

## Sample Exam

### ISO 2025-2026 ENGLISH 1 SAMPLE EXAM

This exam consists of two parts based on a reading text.

Part 1 involves answering 10 comprehension questions. 1 point is allocated for each correct answer (0.5 points) written in correct English (0.5 points).

Part 2 consists of summarising the text using your own words (20 points).

---

Article from: <https://theconversation.com/ai-systems-are-built-on-english-but-not-the-kind-most-of-the-world-speaks-249710>

Published: May 5, 2025 9.04pm BST

#### **Author:**

Celeste Rodriguez Louro

Associate professor, Chair of Linguistics and Director of Language Lab, The University of Western Australia

---

## **AI systems are built on English – but not the kind most of the world speaks**

An estimated 90% of the training data for current generative AI systems stems from English. However, English is an international lingua franca with about 1.5 billion speakers worldwide, and countless varieties.

So whose English is today's technology based on? The answer is primarily the English of mainstream America.

This is no accident. Mainstream American English is entrenched in the digital infrastructure of the internet, in Silicon Valley's corporate priorities, and in the data sets that fuel everything from autocorrect to AI-generated synthetic text.

The consequence? AI models produce a monolithic version of English that erases variation, excludes minoritised and regional voices, and reinforces unequal power dynamics.

### ***The hegemony of mainstream American English***

The proliferation of American English online is a result of historical, economic and technological factors. The United States has been a dominant force in the development of the internet, content creation, and the rise of tech giants such as Google, Meta, Microsoft and OpenAI.

Unsurprisingly, the linguistic norms embedded in products by these companies are overwhelmingly mainstream American.

### **Understand how AI is changing society**

A recent study found that speakers of non-mainstream English were frustrated with the “homogeneity of AI accents” in voice-cloning and speech-generation technologies. One participant noted the predominant mainstream American accents in the voices available, stating the technologies had been built “with some other people in mind”.

Mainstream varieties of English have long reigned as the “standard” against which other varieties are weighed.

To take a single example from the US, linguistics research by John Baugh found that using different accents can determine people’s access to goods and services. When Baugh called different landlords about housing advertised in the local newspaper, using a mainstream accent procured him several housing inspections while using African-American and Latino accents did not.

The prestige of mainstream English also underpins algorithmic decisions. The models behind tools such as autocorrect, voice-to-text, or even AI writing assistants are most often trained on mainstream American-centric data. This is often scraped from the web, where US-based media, forums and platforms dominate.

This means variations in grammar, syntax and vocabulary from other varieties of English are systematically ignored, misinterpreted or outright “corrected”.

### **Whose English is perceived as adding value?**

The stakes of this linguistic bias in favour of mainstream English become even higher when AI systems are deployed around the world.

If an AI tutor fails to understand a Nigerian English construction, who bears the cost? If a job application written in Indian English is marked down by an AI-powered resume scanner, what are the consequences? If an Australian First Nations elder’s oral history is transcribed by voice recognition software and the system fails to capture culturally significant terms, what knowledge is lost or misrepresented?

These questions are unfolding in real time as governments, educational institutions and corporations adopt AI technologies at scale.

### **Englishes, not English**

The idea that there is one “good” or “correct” English is a myth. English is spoken in diverse forms across regions, shaped by local societies, cultures, histories and identities.

As Noongar writer and educator Glenys Collard and I have written, Aboriginal English has “its own structure, rules and the same potential as any other linguistic variety” and the same is true of other forms of English.

Indian English, for example, has lexical innovations such as “prepone” (the opposite of postpone). Singapore English (Singlish) integrates particles and syntactic features from Malay, Hokkien and Tamil.

These are not “broken” forms of English. Each community where English was imposed has gone on to make English its own.

English, and language more generally, is never static. It adapts to meet the needs of an ever-changing society and its speakers.

Yet in AI development, this linguistic diversity is often treated as noise rather than signal. Non-standardised varieties are underrepresented in training datasets, excluded from annotation schemes, and rarely feature in evaluation benchmarks.

This results in an AI ecosystem that is multilingual in theory, but monolingual in practice.

### **Toward linguistic justice in AI**

So, what would it look like to build AI systems that recognise and respect a range of different forms of English?

A shift in mindset is required, from prescribing “correct” language to including many varieties of language. What we need are systems that accommodate linguistic variation.

This may involve supporting community-led efforts to document and digitise linguistic varieties on their own terms, bearing in mind not all linguistic varieties should be digitised or documented.

Collaboration across disciplines is also important. It requires linguists, technologists, educators and community leaders working together to ensure AI development is grounded in principles of linguistic justice.

The goal is not to “fix” language but to create technology that produces just outcomes. The focus should be on changing the technology, not the speaker.

### **Embracing Englishes**

English has been a powerful vehicle of empire, but it has also been a tool of resistance, creativity and solidarity. Around the world, speakers have taken the language and made it their own. AI-enabled systems should be built to be as inclusive of this variability as possible.

So next time your phone tells you to “correct” your spelling, or an AI chatbot misunderstands your phrasing, ask yourself: whose English is it trying to model? And whose English is being left out?

## **PART 1 - Reading Comprehension**

**Answer the following questions in less than 50 words:**

1. What percentage of AI training data comes from English?
2. Which variety of English dominates AI systems, and why?
3. What are the consequences of using only one English variety in AI?
4. Why is mainstream American English so dominant in tech products?
5. What did the study about voice technologies reveal?
6. What example did John Baugh's research provide?
7. How does linguistic bias in AI affect global users?
8. Why is it incorrect to say there is one “good” English?
9. What happens to non-standard varieties of English in AI training?
10. What changes does the author suggest for creating more inclusive AI systems?

**PART 2 – Summary**

**Write a 150-200 (10% margin) word summary of the text.**