
Does the future of English language assessment start with AI?



Speakers



Ron Zeronis
Head of Content Creation



Dr. Kevin Cheung
Head of Marking and Results



Mark Hurrell
Facilitator





Who are we?



Accepted by over

25,000

universities, employers
and government departments
worldwide

Over

6.5 million

qualifications and tests
taken last year

Over

2,800

exam centers in 130
countries



Educational mission

We help people to learn English and prove their skills to the world



Our experience in Higher Education

IELTS™

**The high-stakes English test
for study, migration or work**

Cambridge

English Qualifications

**The most in-depth and reliable
four skills assessment of English**

C1 Advanced

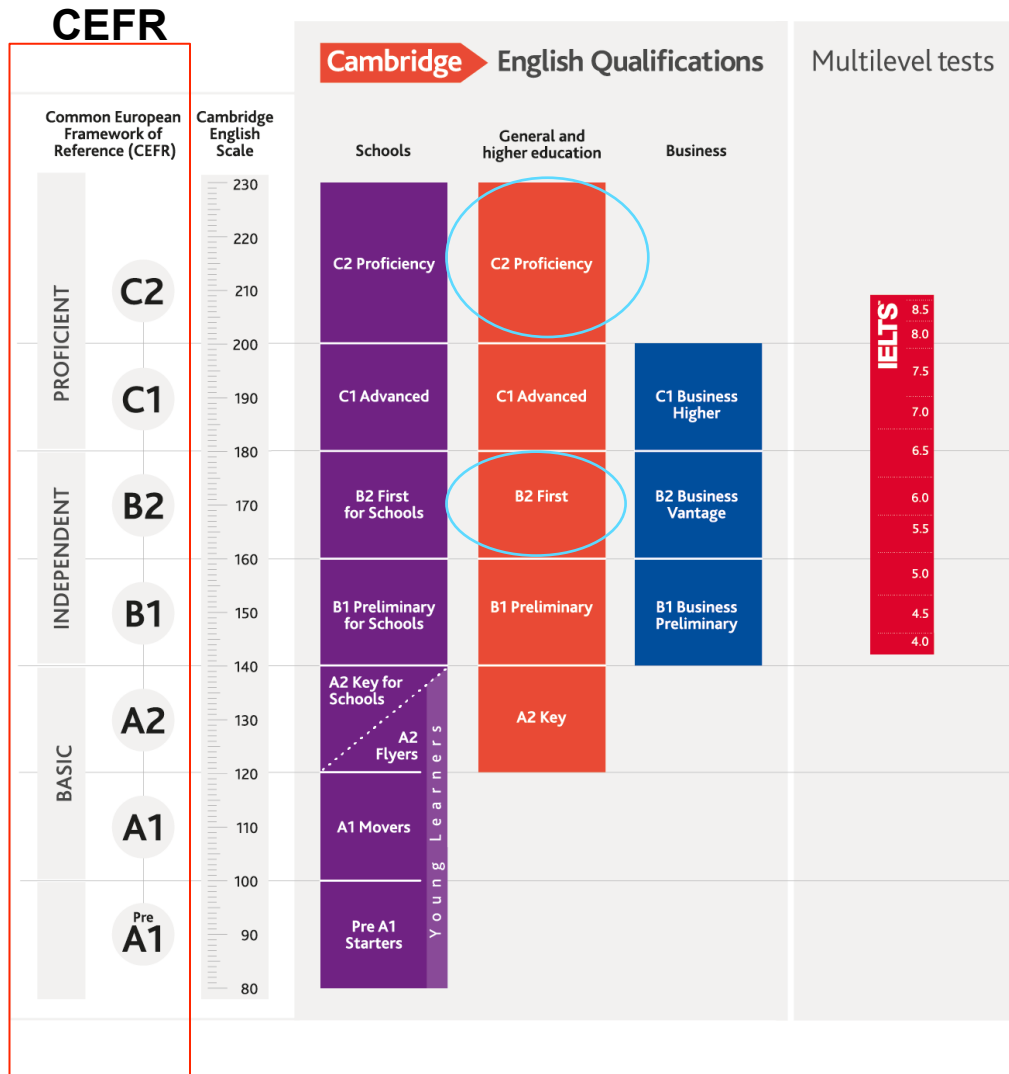
C2 Proficiency



**What good testing
for Higher Education looks like**



International standard for measuring English ability



Easy to understand results

1 Overall score

This is the overall Cambridge English Scale score for the whole exam. It is the average of the five individual scores a candidate receives for the four skills and Use of English.


The overall score is the most important piece of information for recognising institutions when setting requirements to ensure that you ask for the exact level of English you need.

2 Individual scores

Individual scores

Candidates receive a Cambridge English Scale score for each of the four skills (reading, writing, listening and speaking) and Use of English.

With these individual scores it is easy for you to specify the level of English you need in a particular skill.



Reference No.
15BB9615003
To be quoted on all
Correspondence

Certificate in Advanced English

Statement of Results

Candidate name
A N Example

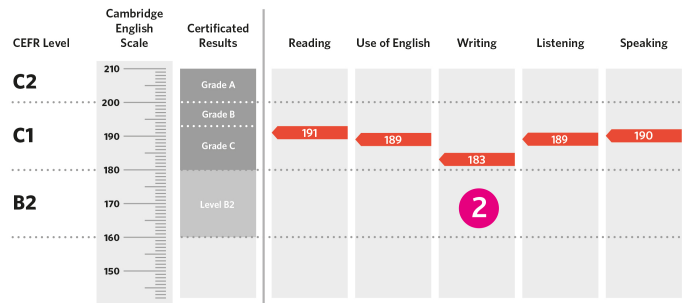
Place of entry
Cambridge

Session
**November (CAE1)
2018**

Result
Pass at Grade C

Overall Score
188

CEFR Level
C1



		Reading	Use of English	Writing	Listening	Speaking
CEFR Level	Cambridge English Scale					
C2	210					
	200					
C1	190	191	189	183	189	190
	180					
B2	170					
	160					
	150					

Advanced is an examination targeted at Level C1 in the Council of Europe's Common European Framework of Reference.

Candidates achieving Grade A (between 200 and 210 on the Cambridge English Scale) receive a certificate stating that they have demonstrated ability at Level C2. Candidates achieving Grade B or Grade C (between 180 and 199 on the Cambridge English Scale) receive a certificate at Level C1.

Candidates whose performance is below Level C1, but falls within Level B2 (between 160 and 179 on the Cambridge English Scale), receive a certificate stating that they have demonstrated ability at Level B2.

Examination results can be quickly and securely verified online at: www.cambridgeenglish.org/verifiers

Results	Score
Pass at Grade A	200 – 210
Pass at Grade B	193 – 199
Pass at Grade C	180 – 192
Level B2	160 – 179

Candidates who take Advanced and score between 142 and 159 on the Cambridge English Scale do not receive a result, CEFR level or certificate.

Cambridge English Scale scores below 142 are not reported for this examination.

Other

X - the candidate was absent from part of the examination
Z - the candidate was absent from all parts of the examination
Pending - a result cannot be issued at present, but will follow in due course
Withheld - the candidate should contact their centre for information
Exempt - the candidate was not required to sit this part of the examination

THIS IS NOT A CERTIFICATE
Cambridge Assessment English reserves the right to amend the information given before the issue of certificates to successful candidates.

What does good testing for Higher Education look like?

Validity

Reliability

Impact

Practicality

Quality

Good testing helps improve learning and teaching

Cambridge
English Qualifications

*What teachers and school leaders say about us:

95%



agree that preparing for Cambridge English Qualifications improves students' English.

88%



agree that Cambridge English Qualifications provide a clear curriculum framework to follow, with clear milestones.

86%



agree that Cambridge English Qualifications help to build students' confidence.

$\frac{3}{4}$



agree that there are lots of support materials available to teach the exams.

$\frac{9}{10}$



agree that Cambridge English Qualifications have a positive effect on students' motivation.

$\frac{3}{4}$



agree that Cambridge English Qualifications are easy to implement.

$\frac{8}{10}$



agree that Cambridge English Qualifications help to improve teaching standards.

*Cambridge English Qualifications Schools Survey asked teachers and school leaders their views about the impact Cambridge English Qualifications have on learning and teaching. It received a total of 5,789 responses from 109 countries.



Good testing helps to identify applicants who are best prepared for university

“ C1 students are of a higher calibre and more prepared to face the challenges and opportunities that exist when studying on a US campus. ”

Shawn Greenfield

Interim Director





International Marketing and Recruitment

University of Idaho



What's in the exam?

C1 Advanced is made up of four papers to test different aspects of candidates' English skills

Paper	Content	Shows students can ...
 Reading and Use of English (1 hour 30 minutes)	8 parts/ 56 questions	deal confidently with various texts and grasp the main topic of a text. This paper also requires candidates to demonstrate their control of grammar and vocabulary.
 Writing (1 hour 30 minutes)	2 parts	produce two different pieces of writing, including a compulsory academic essay.
 Listening (about 40 minutes)	4 parts/ 30 questions	follow and understand a range of spoken materials – such as lectures, interviews, discussions, presentations and everyday conversations.
 Speaking (15 minutes per pair of candidates)	4 parts	communicate effectively in face-to-face situations such as seminars or everyday conversations.

The overall score is calculated by averaging the scores achieved in Reading, Writing, Listening, Speaking and Use of English. The weighting of each of the four skills and Use of English is equal.

C1 Advanced reading task

XXXX XXXXXXX - 123456 90 minutes left Help Hide

Read the article below containing four reviews of a book about architecture. For questions 37 – 40, choose the correct review. The reviews may be chosen more than once.

The Architecture of Happiness

Four reviewers comment on philosopher Alain De Botton's book

A
Alain de Botton is a brave and highly intelligent writer who writes about complex subjects, clarifying the arcane for the layman. Now, with typical self-assurance, he has turned to the subject of architecture. The essential theme of his book is how architecture influences mood and behaviour. It is not about the specifically architectural characteristics of space and design, but much more about the emotions that architecture inspires in the users of buildings. Yet architects do not normally talk nowadays very much about emotion and beauty. They talk about design and function. De Botton's message, then, is fairly simple but worthwhile precisely because it is simple, readable and timely. His commendable aim is to encourage architects, and society more generally, to pay more attention to the psychological consequences of design in architecture: architecture should be treated as something that affects all our lives, our happiness and well-being.

B
Alain de Botton raises important, previously unasked, questions concerning the quest for beauty in architecture, or its rejection or denial. Yet one is left with the feeling that he needed the help and support of earlier authors on the subject to walk him across the daunting threshold of architecture itself. And he is given to making extraordinary claims: 'Architecture is perplexing ... in how inconsistent is its capacity to generate the happiness on which its claim to our attention is founded.' If architecture's capacity to generate happiness is inconsistent, this might be because happiness has rarely been something architects think about. De Botton never once discusses the importance of such dull, yet determining, matters as finance or planning laws, much less inventions such as the lift or reinforced concrete. He appears to believe that architects are still masters of their art, when increasingly they are cogs in a global machine for building in which beauty, and how de Botton feels about it, are increasingly beside the point.

Which reviewer

37 has a different opinion from the others on the confidence with which de Botton discusses architecture?

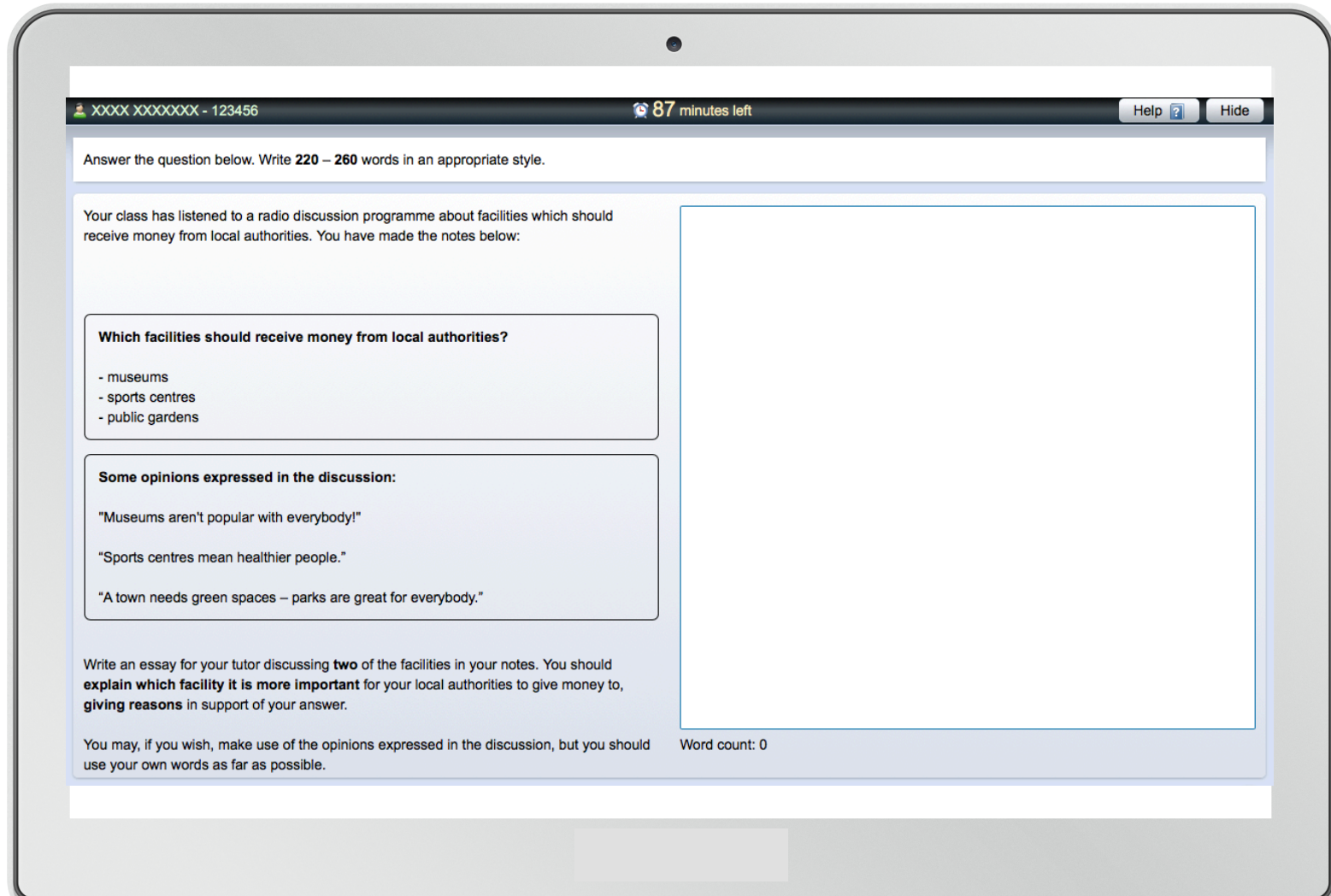
Text A
 Text B
 Text C
 Text D

38 shares reviewer A's opinion whether architects should take note of de Botton's ideas?

39 expresses a similar view to reviewer B regarding the extent to which architects share de Botton's concerns?

40 has a different view to reviewer C on the originality of some of de Botton's ideas?

C1 Advanced writing task



XXXXX XXXXXXXX - 123456 87 minutes left Help ? Hide

Answer the question below. Write **220 – 260** words in an appropriate style.

Your class has listened to a radio discussion programme about facilities which should receive money from local authorities. You have made the notes below:

Which facilities should receive money from local authorities?

- museums
- sports centres
- public gardens

Some opinions expressed in the discussion:

"Museums aren't popular with everybody!"

"Sports centres mean healthier people."

"A town needs green spaces – parks are great for everybody."

Write an essay for your tutor discussing **two** of the facilities in your notes. You should **explain which facility it is more important** for your local authorities to give money to, **giving reasons** in support of your answer.

You may, if you wish, make use of the opinions expressed in the discussion, but you should use your own words as far as possible.

Word count: 0

C1 Advanced listening task

XXXXX XXXXXXX - 123456 36 minutes left Help Hide

You will hear five people talking about changing their jobs. For Task 1, choose from the list the reason each speaker gives for changing job. For Task 2, choose from the list what each speaker feels about their new job.

Task 1

	Speaker 1	Speaker 2	Speaker 3	Speaker 4	Speaker 5
unfriendly colleagues					
poor holiday entitlement					
lacking a sense of purpose					
needing more of a challenge					
the workload					
disagreements with superiors					
no prospect of advancement					
the physical environment					

Task 2

	Speaker 1	Speaker 2	Speaker 3	Speaker 4	Speaker 5
encouraged by early results					
hopeful about future success					
delighted by a change in lifestyle					
relieved the initial uncertainty is over					
glad to be helping other people					
grateful for an increase in salary					
happy to feel in control					
satisfied with the training received					

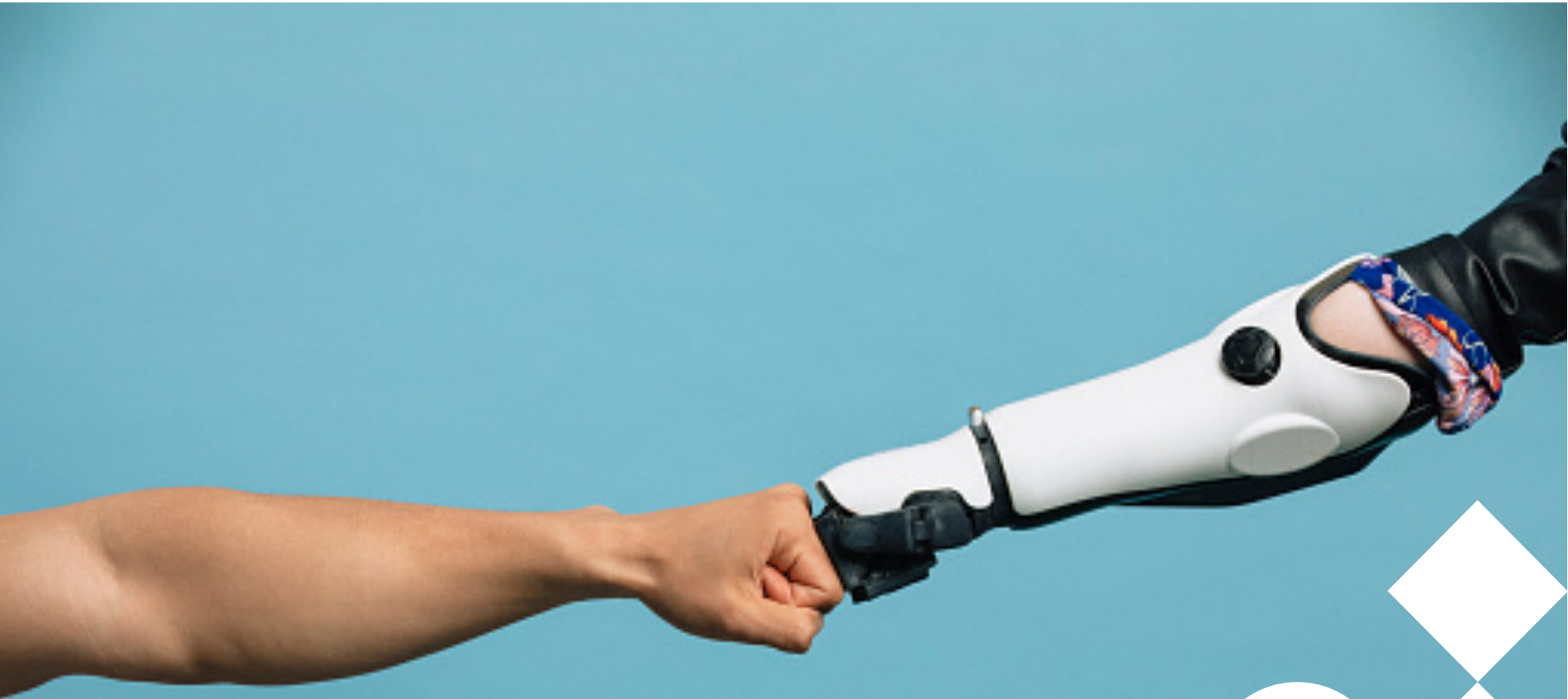
Review

Part 1: 1 2 3 4 5 6 Part 2: 7 8 9 10 11 12 13 14 Part 3: 15 16 17 18 19 20 Part 4: 21 22 23 24 25 26 27 28 29 30

⏪ ⏩

What can students with C1 Advanced do?

- Contribute effectively to meetings and seminars within their own area of work or keep up a casual conversation with a good degree of fluency, coping with abstract expressions.
- Read quickly enough to cope with an academic course, and take reasonably accurate notes in meetings or write a piece of work which shows an ability to communicate.
- Make critical remarks/express disagreement without causing offence.
- Scan texts for relevant information, and grasp the main topic of a text.
- Write a piece of work whose message can be followed throughout.



The future

ALTA Institute (2013-Present)

The Institute of **A**utomated **L**anguage **T**eaching and **A**ssessment

Research in computer systems and platforms, corpus linguistics, computational linguistics, speech processing and machine learning



Cambridge ALTA

Institute for Automated Language Teaching and Assessment

“ ALTA carries out cutting-edge research combining **machine learning** and **natural language processing** to investigate **personalized adaptive** approaches to assessment which will benefit learners and teachers worldwide. ”

Professor Paula Buttery, Director of ALTA

 Cambridge Assessment
English Computer
Laboratory Department of
Engineering Theoretical and
Applied Linguistics

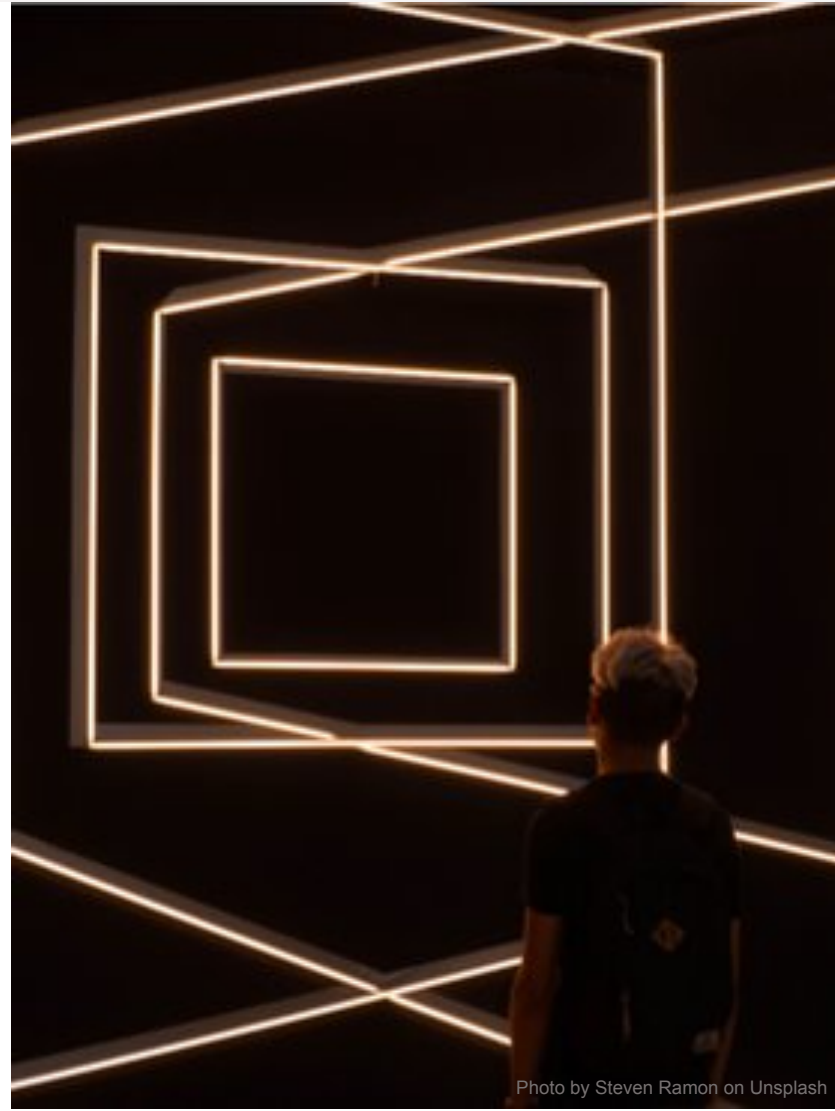
AI technologies currently used in ALTA research

Machine Learning and Natural Language Processing

- Natural language processing
- Supervised and unsupervised machine learning
- Deep neural networks
- Machine translation
- Sequence labelling
- Automatic speech recognition
- Speaker diarisation
- Speech-unit segmentation
- Creating multidimensional representations of learners and tasks

So what does the future of language testing look like?

- 1 How test questions are created
- 2 How exams are marked
- 3 How learners are supported to prepare for their studies



ALTA Institute (2013-Present)

Research Areas

Content Creation
and Calibration

Automated
Assessment

Cambridge ALTA

Institute for Automated Language Teaching and Assessment

Automated
Feedback

Adaptive
Learning

Quick Poll

Which of the following English skills is most important for students at your university?

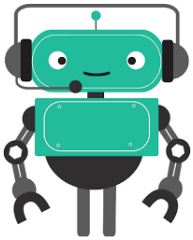
Students at my university need to use English for...

- a) ...finishing incomplete words by typing missing letters
- b) ...writing sentences to describe pictures
- c) ...producing verbatim transcripts of short speech
- d) ...writing structured essays.

So what does the future of language testing look like?

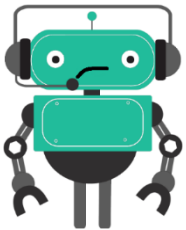
1 How test questions are created and administered

Identifying source material



Finding large amounts
of material

Classifying that material
is right for a level



Considering culture
and taboo

Explaining why material
is right for a level

Generating questions

Creating non-English
words

Generating gap fill or
word reordering tasks

Simulating authentic
contexts

Creative innovation

Delivering tests

Constructing unique
tests automatically

Managing the bank
of items

Interactive
communication

Improvising to deal with
unforeseen issues

AI-led vs AI-powered question creation

Led by AI

Test designed only around tasks that:

- have clear right and wrong answers
- do not need review by human experts
- AI can automatically generate
- AI can automatically mark.

Impact

- Narrow range of English skills assessed
- Increased risk of getting applicants with poor English who have been 'coached'
- Lack of authenticity

Powered by AI

Test content based on:

- how students use English when they study at university and beyond
- authentic communication tasks
- covering a wide range of skills.

AI used for:

- automatically creating simple questions
- finding material for experts to use in more complex questions
- flagging potential issues for review during quality assurance.

So what does the future of language testing look like?

2 How exams are marked

Easier for a
computer to mark

Difficult for a
computer to mark



Inauthentic

Authentic



Multiple choice
& gap filling with
only one answer

Short responses
to questions with
clear right &
wrong answers

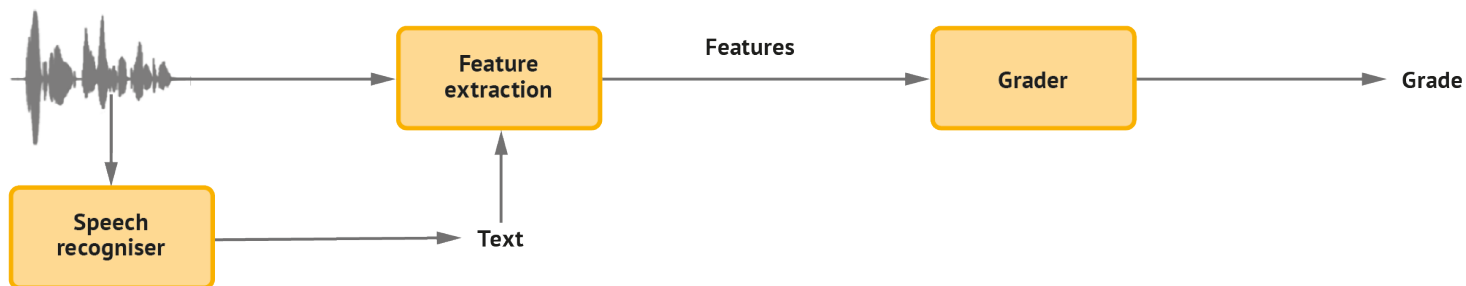
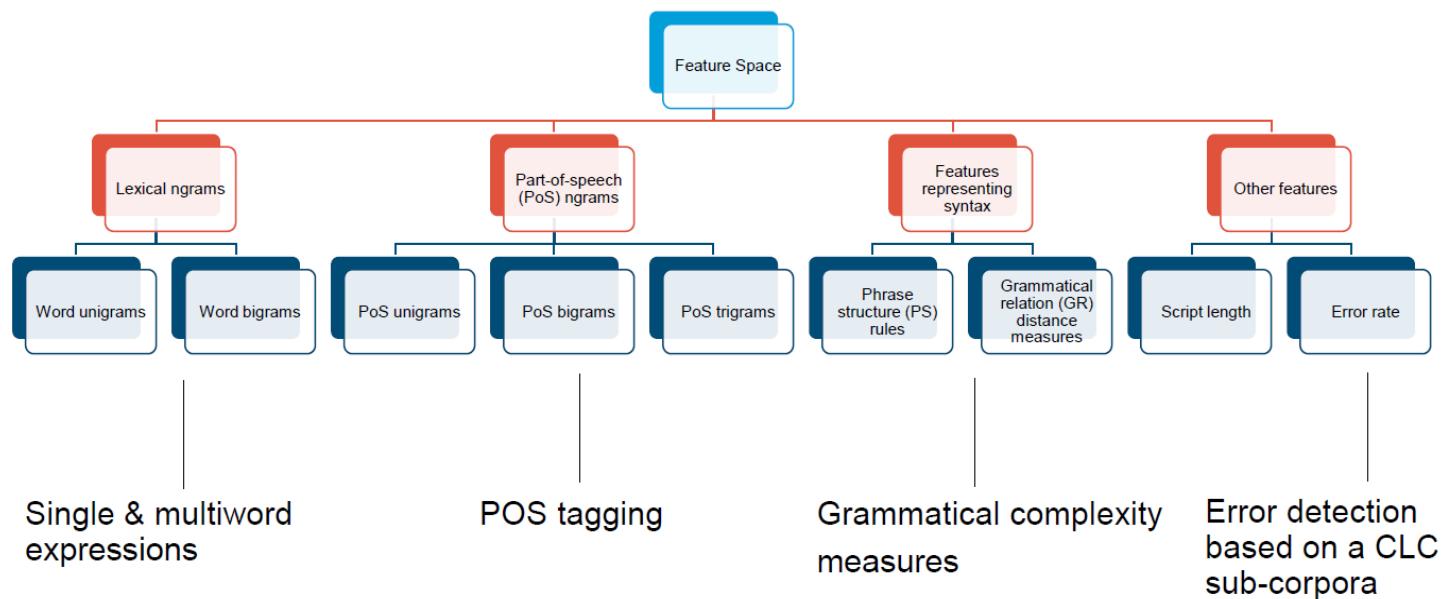
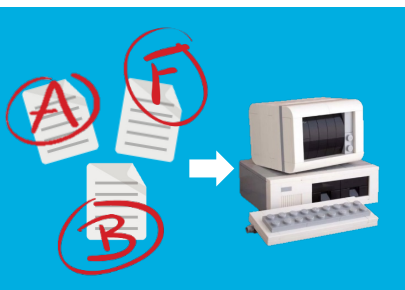
Read aloud &
dictation tasks

Extended
writing

Spontaneous
speech

Multi-speaker
interactions

Writing and speaking automarkers - A brief outline



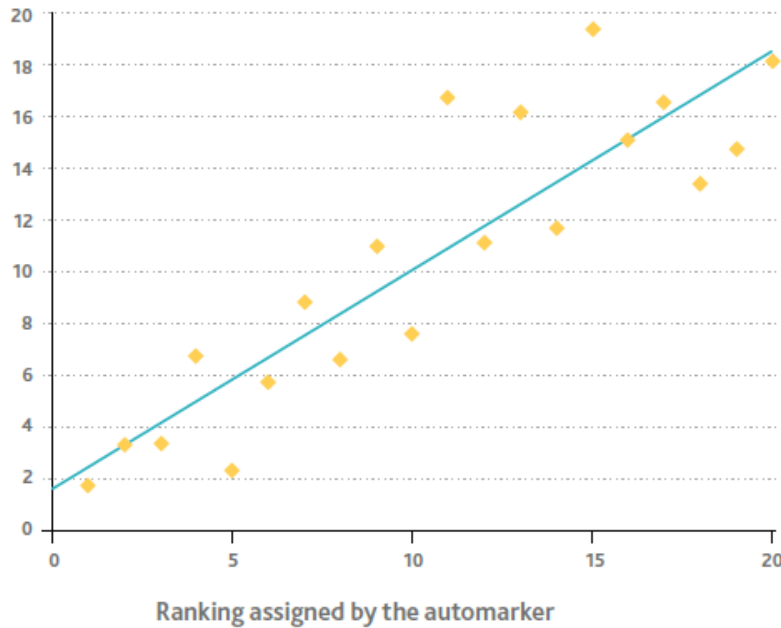
AI automarkers of extended text and spontaneous speech

Performance when compared with humans

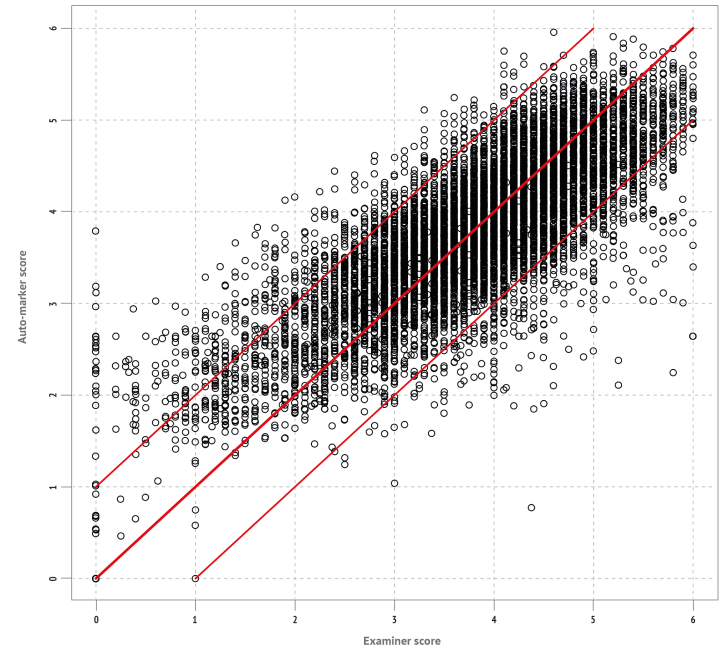
Writing: Spearman Rho = **0.82**

Speaking: RMSE = **0.64**

Average of rankings assigned by human experts



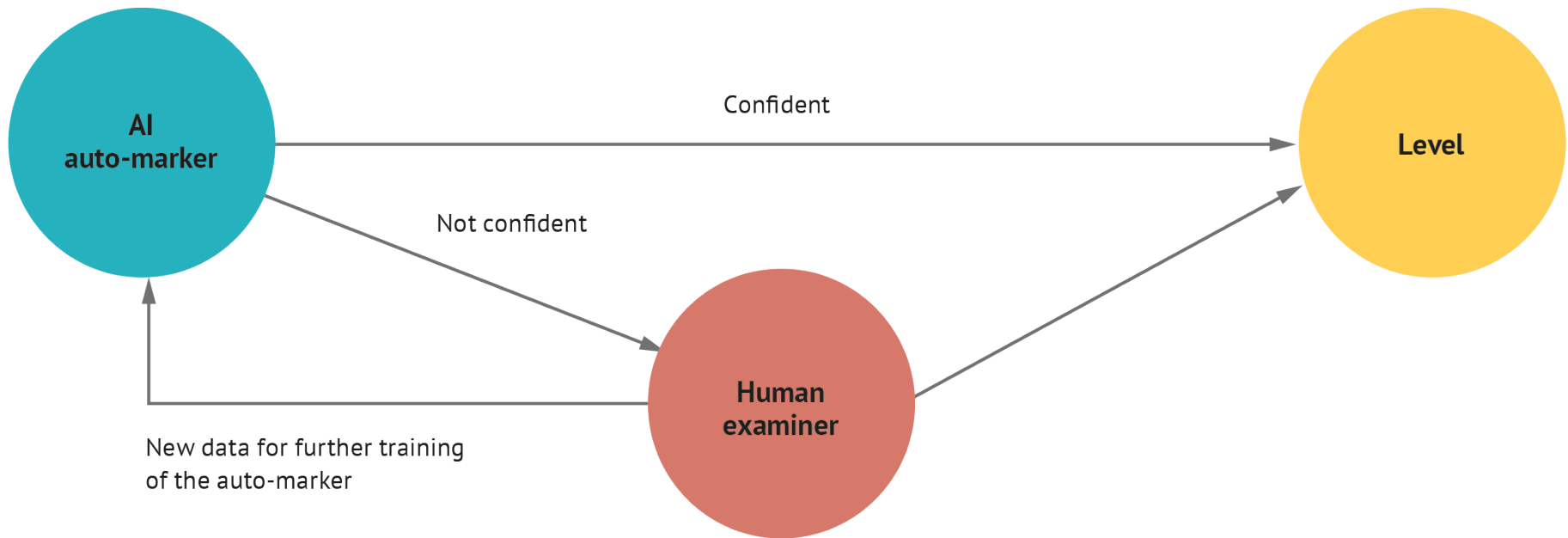
Cheung, Xu & Lim (2017)



Xu et al. (2017)

Hybrid Marking

Combining AI with human marking



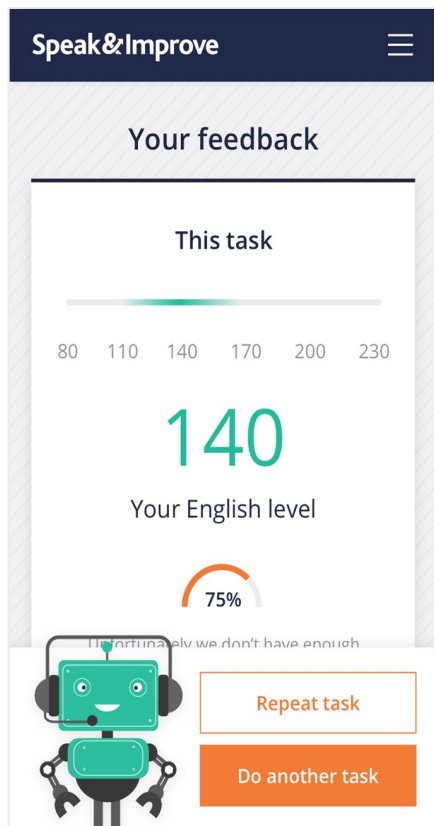
The AI is not quite there....

Interactive Communication



So what does the future of language testing look like?

3 How learners are supported to prepare for their studies



Speak&Improve

Your feedback

This task

80 110 140 170 200 230

140

Your English level

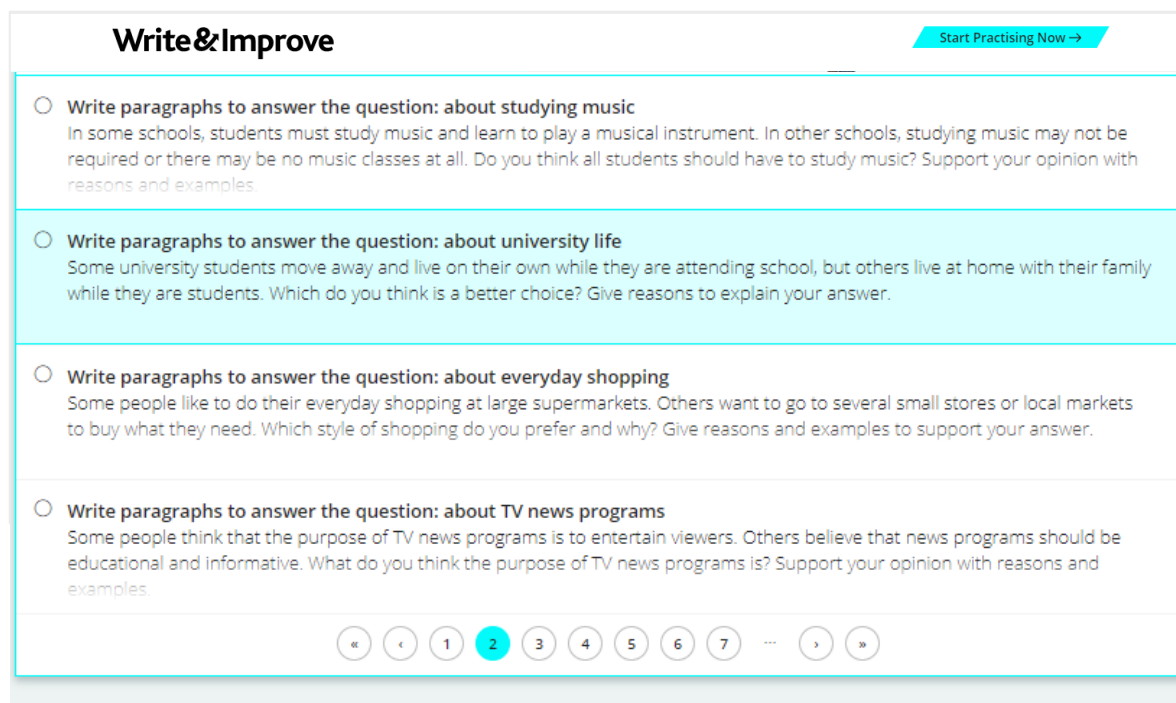
75%

Unfortunately we don't have enough

[Repeat task](#)

[Do another task](#)

“ By failing to prepare, you are preparing to fail.”
Benjamin Franklin



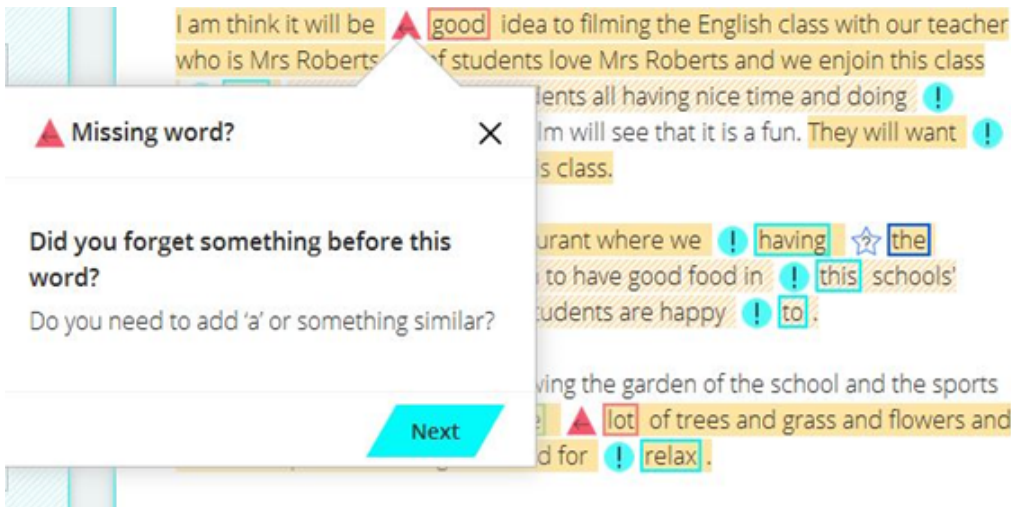
Write&Improve [Start Practising Now →](#)

- Write paragraphs to answer the question: about studying music**
In some schools, students must study music and learn to play a musical instrument. In other schools, studying music may not be required or there may be no music classes at all. Do you think all students should have to study music? Support your opinion with reasons and examples.
- Write paragraphs to answer the question: about university life**
Some university students move away and live on their own while they are attending school, but others live at home with their family while they are students. Which do you think is a better choice? Give reasons to explain your answer.
- Write paragraphs to answer the question: about everyday shopping**
Some people like to do their everyday shopping at large supermarkets. Others want to go to several small stores or local markets to buy what they need. Which style of shopping do you prefer and why? Give reasons and examples to support your answer.
- Write paragraphs to answer the question: about TV news programs**
Some people think that the purpose of TV news programs is to entertain viewers. Others believe that news programs should be educational and informative. What do you think the purpose of TV news programs is? Support your opinion with reasons and examples.

« ‹ 1 2 3 4 5 6 7 … › »

Supporting students to prepare for university

AI that helps students develop genuine communication skills



The screenshot shows a writing interface with a student's text: "I am think it will be good idea to filming the English class with our teacher who is Mrs Roberts of students love Mrs Roberts and we enjojn this class". The text is highlighted in yellow. A feedback pop-up window is open, titled "Missing word?". It asks "Did you forget something before this word?" and "Do you need to add 'a' or something similar?". The pop-up has a "Next" button. Other feedback icons (exclamation marks and stars) are visible on the text, such as "having", "the", "this", "to", "lot", and "relax".

Granular feedback on:

- Spelling
- Grammar
- Vocabulary choice

Highlights areas that need attention for some examples, why.

How the learner is doing **now**, compared to their last performance

- Personal
- Adaptive
- Formative

Feedback

Help ?

Level
B1



*That's amazing! Your writing is really improving.
Pay attention to the feedback. What changes can you make? Why not try to write longer sentences?
Keep writing to keep improving!*

Recent ALTA research

Topics that have been published on in the last 2 years

- Detecting off-topic responses
- Assessing text coherence
- Detecting disfluency in speech
- Novel approaches to Grammatical Error Correction
- Investigating the impact of ASR performance on assessment
- Predicting learner performance
- Automatic homework selection
- Identifying offensive and abusive language
- Automatic text simplification
- Assessing pronunciation in speech
- Advances in Automated Speech Recognition

The future of language testing - AI powered assessment

“ Some people call this artificial intelligence, but the reality is this technology will enhance us. So instead of artificial intelligence, I think we'll augment our intelligence. ”

***Ginni Rometty, Executive
Chairman, IBM***



Get in touch



Recognition@CambridgeEnglish.org



Discussion

