The Prediction of Fatigue Using Speech as a Biosignal

Khan Baykaner, Mark Huckvale, Iya Whiteley

University College London

Oleg Ryumin, Svetlana Andreeva

Gagarin Cosmonaut Training Centre

iVOICE

integrated VOIce analysis of satellite Communications Embedded in time & safety-critical environment

ARTES (Advanced Research in Telecommunications Systems)

Fatigue in Safety-Critical Environments

- **Driver fatigue said to be involved in 15-20% of all transportation accidents**
- **Critical problem in aerospace**
  - pilots, air-traffic controllers
  - passenger safety
- **Critical problem in mining**
  - drivers
  - lost revenue from accidents

Assessing Fatigue from Biosignals

- Equipment operation
  - e.g. steering of vehicle
- Computer-vision
  - e.g. head movement, eye blinks
- Electro-encephalography
  - e.g. visual cortex activity
- Electro-cardiology
  - e.g. heart rate
- Electro-oculography
  - e.g. eye muscle movements
- Pulse oximetry
  - e.g. blood oxygenation

Predicting Fatigue from Speech Recordings

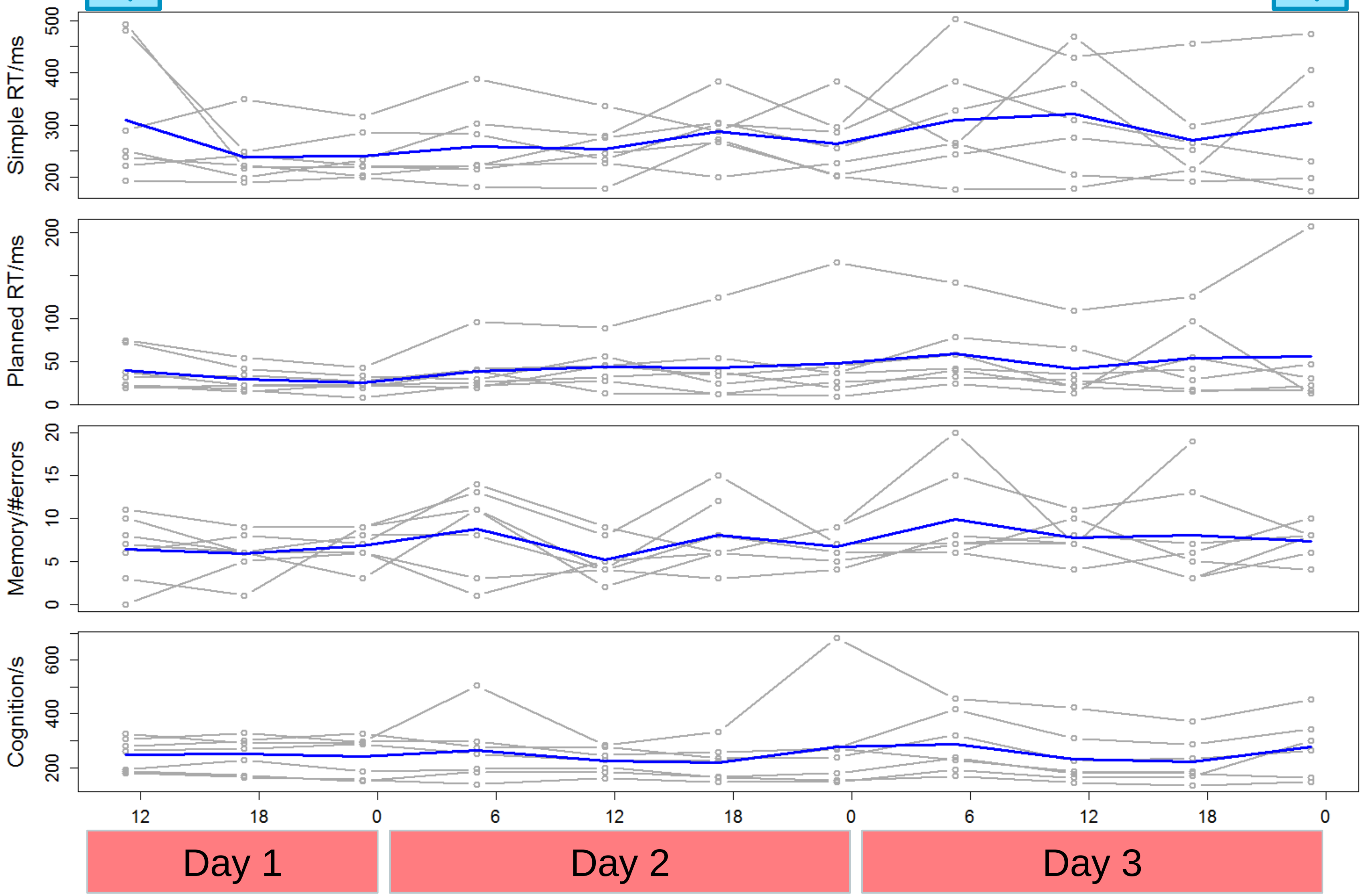
- **To obtain cheap, non-obtrusive means for assessing fatigue**
  - May be used in combination with other methods
- **Most published work uses self-reporting of "sleepiness"**
  - Karolinska Sleepiness Scale
- **But KSS ratings only have average correlations with behavioural measures of fatigue**
  - r ~ 0.57 (Kaida et al., 2006), r ~ 0.49-0.71 (Gillberg et al., 1994), no significant correlation (Åhsberg et al., 2000)
- **Our goal was to collect speech recordings labelled with objective measures of fatigue**
  - Time spent awake
  - Performance on psycho-physiological tests

Data collection on fatigue and speech

- **Isolation experiment in collaboration with GCTC**
- **7 aerospace trainees**
- **In chamber and awake 60hr**
  - morning day 1
  - evening day 3
- **Physiological/psychological tests every 6 hours**
  - reaction time
  - memory
  - cognitive load
- **Read 3min passage from novel every 6 hours**

Example Isolation Chamber

Test scores over 60hr awake



Speech signal feature extraction

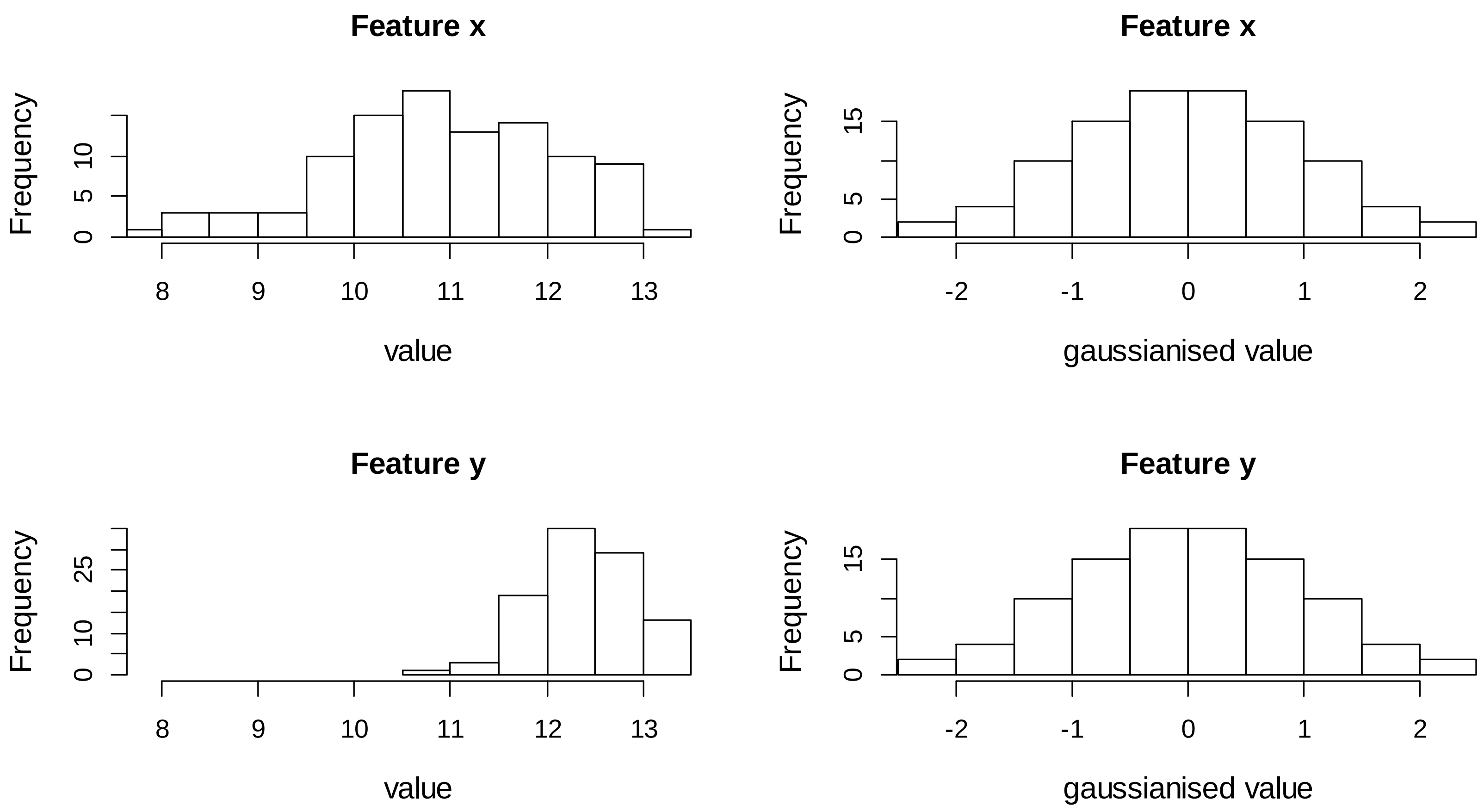
- **iVOICE Feature Analysis**
  - designed to generate features robust to added noise
  - C++ implementation
- **Temporal domain analysis**
  - features derived from autocorrelation function
- **Spectral domain analysis**
  - features derived from spectrum
- **Modulation domain analysis**
  - features derived from modulation spectrum
- **Statistical functionals**
  - percentiles, dispersion, robust skewness, robust kurtosis
- **Generates 1100 features per recording**

Machine Learning

- **Classification Task**
  - Divide time awake into day 1 vs days 2 & 3
  - 24hour threshold or 16hour threshold
  - Support Vector Machine, linear kernel
  - Leave-one out cross-validation
- **Regression**
  - Predict time awake from speech features
  - Predict test scores from speech features
  - Support Vector Regression
  - 10-fold cross-validation

Speaker-dependent Feature Normalisation

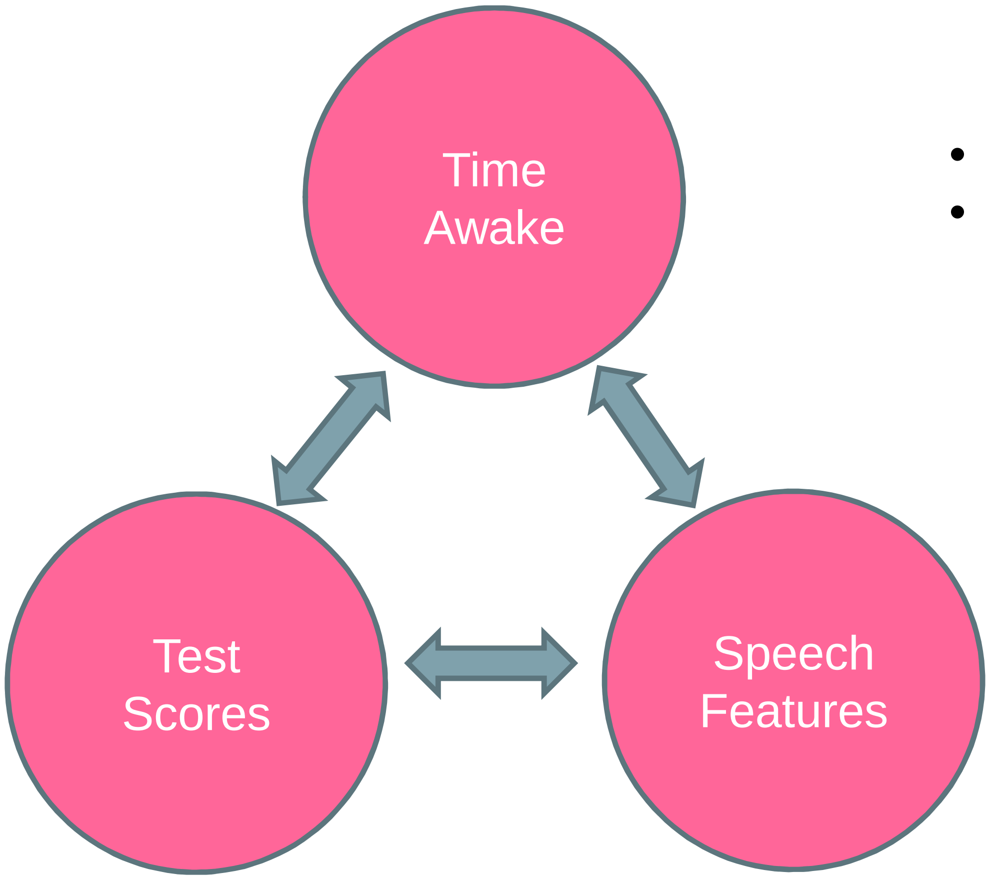
Gaussianization



Summary of Classification Results

Configuration	Unweighted Accuracy
24hr threshold, raw features	82.2%
24hr threshold, Gaussianized features	86.1%
16hr threshold, raw features	82.6%
16hr threshold, Gaussianized features	93.9%
24hr threshold, split validation, Gaussianized features	93.8%

Building Predictive Models

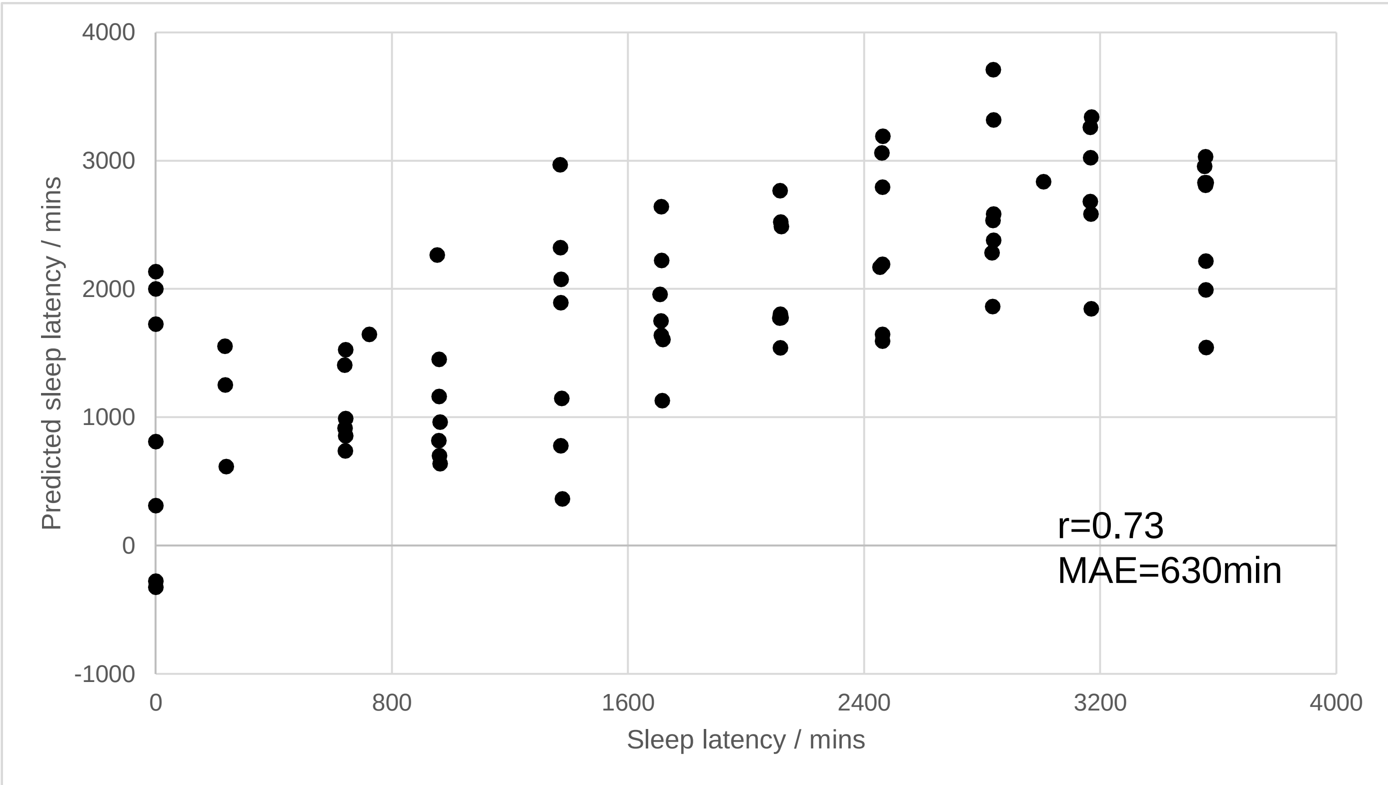


- Multiple Linear Regression
- Support Vector Regression

Prediction of Time Awake from Speech

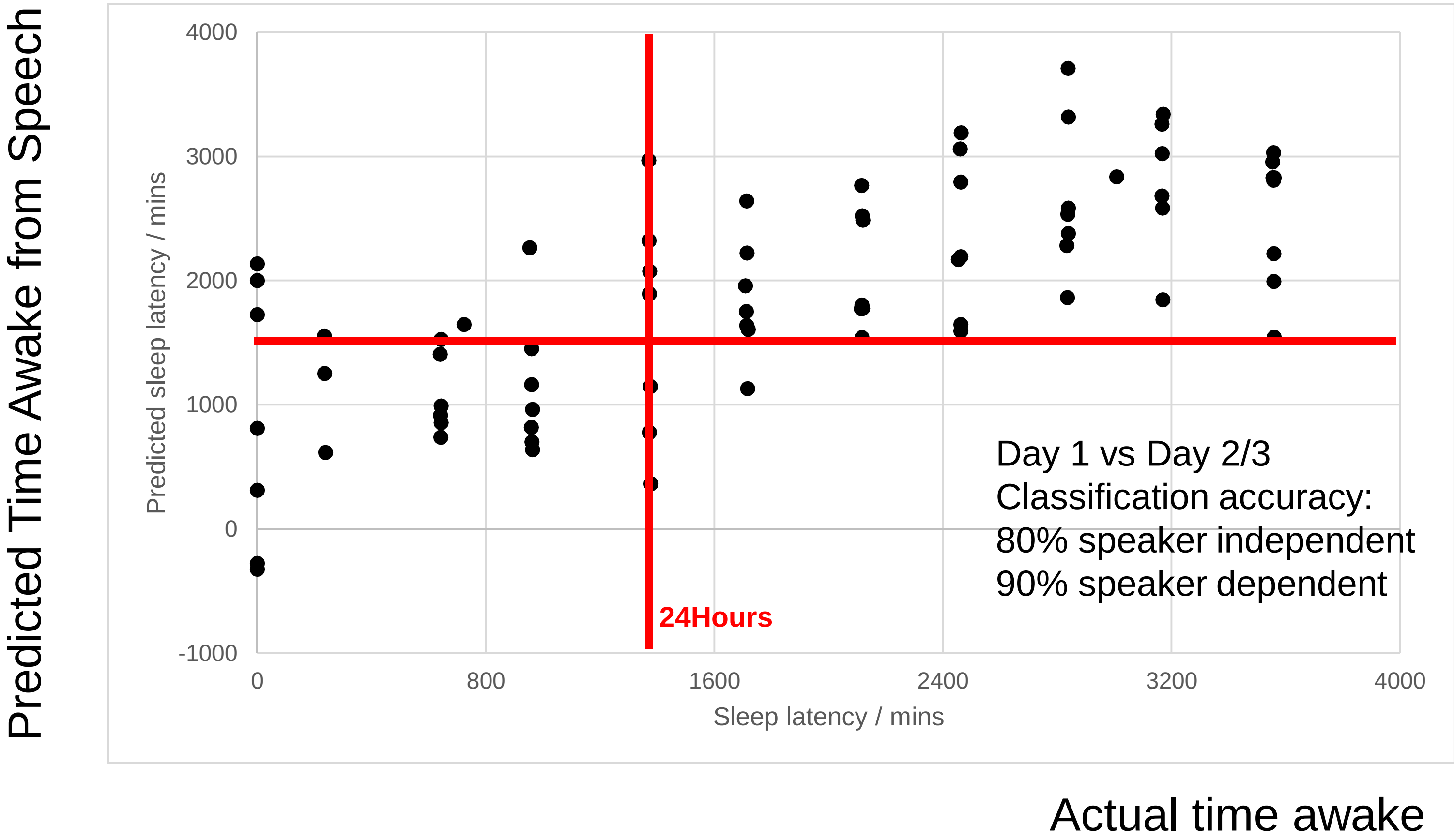
SVR, Gaussianized features, 10-fold CV

Predicted Time Awake from Speech

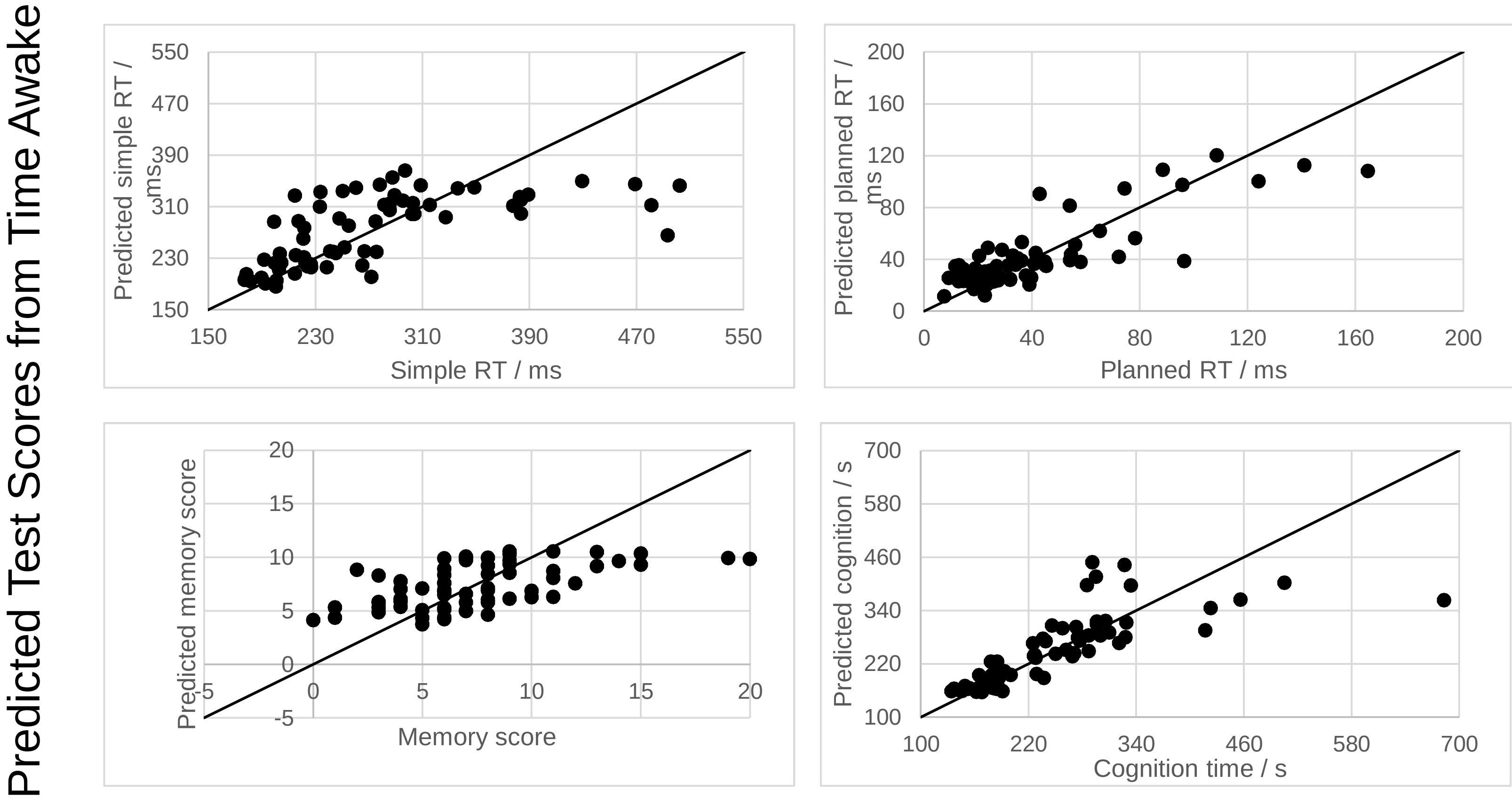


Actual time awake

Prediction of Time Awake from Speech



Prediction of Test Scores from Time Awake

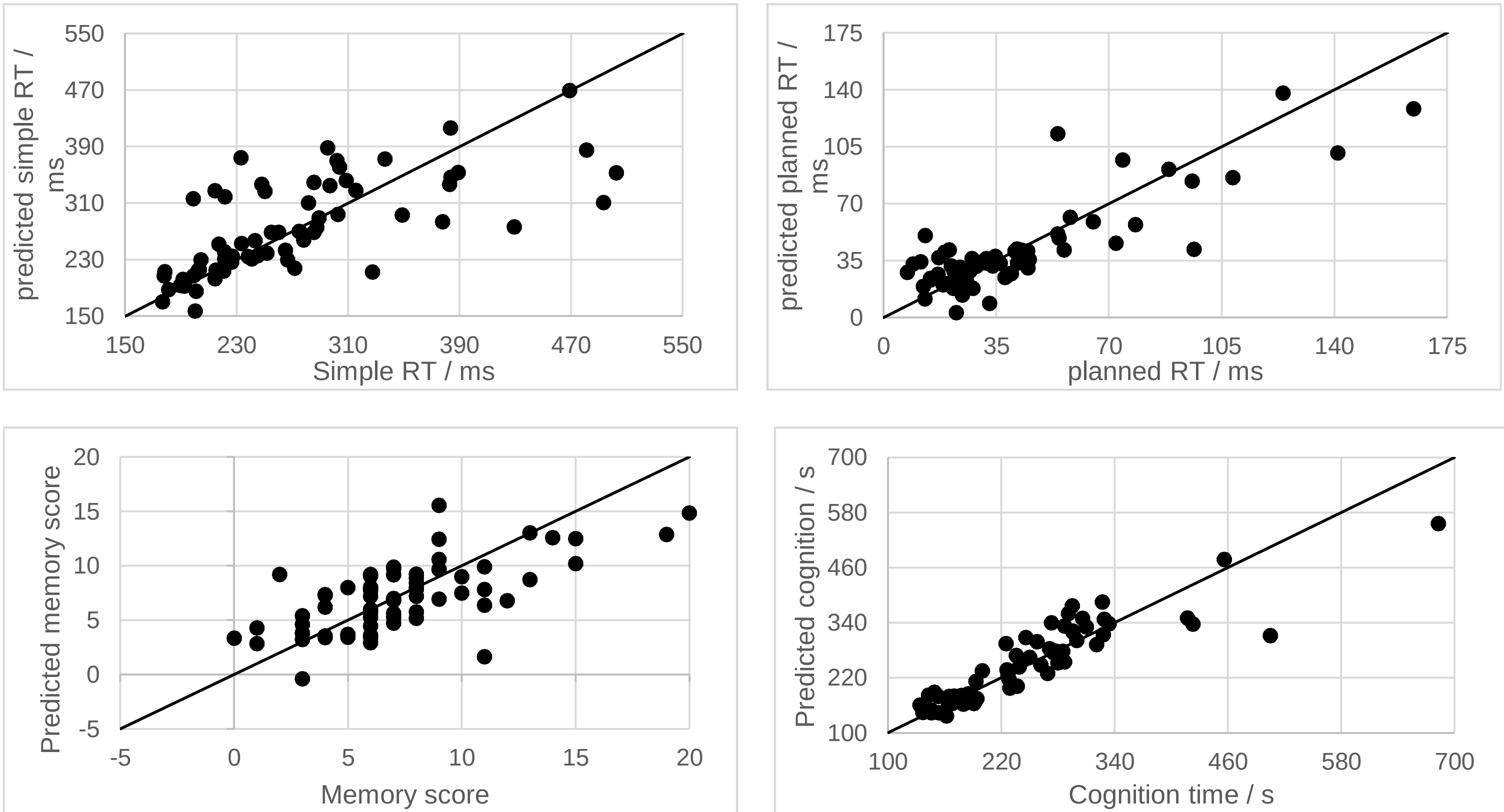


Actual test scores

r: 0.06-0.32, MAE: 0-3% reduction over NULL model

Prediction of Test Scores from Speech

Predicted Test Scores from Speech

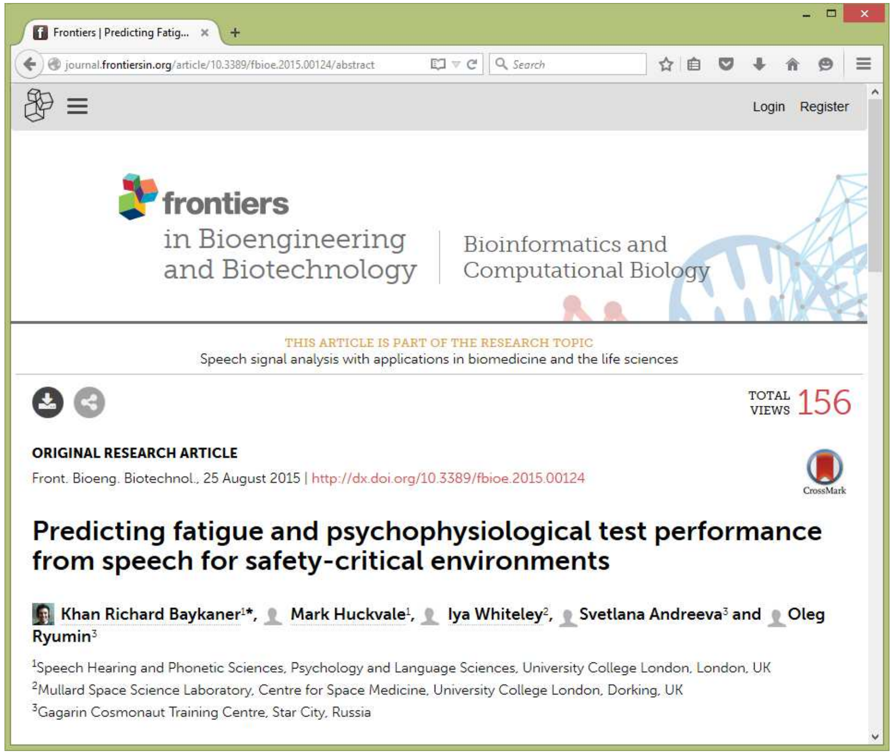


Actual test scores

r: 0.44-0.58, MAE: 7-21% reduction over NULL model

Summary

- Useful information is present in speech signal for predicting fatigue
- Regression: time awake predicted to about 1 part in 6 (r=0.73, MAE=630min)
- Classification: day-1 vs day-2/3 can be discriminated ~80% speaker independent, ~90% speaker dependent
- Availability of speech improves prediction of test scores over latency alone



Conclusions

- Changes in speech with fatigue in this task were detectable and reliable enough for classification of time spent awake.
- Better performance was achieved through Gaussianization of features, although in practice this would require an enrolment stage for speakers.
- That test scores were better predicted from the speech features than from time may be due to some common cognitive or physiological basis for test performance and speech performance.
- However, this is a very small study (7 subjects, 10 recordings over 60 hours) and concept needs to be trialled in more realistic application scenarios.