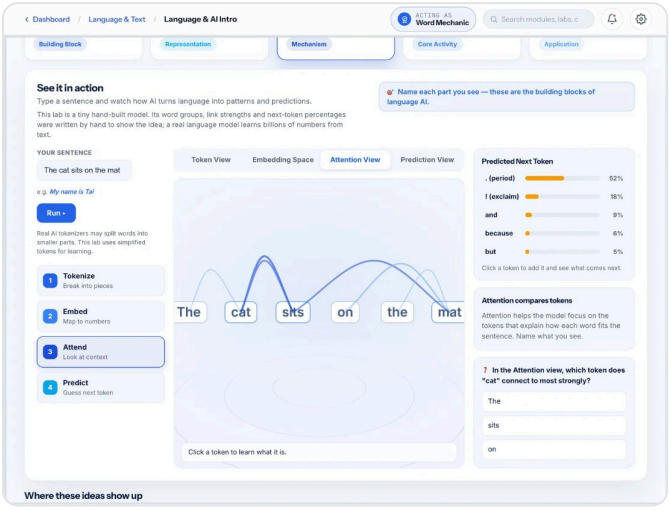
Module 03 · Language & Text · Unit 01 · Stage 2, Capabilities

Language & AI Intro

Key Ideas You'll Use

Grades 7–10 | 12 minutes | Role: **Word Mechanic**

Grade 7: About age 12–13 · Year 8 in the UK · Class 7 in India



The big idea

A language model does not read the way we do. It cuts text into tokens, turns each token into numbers that place similar meanings close together, uses attention to weigh which tokens matter to each other, and then predicts a likely next token. A high percentage means a common pattern in its training text, not a true fact.

The learner's job

Learners are the **Word Mechanic** at Bureau of Plain Language · Workshop. Everyone in this building uses these machines. Almost nobody has looked inside one. Their task: take a language machine apart. Name every piece.

By the end, learners can say

- I can define tokens and tokenization
- I can explain embeddings as number-based representations of meaning
- I can describe attention as a way to compare relationships between tokens
- I can explain next-token prediction as the core behavior of language models
- I can identify real-world NLP applications such as chatbots, translation, summarization, and sentiment analysis

Before class

- One computer or tablet for each learner or pair, with EdSutra open (this Module is part of the full EdSutra studio).
- Open the studio once on each device before class, so the Unit still runs if the internet drops.
- Play the Remember and Understand steps yourself, so you know where learners type and click.
- Write "Sentence → Tokens → Embeddings → Attention → Prediction" across the board. Learners will point to each stage as they go.
- Best after What is AI? and How Machines Learn.
- Units open in order, and Module 1 comes first. To start a class here, switch on Open every Unit in Settings → Teachers & parents.

Words to introduce

Token: A small piece of text: a word, part of a word, or punctuation.

Tokenization: Splitting a sentence into tokens before the model can use it.

Embedding: A list of numbers for a token; similar meanings sit close together.

Attention: How the model weighs which other tokens matter for each token.

Next-token prediction: Choosing a likely next token, one at a time, to build text.

Probability: How likely each next token is rated, shown as a bar and percentage.

Running the lesson 21 minutes

Open 4 min

"Almost everyone uses a language machine every day, from phone keyboards to chatbots. Almost nobody has looked inside one. What happens to your sentence first?"

Before anyone clicks, write "The cat sits on the ___" on the board. Ask: which word will the machine pick, and how sure will it be? Write the guesses down and come back to them in Evaluate.

Step	What learners do	Ask the class	Listen for
1. Remember 3 min	Learners read the five idea cards, then answer the Token View check: how many tokens is "The cat sits on the mat" split into?	• What is a token? Can a token be smaller than a word? • Why does a machine need to cut text into pieces first?	Six tokens here. A token is one piece of text; real tokenizers can also split words into smaller parts.
2. Understand 3 min	In Embedding Space each token becomes a point, with faint extra words like dog and chair for comparison. Learners drag to turn it and find the word closest to "cat".	• Which words ended up near "cat"? What do they have in common? • Where do little words like "the" and "on" sit?	Words with similar meanings sit close together: animals near animals, places near places. Closeness in numbers stands for closeness in meaning.
3. Apply 4 min	Learners type their own sentence and press Run, or click a predicted token to add it. They watch their sentence change in each of the four views.	• What did the lab do with a word it had never seen before? • Did the predicted next tokens change when your sentence changed?	Every sentence takes the same four steps. This tiny hand-built lab knows few words; real models learn from huge amounts of text.
4. Analyze 3 min	In the Attention View, curved lines link the tokens; the thickest, darkest line is the strongest link. Learners answer the check: which token does a word, such as "cat", connect to most strongly?	• Why would "cat" link most strongly to "sits", not to "the"? • How is being close in Embedding Space different from a strong attention link?	Attention shows which words help explain each other in this sentence, like who sits. Embeddings show which words mean similar things.
5. Evaluate 4 min	In the Prediction View, learners find the longest bar, then answer the judgment question: after "The cat sits on the", the lab gives "mat" 82%. What does that number tell you?	• Does 82% mean the cat really is on a mat? • Look back at our guesses on the board. Did we pick the same word as the machine?	The number is about word patterns, not the world. A likely word can still be wrong.

Steps follow Bloom's revised taxonomy, in the order the Unit presents them.

Deepen the thinking

Critical thinking	A chatbot answers in smooth, confident sentences. Using what you saw in the Prediction View, why might a fluent answer still be wrong?
AI thinking	If a model only ever picks the single most likely next token, what would its writing be like? When might you want a less likely word?
Ethics and fairness	A model learns word patterns from the text it was trained on. If most of that text came from one country or group, who might it serve less well?
Connect to life	When you read a word you don't know, how do the words around it help you guess its meaning? How is that like attention?

Misconceptions to listen for

"AI reads a sentence the way a person does."

The Token View shows it first cuts text into pieces and turns them into numbers. Then it predicts likely next pieces from patterns.

"A high percentage means the answer is true."

The Evaluate question shows 82% is about which word usually comes next, not about the world. A likely word can still be wrong.

"Embeddings are just a list of words."

They are positions in a space of numbers. In Embedding Space, words with similar meanings, like cat, dog and pet, sit close together.

In real life the card learners see at the end of the Unit

- Your phone keyboard's next-word suggestions are a small version of next-token prediction.
- Chatbots and translation apps split your text into tokens: small pieces like words or parts of words.
- When you hear "bat", nearby words tell you which kind. Attention helps language AI weigh those words too.

Try it this week. Say the start of a sentence and stop before the last word. Ask three people to finish it. Did they all pick the same word?

Watch out for. A language model doesn't read like you do. It predicts likely next tokens from patterns, so a fluent answer can still be wrong.

Talk about it. If a machine can finish your sentences, does it know what you mean? How could you tell?

Exit ticket

1. Name the four steps a sentence goes through in this lab, in order.

   _____

2. What does it mean when two words sit close together in Embedding Space?

   _____

3. The lab gives "mat" 82%. Does that make the sentence true? Explain.

   _____

Going further

The AI Story Canvas Lab lets learners build a choose-your-own-adventure story one choice at a time. Ask them how a person picking the next scene differs from a model predicting the next token. For homework, use the "Try it this week" activity.