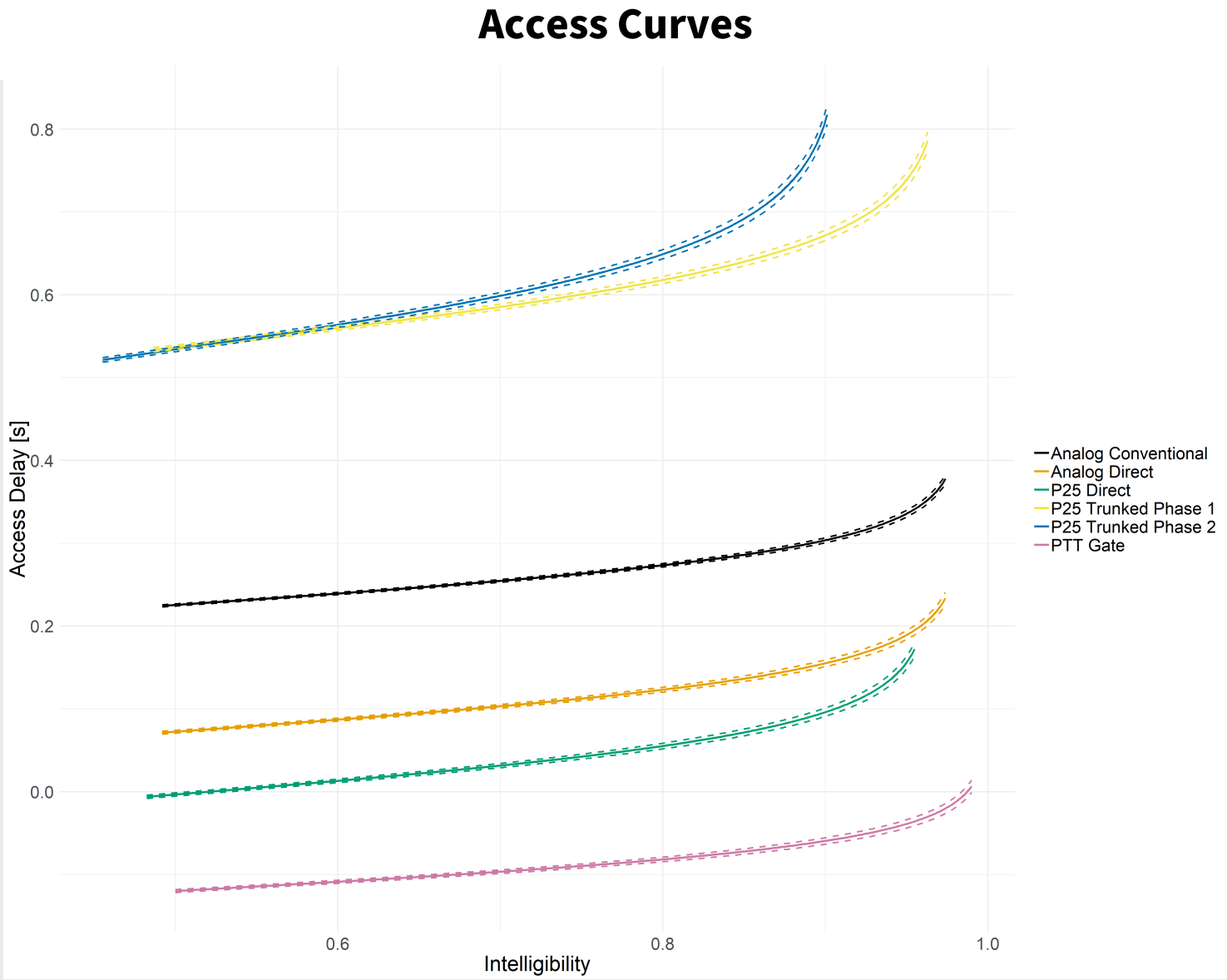
Uncertainty from variance of estimate:

$$\mathrm{Var}(\hat{t}) = C^2\mathrm{Var}\left(\hat{\lambda}\right) + \mathrm{Var}(\widehat{t_0}) + 2C \cdot \mathrm{Cov}(\hat{\lambda}, \widehat{t_0})$$

Measurement Concept



Measurement Concept



End-to-end Access Time Results for 85% Intelligibility

PTT Technology*	M2E Latency (ms)	Access Delay (ms)	End-to-End Access Time (ms)
Analog Direct	76.5 ± 0.3	136.5 ± 3.3	213.1 ± 3.3
Analog Conventional	78.5 ± 0.3	286.1 ± 2.5	364.7 ± 2.5
P25 Direct	220.9 ± 0.3	71.6 ± 4.1	292.4 ± 4.1
P25 Trunked (Phase 1 – FDMA)	356.6 ± 3.8	640.1 ± 5.1	996.7 ± 6.3
P25 Trunked (Phase 2 – TDMA)	575.9 ± 8.1	692.2 ± 7.1	1268.1 ± 10.7

*Analog Conventional operates in VHF band (30 MHz to 300 MHz).
All P25 technologies operating in 700 MHz band.

Please note all information and data presented is preliminary/in-progress and subject to change.

Future Work

Limitations of the Measurement

- **Aggregate Measure**
  - **Only get an access delay value after extensive testing/data collection**
  - **Cost of measuring the whole curve is very high**
- **ABC-MRT never tested under these impairments**
  - **No guarantee that the MRT estimates are accurate for partially muted words**

Future Work

Improving the Measurement System

- Collect truth data
  - Perform MRT tests with real people
  - Focus on specific impairment of access delay problem
    - E.g. receiving portions of words
- Compute accuracy of ABC-MRT16 access delay system
- Develop specialized intelligibility estimator
  - Focused on partially muted words
- Develop algorithm to differentiate intelligibility on a more fine tuned scale than whole words
  - One-to-one measurement
  - Requires fine time resolution

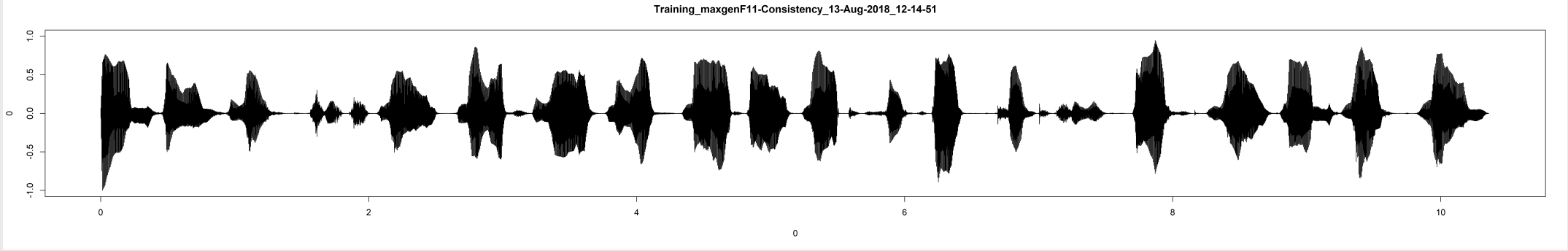
Future Work

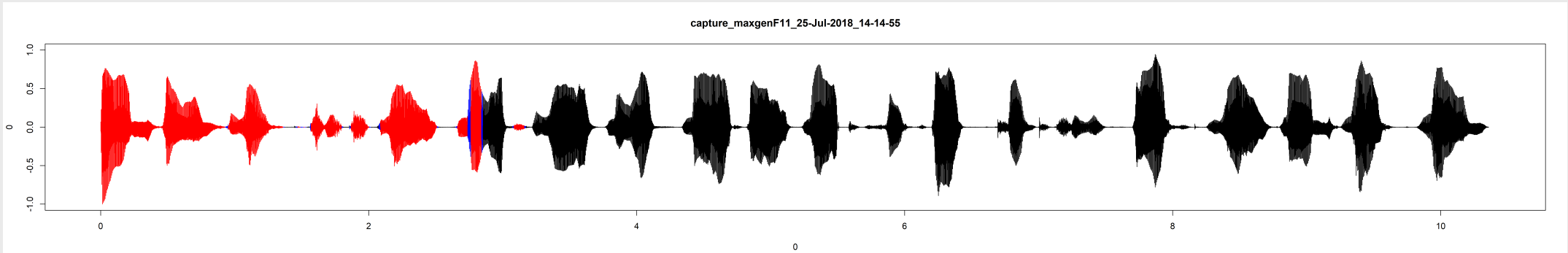
MFCC Intelligibility

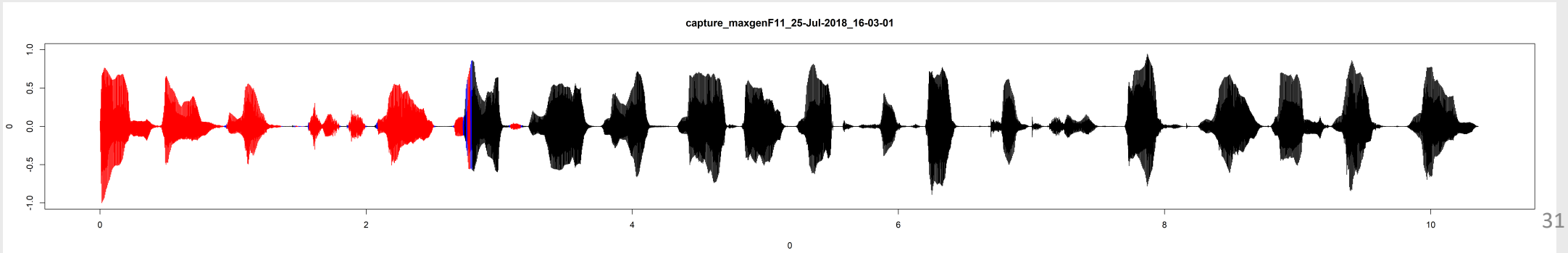
- Mel Frequency Cepstral Coefficients (MFCCs)
  - Turns speech into a musical score
  - Related to the perceived loudness in the human ear of different frequencies
  - Can we make intelligibility distinctions on a smaller time scale?
    - e.g. within a word
- Train a model with intelligible and unintelligible speech
  - Labeled by human MRT trials
  - Tune to the structure of intelligible speech
  - Example model: Multivariate Normal Distribution
- One-to-one measurement
  - Each trial yields an end-to-end access time value

Future Work

Small timescale intelligibility decisions

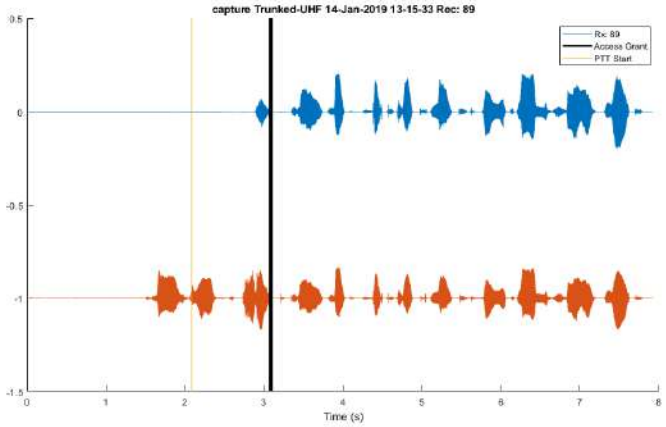


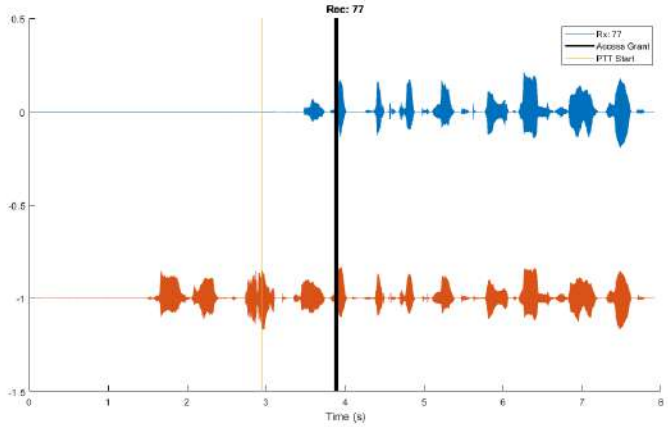


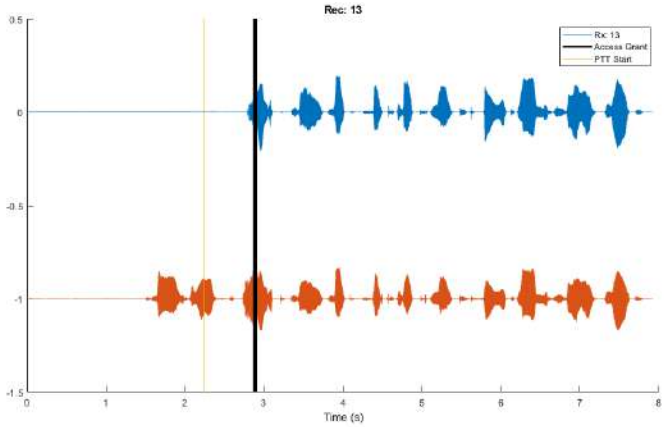


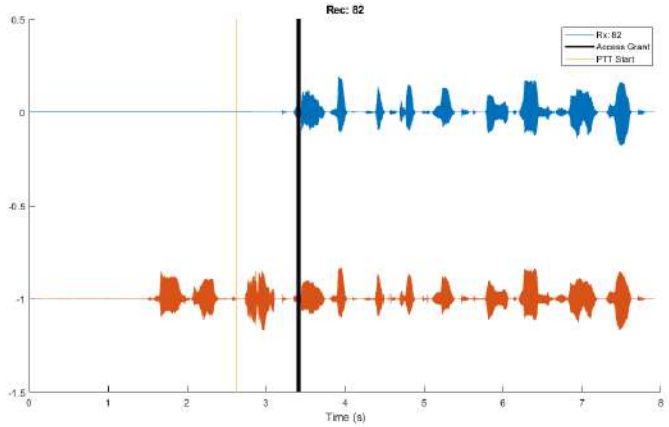
Future Work

MFCC Intelligibility









Team Members

- **Back Row**
  - **Steve Voran**
  - **Tim Thompson**
  - **Jesse Frey**
  - **Zainab Soetan**
- **Front Row**
  - **Hossein Zarrini**
  - **Don Bradshaw**
  - **Chelsea Greene**
  - **Jaden Pieper**


THANK YOU

#PSCR2019

Come back for the

Next Session

2:40 PM