



Pearson

Pearson Test of English

Academic:

Artificial Intelligence

Applied to English

Assessment


20 August 2017

William Bonk, PhD

Director, Psychometrics & Research

Pearson

Pearson Test of English (PTE) - Academic

- The world's first fully automatically scored, high-stakes test of academic English
- Computer-based test of international, academic English
- All four skills (Listening, Speaking, Reading, Writing)
- 3 hours of testing (TOEFL = 4 hours)
- Administered at Pearson's certified test centers for high security
- Objectively and consistently scored by  automated scoring systems, including

Pearson Test of English (PTE) - Academic

- 20 different tasks
- 11 performance-based tasks integrating multiple skills
- Assesses English proficiency reliably from Basic to Advanced levels

PTE-A Score Report

PTE PEARSON TEST OF ENGLISH Academic

PEARSON

Test Taker Score Report

Example, Student

Address: 123 Example Lane
Example City, EX 12345
Example Country

Email Address: example.student@xlentstudent.com

Telephone Number: +999 321 12345678

Date of Birth: 12 August 1985

Country of Citizenship: Example Country

Gender: Male


Registration ID: 990877665

Report Issue Date: 29 October 2009

Test Date: 26 October 2009

First-Time Test Taker: No

Scores Valid Until: 26 October 2011



Overall Score: 56

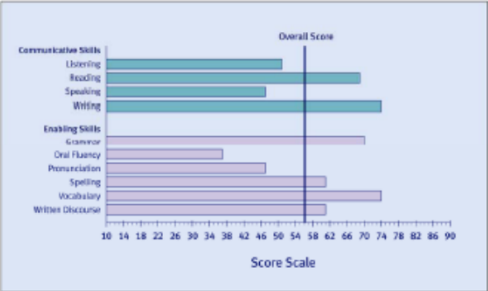
The Overall Score for PTE Academic is based on the test taker's performance on all items in the test. The scores for Communicative Skills and Enabling Skills are based on the test taker's performance on only those items that pertain to these skills specifically. As many items contribute to more than one Communicative Skill or Enabling Skill, the Overall Score cannot be computed directly from the Communicative Skills scores or from the Enabling Skills scores.

The graph below indicates this test taker's Communicative Skills and Enabling Skills relative to his or her Overall Score.

When comparing the Overall Score and the scores for Communicative Skills and Enabling Skills, please be aware that there is some inconsistency in all measurement, depending on a variety of factors. For more information on interpreting PTE Academic scores, please refer to *Interpreting the PTE Academic Score Report* which is available at www.pearsonpte.com/pteacademic/scores.

Skills Profile

Communicative Skills	
Listening	51
Reading	69
Speaking	47
Writing	74
Enabling Skills	
Grammar	70
Oral Fluency	37
Pronunciation	47
Spelling	61
Vocabulary	74
Written Discourse	61



NOTE TO INSTITUTIONS:
To obtain official, authenticated PTE Academic score information for this test taker, visit our secure website at www.pearsonvue.com/ptescores

© Edment 2009
All rights reserved; no part of this publication may be reproduced without the prior written permission of Edment Limited (a Pearson company)

Overall score

Communicative Skills

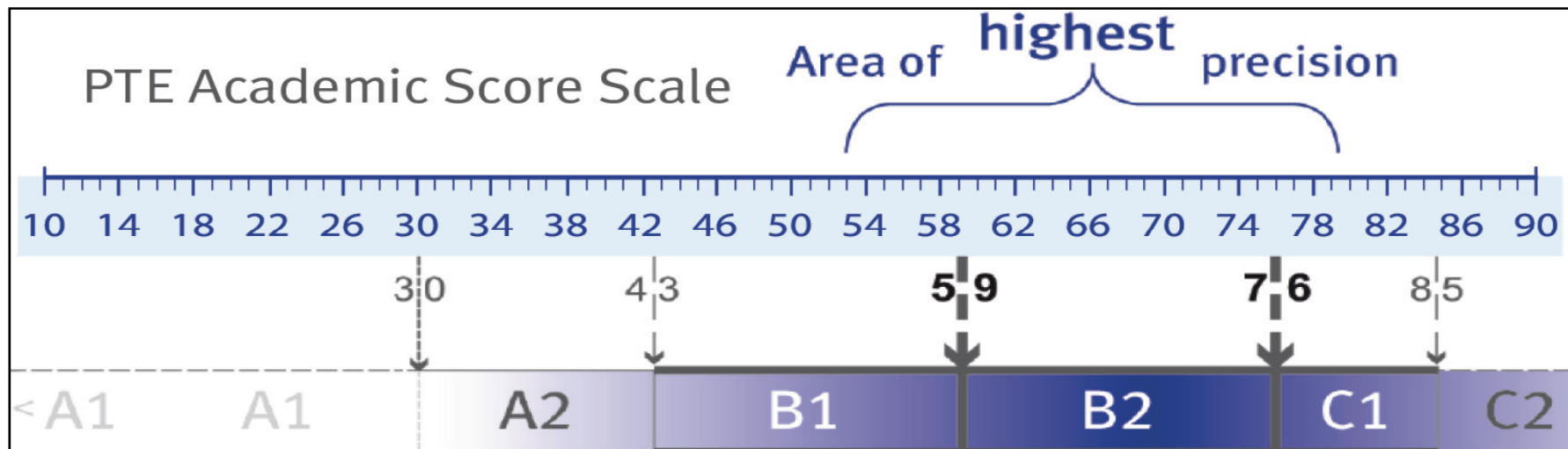
- Speaking
- Writing
- Reading
- Listening

Enabling Skills

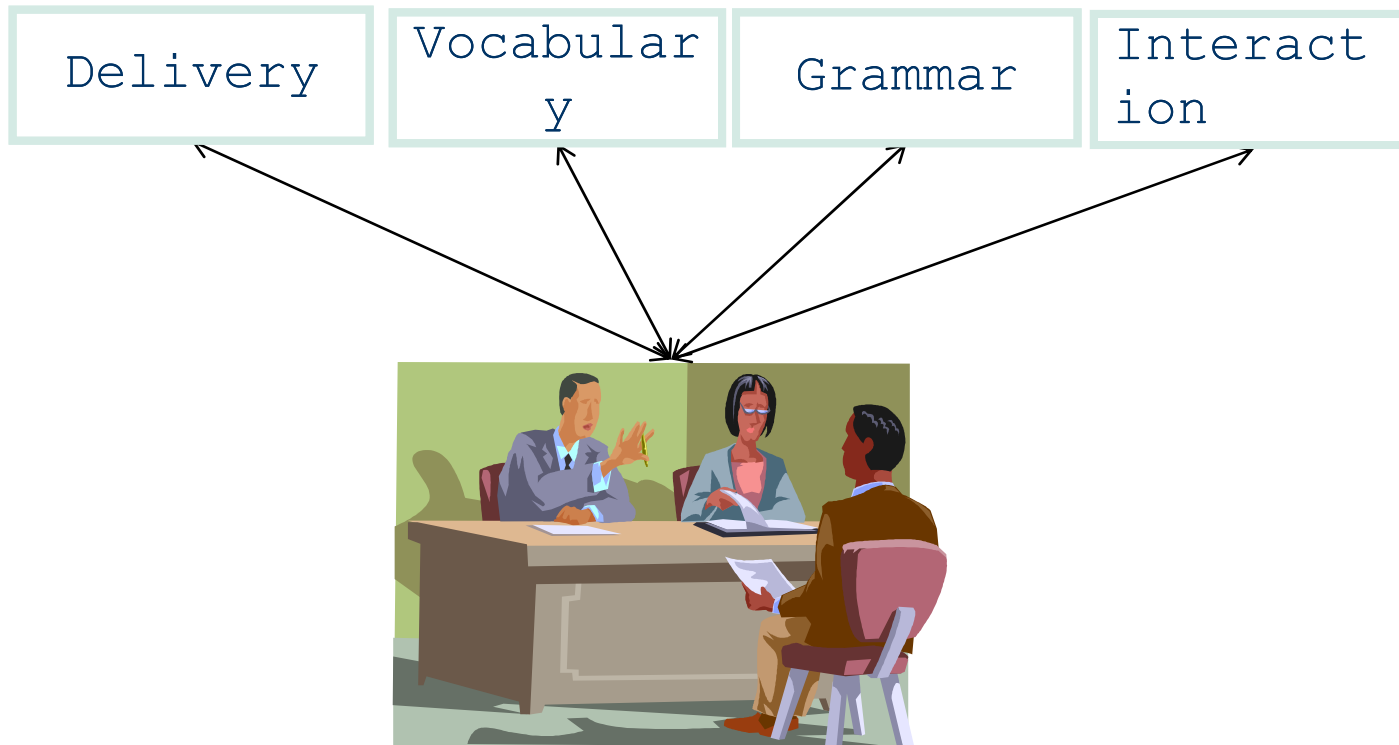
- Grammar
- Oral Fluency
- Pronunciation
- Spelling
- Vocabulary
- Written Discourse

PTE-A Score Report

- Reported on the [Global Scale of English \(GSE\)](#)
- A 10-90 scale (81 possible scores)
- Linked to Common European Framework of Reference (CEFR)
- Scores returned within 5 business days



Face-to-face interview spoken test format



Interview tests

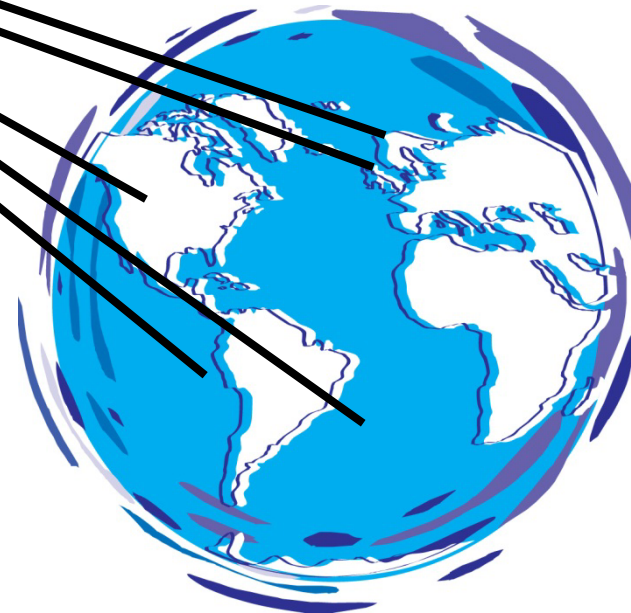
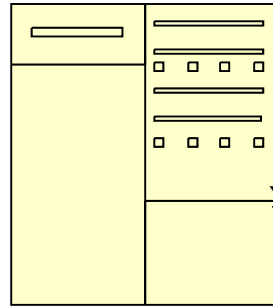
Many positive aspects of speaking skills are involved in the task, but there are also some problems:



- Power difference creates performance problems
- Inconsistent use of scoring rubrics across human raters
- Inconsistent use of scoring rubrics from test to test for individual raters
- Inconsistent use of scoring rubrics over time (drift)
- Human raters do not consistently maintain independence of scoring for different traits
- Raters can inappropriately react to gender, culture, ethnicity, appearance, accents, etc.

Automated scoring

Automated scoring systems



Standardized scoring

Speed of scoring

Objective, bias-free measurement

Data-driven models from 10,000 candidates

Accumulation of measures from multiple expert raters

Auto-scoring can assess many skills accurately


Written Scoring

- Word choice
- Grammar & Mechanics
- Progression of ideas
- Organization
- Style, Tone
- Paragraph structure
- Development, Coherence
- Point of view

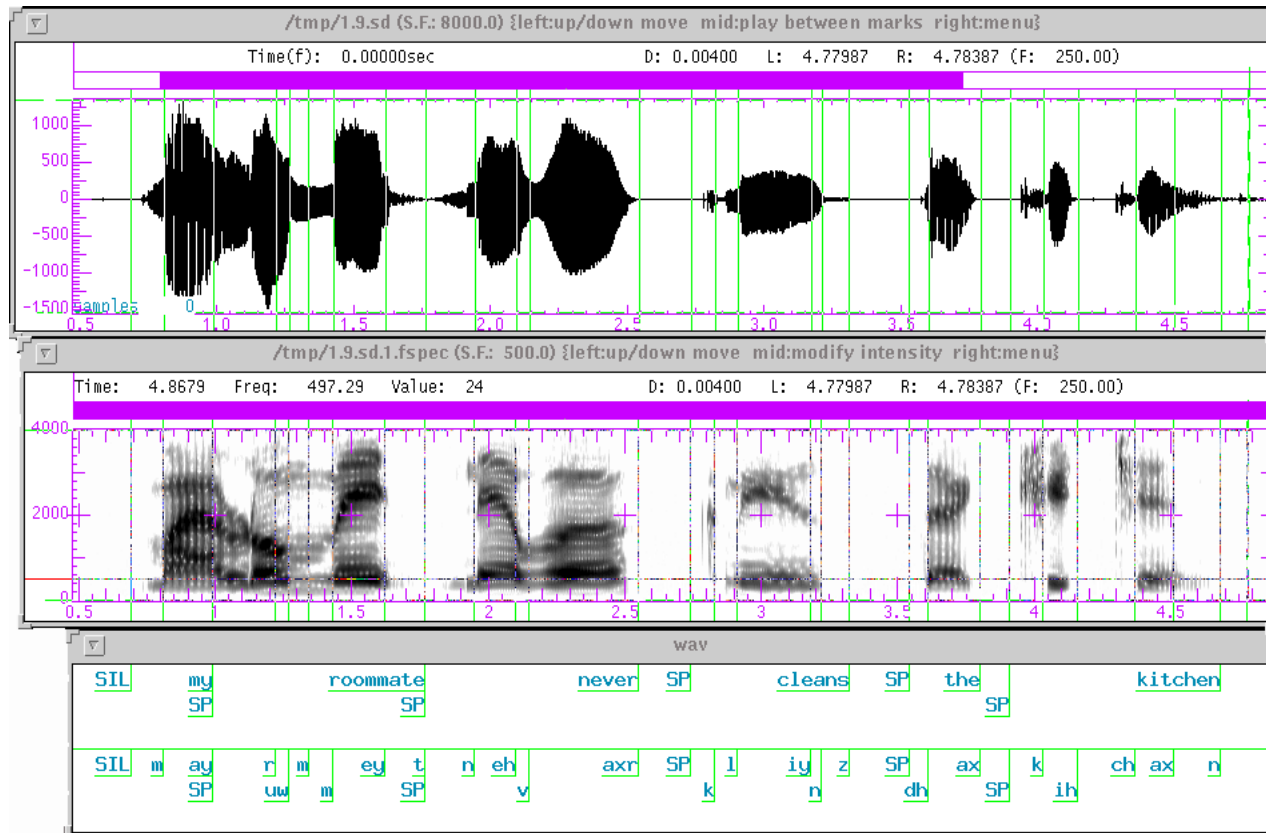
Spoken Scoring

- Sentence Mastery
- Content
- Vocabulary
- Accuracy
- Pronunciation
- Intonation
- Fluency
- Expressiveness
- Pragmatics

Latent Semantic Analysis (LSA)

- LSA reads a huge corpus of text
- Words are all fit into a **300**-dimensional model - this semantic space represents word meanings mathematically
- Words with similar meanings are close to each other in this space
- When a test-taker produces some language, the meaning of that response is basically a weighted sum of the meanings of the words
- Machine learning techniques take the “meaning” of each student response, then identify other test-taker essays that were similar in meaning
-  The machine-generated quality

Automatic Speech Recognition



Waveform

Spectrum

Words
Segmentation

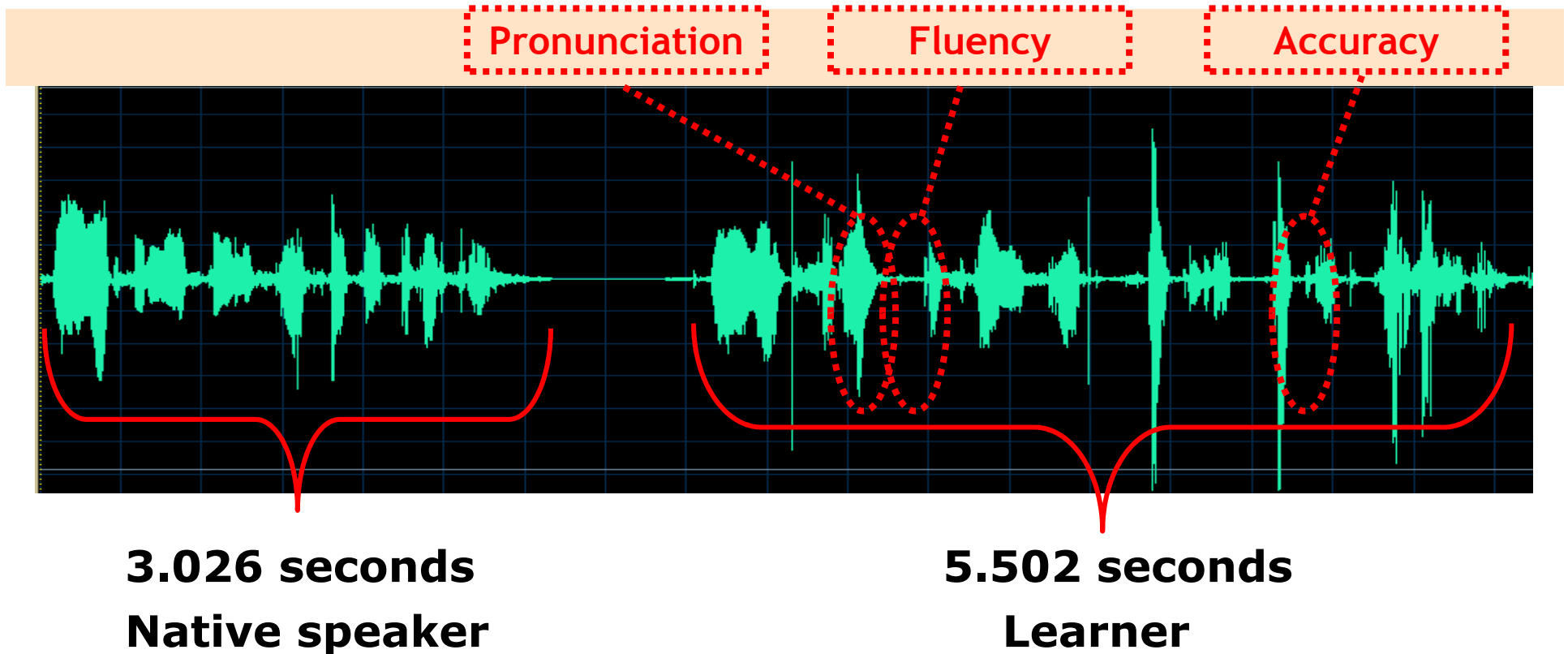
w1 w2 w3 w4 w5 w6

75-90 Words/Min

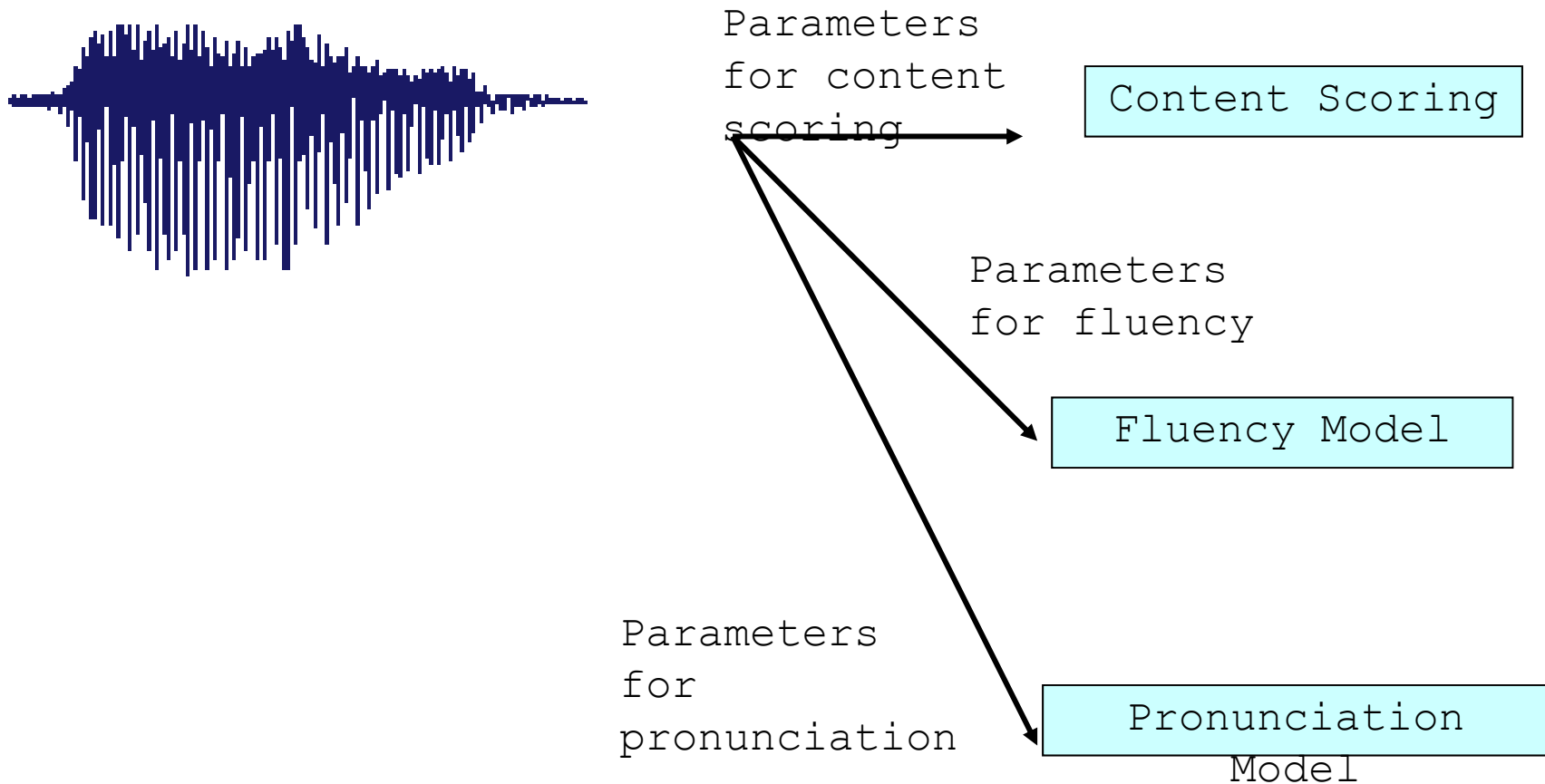
p p pppp p p p p pp ppp pp p p p p p

5.8 Phones/Sec

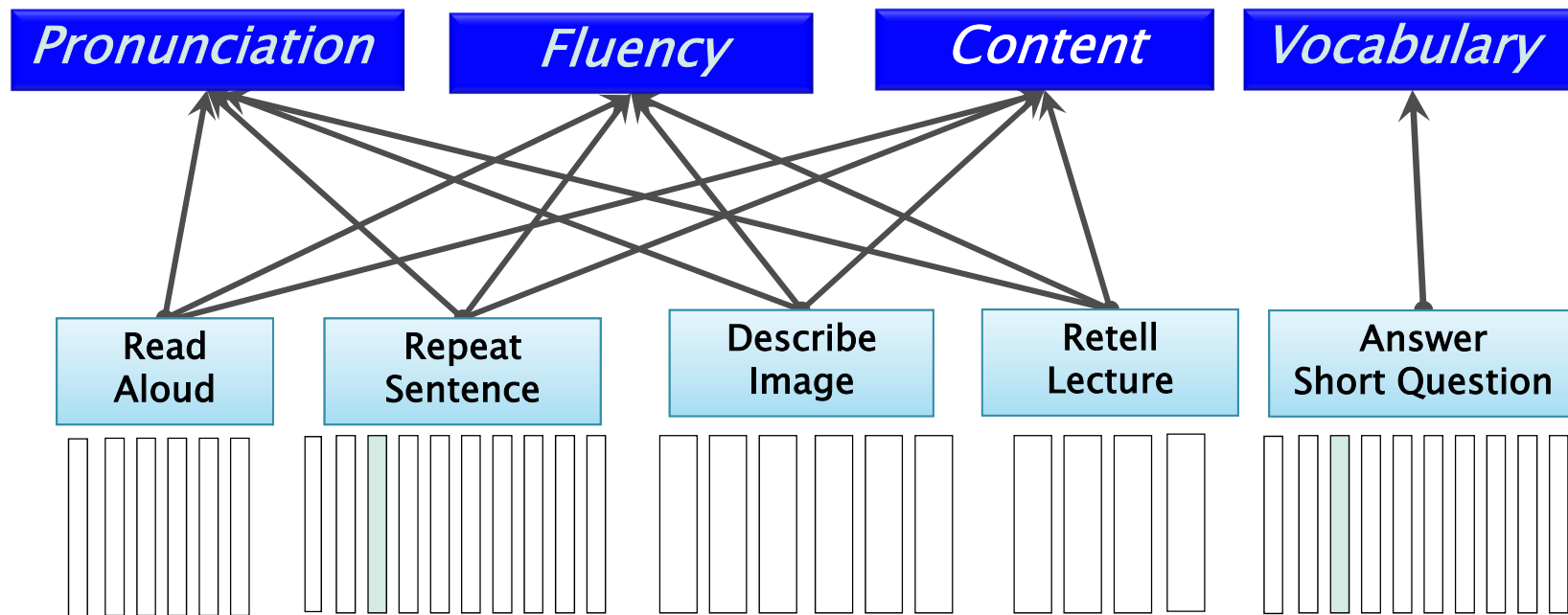
Performance Comparison



Multiple Aspects per Response



PTE Academic: Speaking



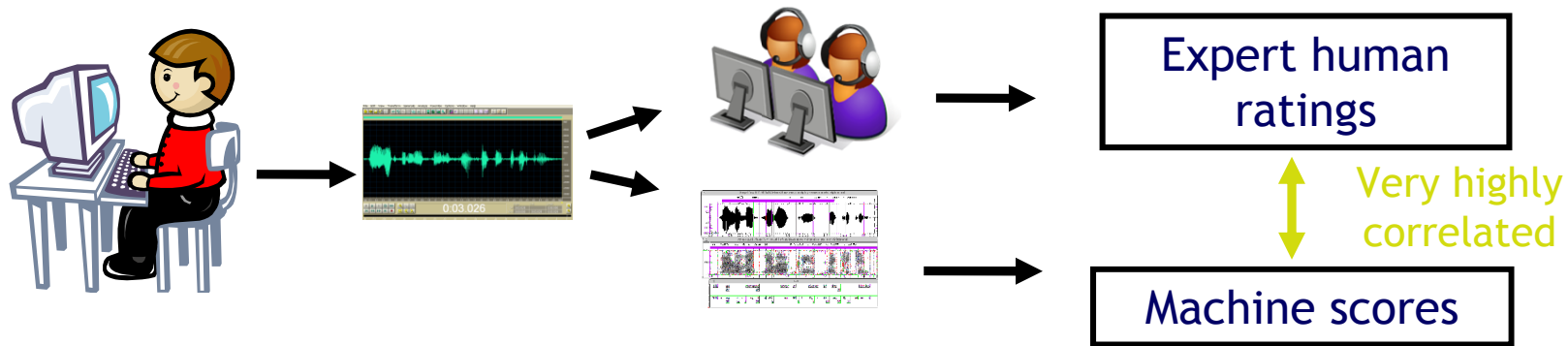
- ✓ Probing & topic shifting
- ✓ Prepared & immediate responses

- ✓ 36 responses, 8 mins speech
- ✓ Various inputs

Pearson's Automated Spoken Tests

Language	Correlation to human raters
Spanish	.97
Dutch	.93
Arabic	.98
French	.97
Chinese	.96
English	.97

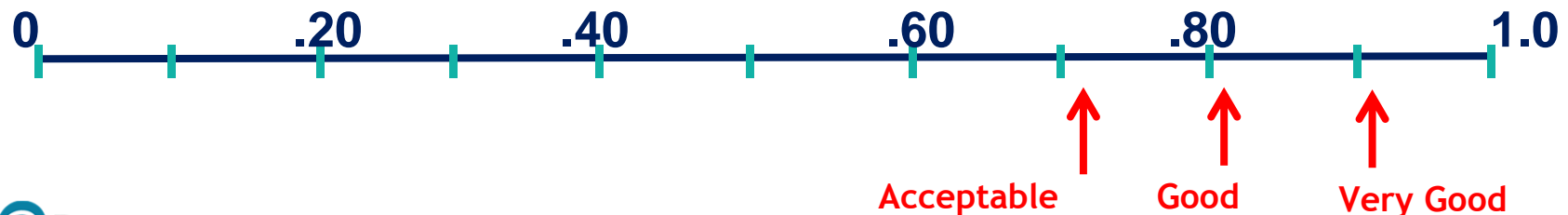
PTE-A Speaking Scores – Accuracy



	Machine-Human Correlation ($N=158$)
Pronunciation	.81
Fluency	.82
Content	.92
Vocabulary	.90
Accuracy	.95
Overall	.96

PTE Academic Test Reliability

	PTE Academic	IELTS	TOEFL
Overall	.97	.96	.94
Reading	.92	.90	.85
Listening	.91	.91	.85
Writing	.91	.81–.90	.74
Speaking	.91	.83–.86	.88



In summary, PTE Academic provides

Convenient	Secure	Accurate	Relevant & Objective
Testing over 360 days/year	Advanced multi layer security measures	Computer based marking ensures impartial, accurate marks	Uses genuine academic content and integrated tasks testing multiple skills
In over 50 countries	Most secure tests	Institutions can trust the English ability of students	Reflects the use of language students need
Book up to 24 hours before	Confidence in results		
Fast - 85% of results within 2			

