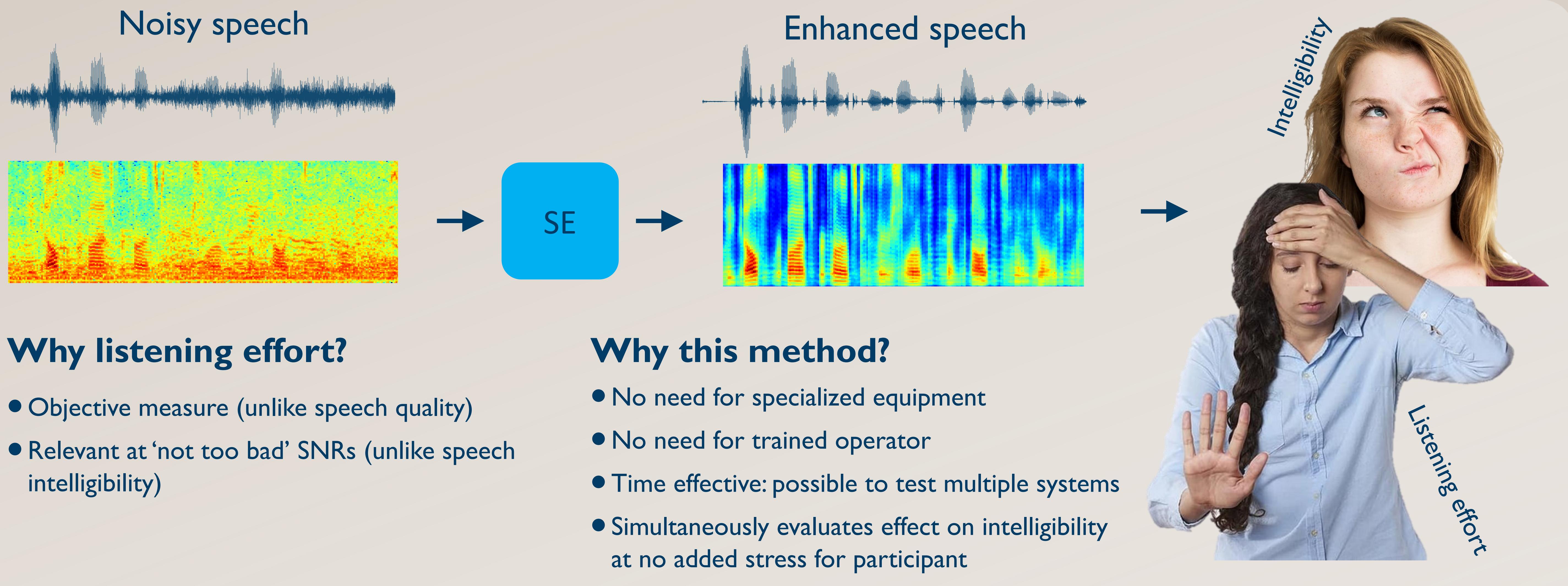


Evaluating Speech Enhancement: Listening Effort

F. B. Gelderblom, T.V. Tronstad and I. López-Espejo

Proposing a simple method to evaluate the impact of an SE system on listening effort

Motivation



Why listening effort?

- Objective measure (unlike speech quality)
- Relevant at 'not too bad' SNRs (unlike speech intelligibility)

Why this method?

- No need for specialized equipment
- No need for trained operator
- Time effective: possible to test multiple systems
- Simultaneously evaluates effect on intelligibility at no added stress for participant

Experiment

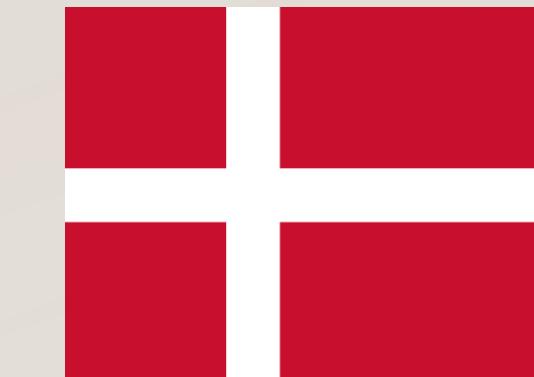
Common setup

- Single task: speech *intelligibility*
- Matrix test with 5 words × 10 options: *name, verb, numeral, adjective, object*
- Limited speech material required, yet 100 000 unique sentences
- Listener quickly becomes familiar with the material (training effect disappears)
- Suitable for testing different processing conditions
 - Cognitive challenging sentences
 - **Secretly track time to first click**
 - **Discard incorrect clicks**



Norwegian test

- 50 office workers
- 25 male, 25 female
- Ages: 26 - 72
- Includes native, non-native, and self-reported normal hearing/hearing loss
- 6 processing conditions
- Tests adapt SNR to responses



Danish test

- 26 native speakers
- 19 male, 7 female
- Ages 18 - 30
- No self-reported hearing loss (one slight bilateral tinnitus)
- 3 processing conditions
- Tested at fixed SNRs: -5 dB and -10 dB

Results

Conclusions

We were able to obtain a measure of listening effort that allowed us to compare different processing conditions.

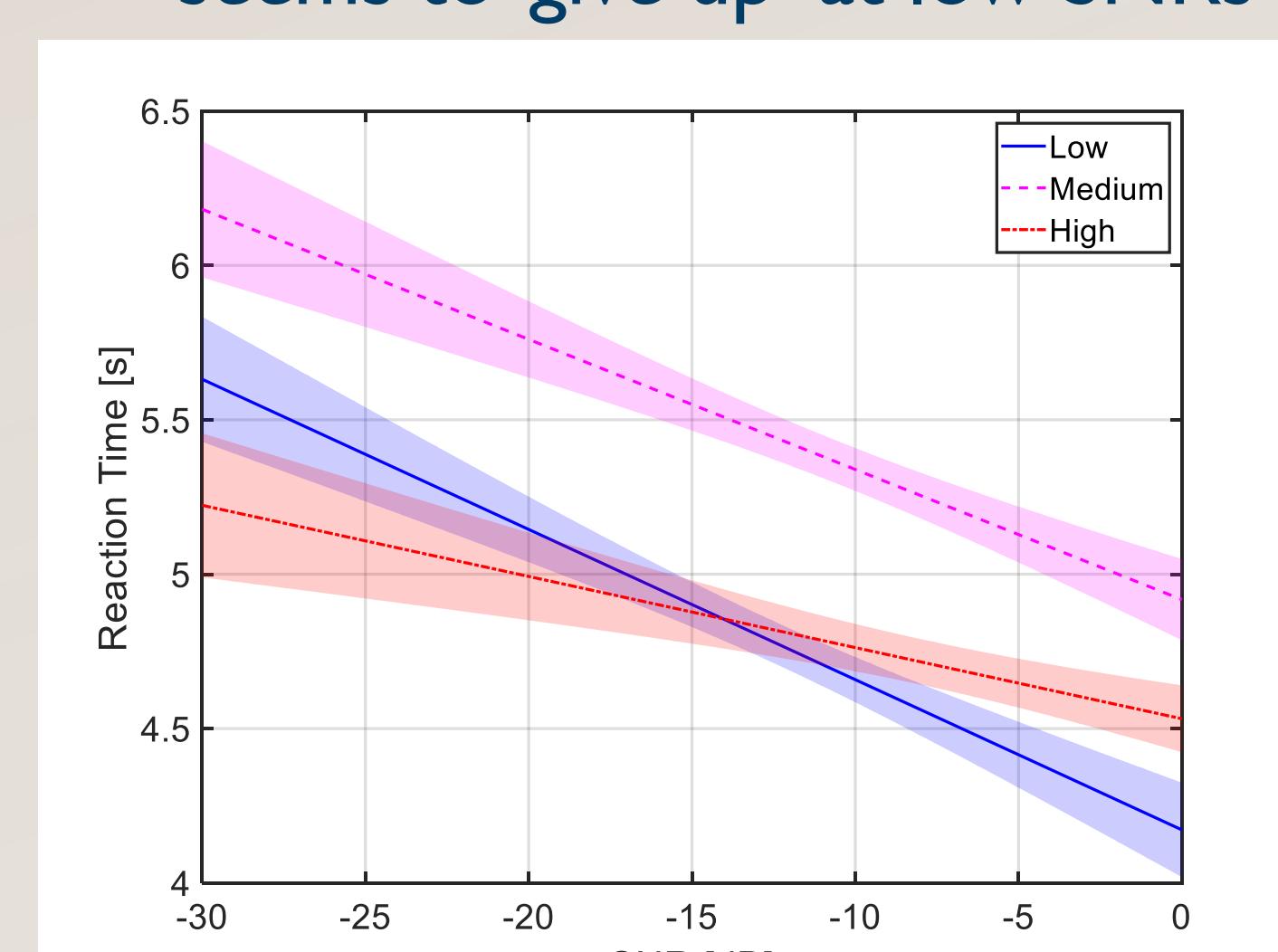
The tested processing conditions did not affect the measured LE, but as expected SNR and language familiarity/hearing ability did.

Our paper:



Norwegian test

- SNR significant effect on LE
- Listening group significant effect on LE
- Processing models NO significant effect on LE
- High SRT group (non-native/hearing loss) seems to 'give up' at low SNRs



Danish test

- SNR significant effect on LE
- Processing model only once significant effect on LE

Table 1: reaction times in seconds

Processing	SNR -10 dB	SNR -5 dB
Unprocessed	6.02	5.66
Unmatched	6.18	5.90
Matched	6.13	5.62

Table 2: p-values for comparison between -10 dB and -5 dB SNRs

Processing	p-value
Unprocessed	0.035
Unmatched	0.019
Matched	<0.001