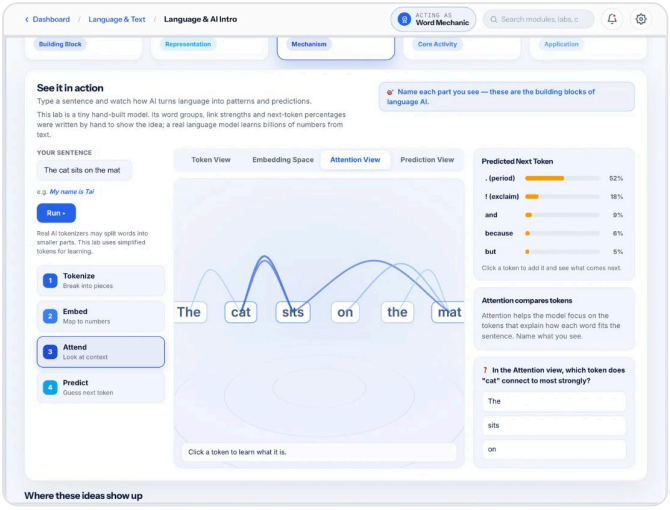
Module 03 · Language & Text · Unit 01 · Stage 2, Capabilities

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.

EdSutra

Teaching guide

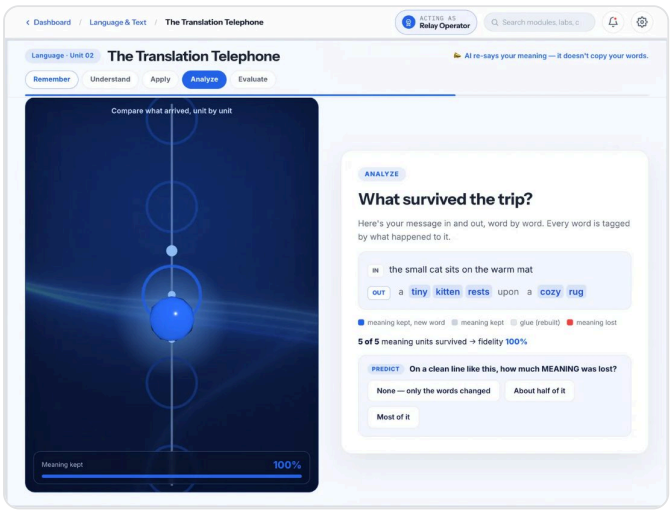
Module 03 · Language & Text · Unit 02 · Stage 2, Capabilities

The Translation Telephone

How words become meaning — and back

Grades 7–10 | **20 minutes** | Role: **Relay Operator**

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



The big idea

When AI rewrites or passes on a message, it does not copy your words. It keeps the meaning, the who, did what, to what and where, and says it again in new words. On a clear message the meaning survives completely. Meaning breaks when a message can be read two ways and the machine picks the wrong one.

The learner's job

Learners are the **Relay Operator** at Bureau of Plain Language · Relay Room. Someone sent a sentence through this room. What arrived was not what they said. Their task: send a message down the line. Find out what fell off.

By the end, learners can say

- I can name the journey a message takes: encode, whisper chain, decode
- I can explain that meaning is preserved while the exact words change
- I can compare an input and output and count how much meaning survived
- I can find an ambiguous message's breaking point and judge whether its meaning survived

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 first two steps yourself; the Unit scrolls down the page. The line is a small model built for this lesson, as it says.
- Write two headings on the board, "Meaning words" and "Glue words". You will sort words under them in Understand.
- Best after Language & AI Intro.
- 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

Meaning unit: A word that carries meaning: who, did what, to what, where, which kind.

Glue word: A small word like the, on or a, which the model rebuilds freely.

Encode: Turn words into their meaning, dropping the exact wording.

Decode: Build a new sentence, in fresh words, from the meaning.

Fidelity: Meaning units that survived out of the total, as a percentage.

Ambiguity: When a message can honestly be read in two different ways.

Paraphrase: Saying the same meaning in different words.

Running the lesson 25 minutes

Open 4 min

"In the game of Telephone, a message changes as it passes along. When an AI passes on your message, does it keep your words, your meaning, or both?"

Write "I saw her duck" on the board. Ask: what are two things this could mean? If a machine passed it on in new words, which meaning would arrive? Come back to it in Evaluate.

Step	What learners do	Ask the class	Listen for
1. Remember 3 min	Learners read how the line works, with its three stations (Encode, Whisper chain, Decode), and predict what actually travels down the line.	• What travels down the line: the words, or what they mean? • What happens at Encode, and what happens at Decode?	The meaning travels. Encode keeps the meaning and drops the wording; Decode builds new words from it.
2. Understand 4 min	For "the small cat sits on the warm mat", learners tap the words the AI must keep, then press Check my picks and compare what went in with what came out.	• Which words carry the meaning, and which are just glue? • The line said "a tiny kitten rests upon a cozy rug". Is that wrong?	Five meaning words (small, cat, sits, warm, mat) came back in new words. Glue words like "the" and "on" were rebuilt.
3. Apply 5 min	Learners pick a postcard or type their own message, tap the meaning words they predict must survive, and send it. The meaning capsule shows what was kept.	• Did your predicted meaning words match the meaning capsule? • If you typed your own message, which words did the line treat as glue?	Each kept unit has a job label (who, did what, where). The words out are new; the meaning is the same.
4. Analyze 4 min	Learners read their message in and out, word by word, with each word colored by what happened to it. They count the units that survived and predict how much meaning was lost.	• How many meaning units survived? What is the fidelity? • If every word changed, how can no meaning be lost?	All units survived, so fidelity is 100%. Different words are not wrong; only lost meaning counts.
5. Evaluate 5 min	Learners pick an ambiguous message ("I saw her duck" or "the chicken is ready to eat"), predict what happens, then drag the fuzziness dial to find where it flips and judge whether the meaning survived.	• At what point on the dial did the message flip to the wrong meaning? • Why did this message break when the cat message never did? • Look at our two meanings on the board. Which one arrived, before and after the flip?	It breaks because the message has two readings. Past a tipping point (40 for duck) the line picks the wrong one.

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

Deepen the thinking

Critical thinking	When 2 of 3 units survived, the line called it "Mostly". When can losing one unit matter more than the percentage suggests?
AI thinking	How would you rewrite "I saw her duck" so no machine could misread it? Which words would you add?
Ethics and fairness	A doctor's note or a school rule is passed through an AI rewriter and one word is misread. Who is responsible for checking the meaning arrived?
Connect to life	When you retell a story or a movie to a friend, which parts do you keep exactly, and which do you say in your own words?

Misconceptions to listen for

"The AI stores and repeats my exact words."
Understand shows the words coming back new: "small cat" becomes "tiny kitten". It keeps the meaning units and builds fresh words.

"Different words mean the AI got it wrong."
Analyze counts the units: on a clean line all of them survive, 100%. Same meaning is what matters, not same words.

"When AI misunderstands, it is just random."
Evaluate shows the break happens at an ambiguous word, at a tipping point you can find. Clear messages did not break on this line.

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

- When a translation app rewords your message, new words can still carry exactly the same meaning.
- An AI rewrite of a sentence may use all-new words. What matters is whether the meaning survived.
- Headlines and signs can be read two ways. When a message is ambiguous, AI can pick the wrong meaning.

Try it this week. Play telephone with friends or family using a sentence with two meanings, like "I saw her duck". Which meaning survives to the end?

Watch out for. Different words don't mean the AI got it wrong; lost meaning does. Meaning usually breaks at ambiguous words, not at random.

Talk about it. Have you ever been misunderstood because your words could mean two things? What happened?

Exit ticket

1. What travels down the line: your words or your meaning?

   _____

2. Name one meaning word and one glue word in "the happy child ran to the big park".

   _____

3. Why did "I saw her duck" break when the cat message did not?

   _____

Going further

For homework, use the "Try it this week" telephone game. Learners can also rewrite one sentence in all-new words, swap with a partner, and count how many meaning units survived.

EdSutra

Teaching guide

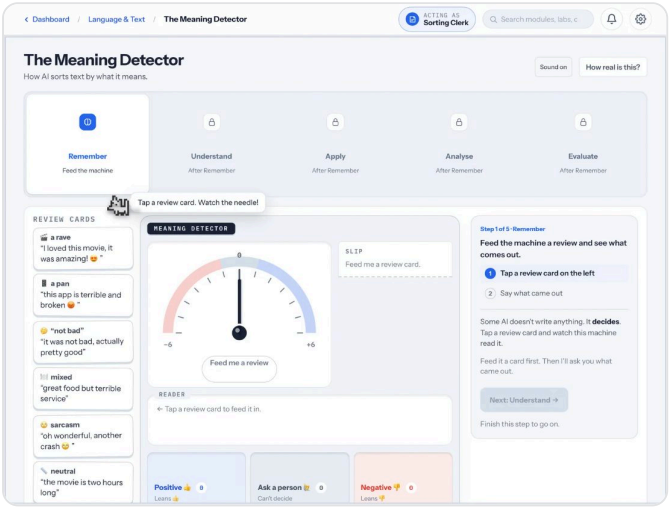
Module 03 · Language & Text · Unit 03 · Stage 2, Capabilities

The Meaning Detector

How AI sorts text by what it means

Grades 7–10 | 20 minutes | Role: **Sorting Clerk**

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



The big idea

Not all AI writes. A classifier reads text and makes a call: a label and how sure it is. The Meaning Detector adds up the weights of cue words, so it is easy to explain, and easy to fool: sarcasm tricks it, and mixed reviews leave it unsure. A trust bar sends shaky calls to a person.

The learner's job

Learners are the **Sorting Clerk** at Bureau of Plain Language · Sorting Hall. Every message in this hall goes to a person. The wrong bin means the wrong person, or nobody. Their task: find out which words decide the bin. Then find who gets misfiled.

By the end, learners can say

- I can name what a classifier outputs: a label and a confidence, not new text
- I can explain that the label is a sum of weighted cue words
- I can scan a review and predict its label from the cues
- I can find what fools the detector — negation, sarcasm, mixed signals
- I can judge when to trust a call versus send it to a human

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 first two steps yourself, and read "How real is this?" at the top: it is a real, simple cue counter.
- The machine makes sounds. Use headphones, or press the Sound button at the top of the Unit to switch them off.
- Best after Language & AI Intro and The Translation Telephone.
- 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

Classifier: AI that sorts text into a fixed label instead of writing new text.

Sentiment: Whether a piece of text sounds positive or negative.

Cue word: A word or emoji that leans positive or negative, like "amazing".

Weight: How strongly a cue leans: from +2 to −2 in this machine.

Confidence: How sure the call is: how far the total sits from zero.

Negation: A word like "not" that flips the next cue.

Trust bar (threshold): The confidence a call needs before its label is trusted without a person.

Running the lesson 27 minutes

Open 4 min

"Shops get thousands of reviews a day, and nobody can read them all. How could a machine sort them into happy and unhappy, without understanding any of them?"

Before anyone clicks, write "oh wonderful, another crash" on the board. Ask: is this review positive or negative? What would a machine that counts happy and sad words say? Come back to it in Analyze.

Step	What learners do	Ask the class	Listen for
1. Remember 3 min	Learners tap a review card. A scan light reads it word by word, the needle moves, and the card drops into a bin. Then they say what came out.	• Did the machine write anything new? What did it give you instead? • What are the three bins, and why is one of them "Ask a person"?	It gives a label and how sure it is. It writes nothing new; that is classification.
2. Understand 4 min	On "I loved this movie, it was amazing!", learners tap the words they think the machine counts, then press Check my picks and watch the needle step once for each cue.	• Which words did the machine count? Which did it ignore? • Add up the weights on the slip. Does the total match the label?	Three cues, "loved", "amazing" and the heart-eyes emoji, each +2, add up to +6: positive and fully sure.
3. Apply 5 min	Learners pick a review and guess its bin before the machine reads it: Positive, Unsure (ask a person) or Negative. After that, they type their own review and watch the needle.	• Which cue words told you which bin it would pick? • In your own review, which single word moved the needle most?	A guess backed by the cue words and their weights, checked against the words that lit up on the slip.
4. Analyze 6 min	Learners pick tricky reviews ("not bad", sarcasm, mixed, neutral) and guess what the machine will do. Then they try to write a review it gets wrong, and say what they meant.	• Why did "oh wonderful, another crash" land on Positive? • Why is the machine not very sure about "great food but terrible service"? • Look back at our board. Did we predict the sarcasm trap?	It reads words literally, so it can't hear sarcasm. Praise and complaint nearly cancel, so the total sits near zero.
5. Evaluate 5 min	Learners set the trust bar, then decide for incoming reviews: trust the label, or send it to a person. Swapping the labels turns the same machine into a spam filter.	• Where would you set the bar for a spam filter? For a safety filter? • The sarcastic review was low confidence. What if a sarcastic review were very confident?	Higher stakes need a higher bar. A confident wrong call still slips past, so the bar alone is not enough.

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

Deepen the thinking

Critical thinking	The machine was 100% sure about the movie review. Does 100% sure mean 100% right? What does the number really measure?
AI thinking	Real classifiers learn their cue weights from thousands of examples. Where might those examples come from, and what blind spots could they bring?
Ethics and fairness	A comment filter uses cue words to block unkind messages. Whose way of speaking, such as slang or another dialect, might get blocked unfairly?
Connect to life	Think of a message a friend once read the wrong way. What made it hard to read? Would a cue-counting machine have done better?

Misconceptions to listen for

"AI understands tone and sarcasm like a person."
Analyze shows "oh wonderful, another crash" landing on Positive. The machine reads cue words literally; it cannot hear what you meant.

"A negative word always makes a review negative."
"Not bad, actually pretty good" lands on Positive: "not" flips the next cue. A word's meaning depends on the words around it.

"More cue words always make the machine more sure."
Confidence is how far the total sits from zero. "Great food but terrible service" has two cues but low confidence, because they cancel.

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

- Shops and review sites can sort thousands of customer reviews into positive or negative automatically.
- Comment filters on games and video sites label messages safe or unsafe, and sarcasm can fool them.
- TV reports on how "people online feel" can come from tools that count positive and negative posts.

Try it this week. Write a one-line snack review that a word-counting detector would get wrong, using "not" or sarcasm. Would a friend get it right?

Watch out for. A detector counts cue words; it doesn't get tone. "Oh wonderful, another crash" can score positive. Confidence only shows how one-sided the cues are.

Talk about it. When should a computer's yes-or-no call about a message be checked by a person?

Exit ticket

1. What two things does a classifier give you?

2. Why did the sarcastic review fool the machine?

3. When should a call go to a person instead of being trusted?

Going further

For homework, learners write five one-line reviews of a snack, game or movie, score each cue word by hand (+2 to −2), and mark which ones a person should check.

EdSutra

Teaching guide

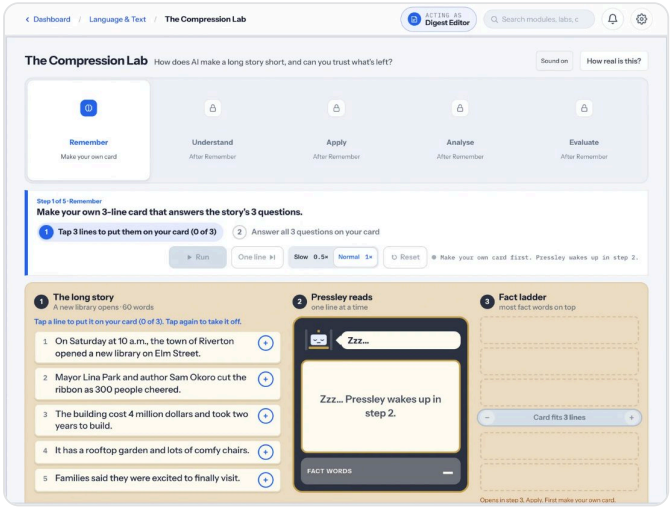
Module 03 · Language & Text · Unit 04 · Stage 2, Capabilities

The Compression Lab

How AI shortens a story without changing what it says

Grades 7–10 | 12 minutes | Role: **Digest Editor**

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



The big idea

A good summary keeps every key fact and invents nothing. Pressley, a digest press, counts the fact words in each line (names, days and times, numbers) and copies the lines with the most onto a small card. Squeeze too hard and a key point falls in the bin. Counting cannot tell which fact matters, so a person checks the card.

The learner's job

Learners are the **Digest Editor** at Bureau of Plain Language · Digest Desk. People will act on your summary without ever reading the original. Their task: cut it short. Then prove you did not change what it says.

By the end, learners can say

- I can explain what a good summary does: answer the key questions, add nothing new
- I can explain how an extractive summarizer chooses lines: it counts fact words
- I can squeeze a story to the shortest card that still keeps every key point
- I can explain why a fact-counting machine can drop the most important fact
- I can judge which summary is safe to send: faithful and complete

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 tap and where Pressley's lever is.
- Write the three race questions on the board: Who won? How fast? How many ran? You will check cards against them in Evaluate.
- Best after The Translation Telephone and The Meaning Detector.
- 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

Summary: A shorter version that keeps the key points and adds nothing new.

Extractive summary: A summary made by copying whole lines from the original.

Fact word: A word Pressley counts: a name, a day or time, or a number.

Key point: A fact the reader must get, such as "Who won?"

Faithful: Adds nothing the original does not say.

Complete: Keeps every key point.

Rewrite (abstractive summary): A summary in new words, which can add or change facts.

Running the lesson 24 minutes

Open 4 min

"People often act on a summary without ever reading the original. What could go wrong if the summary leaves out the one fact that matters?"

Before anyone clicks, ask: if a machine shortened a news story by keeping the sentences with the most names and numbers, would it always keep the most important fact? Take a show of hands and come back to it in Analyze.

Step	What learners do	Ask the class	Listen for
1. Remember 3 min	Learners read the library story and tap three of its five lines onto their own card. The card must answer What opened? Who opened it? What did it cost?	• Which two lines did you leave out, and why? • Did your card add anything the story does not say?	A good card answers every key question and adds nothing new. The rooftop garden and the excited families are filler.
2. Understand 4 min	Learners tap every fact word in line 1 (names, days and times, numbers) and check with Pressley. Then they guess which line goes on top and press Run to watch it count.	• Which words in line 1 are fact words? How many did you find? • Does Pressley know what a library is, or only what it can count?	Line 1 has four fact words. Pressley ranks lines by counting them; it does not know what they mean.
3. Apply 4 min	Learners pull the lever to shrink the card one line at a time, squeeze until a key point falls in the bin, then find the shortest card that keeps all three.	• How many lines did the shortest complete card need? • Which key point fell in the bin first, and why that one?	For the library, three lines keep all three key points. With fewer, the key line with the fewest fact words drops out.
4. Analyze 5 min	On the marathon story, learners guess whether the winner makes a three-line card, run Pressley, and watch "Who won?" fall in the bin. They say why, then give the card more room.	• Why did Pressley keep the weather line and drop the winner? • Look at our show of hands. Were we right about the machine?	The weather line is stuffed with numbers, so it out-counts the winner's line. Pressley counts facts; it cannot tell which one matters.
5. Evaluate 4 min	Learners make their own three-line race card, then choose which card goes to busy readers: Pressley's three lines, Pressley's four, a chatbot's rewrite, or theirs. They back the choice with a fact.	• The chatbot's card is shortest and reads well. Why can't it go out? • Check each card against the three questions on the board. Which pass?	Send a card that is faithful and complete. The chatbot adds "Sunday" and 12,000 runners, which the story never says.

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

Deepen the thinking

Critical thinking	The chatbot's card sounded smooth and sure, but added two facts. How would you check an AI summary of a long article before you trust it?
AI thinking	Pressley counts names, days and numbers. What other rule could help it find the line that matters most, and how might that rule fail?
Ethics and fairness	A summary of a news story leaves out one side's point of view because it had fewer facts to count. Who is harmed, and who should notice?
Connect to life	When you tell a friend what happened at school today, what do you keep and what do you leave out? How do you decide?

Misconceptions to listen for

"The machine knows which fact matters most."
Pressley counts fact words. In the race it keeps the weather line, full of numbers, and drops the winner.

"A shorter summary is always a better one."
In Apply, one line too few and a key point falls in the bin. The shortest good card is the shortest complete one.

"A smooth AI summary must be right."
The chatbot's race card reads well but invents a day and a runner count. Check every fact against the original.

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

- Search pages and apps now show AI summaries of articles, emails and reviews.
- A TV news roundup squeezes a long story into a minute, so someone chose which key points stay.
- Some study apps turn a chapter into flashcards by pulling out names, dates and numbers. That's extraction.

Try it this week. Summarize a TV story or book chapter in three sentences. Then check: did you keep every key point, and add nothing new?

Watch out for. A shorter summary isn't always better. Trust one only when it invents nothing and keeps every key point; AI rewrites can add facts.

Talk about it. When is reading only the summary fine, and when do you need the whole thing?

Exit ticket

1. What two things make a summary safe to send?

2. How does Pressley decide which lines to keep?

3. Why did the winner of the race fall in the bin?

Going further

The AI Story Canvas Lab lets learners build a choose-your-own-adventure story. Have them write a three-line card for each ending that answers who, what happened and how it ended, and invents nothing. Or use the "Try it this week" activity.

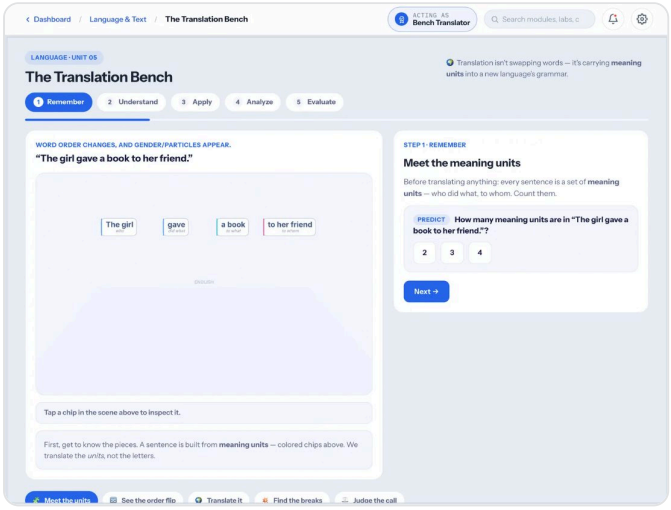
Module 03 · Language & Text · Unit 05 · Stage 2, Capabilities

The Translation Bench

How meaning moves across languages

Grades 7–10 | **30 minutes** | Role: **Bench Translator**

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



The big idea

Translation is not swapping words. A sentence is a set of meaning units (who did what, to whom), and translating carries those units into another language's grammar: the order can flip, and words gain gender or small grammar markers. Idioms and culture words can break or have no match, so a person checks anything important.

The learner's job

Learners are the **Bench Translator** at Bureau of Plain Language · Translation Bench. Two languages. One message. They do not carry the same things. Their task: move the meaning across. Say what could not come with it.

By the end, learners can say

- I can explain that translation preserves meaning units, not individual words
- I can describe how word order, gender, and particles differ across languages
- I can translate a sentence and predict which meaning units survive
- I can analyze why idioms and culture-specific words break in translation
- I can judge when a translation is faithful enough to ship vs. needs a human

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. Each prediction reveals the translation; tap a chip in the 3D bridge to read its note.
- If any learners speak Spanish, Hindi, Japanese or another language, invite them to check the Unit's translations and add their own.
- Best after Language & AI Intro and The Translation Telephone.
- 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

Meaning unit: One piece of meaning with a job: who, did what, to what, to whom.

Word order: The order a language puts its units in; Hindi and Japanese put verbs last.

Gender (in grammar): Words marked masculine or feminine, like Spanish la niña and un libro.

Particle: A small marker showing a word's job, like Japanese "o" for the object.

Idiom: A phrase whose meaning isn't its words, like "break a leg".

Borrowed word: A word taken from another language when there is no native one.

Fidelity: Meaning units that survived out of the total, like 1 of 2.

Running the lesson 30 minutes

Open 5 min

"If you swap every English word for its partner in Japanese, in the same order, will a Japanese reader understand the sentence?"

Write "Break a leg!" on the board. Ask: what does it mean, and what would it say if you translated it word for word? Write the guesses down and come back to them in Analyze.

Step	What learners do	Ask the class	Listen for
1. Remember 4 min	Learners see "The girl gave a book to her friend." split into colored chips on the bench and predict how many meaning units it has.	• What job does each chip do: who, did what, to what, to whom? • Why is "to her friend" one unit, not three words?	Four units: the girl (who), gave (did what), a book (to what), to her friend (to whom).
2. Understand 5 min	Learners predict where the verb "gave" goes in Japanese, then watch the chips fly across the 3D bridge into Japanese order. Tapping a chip shows its note.	• Where did the verb end up? Did any unit get lost? • Look at the small markers like "wa" and "o". What job do they do?	Japanese puts the verb last (Subject, Object, Verb). Every unit survives; only the order and the markers change.
3. Apply 6 min	Learners pick a sentence ("The girl gave a book..." or "The red car is fast.") and a language (Spanish, Hindi or Japanese), predict whether every unit survives, then reveal.	• What changed in Spanish that did not change in English? Look for gender. • Which language borrowed the English word "car"?	Plain sentences keep every unit. Only the surface changes: order, gender (la niña), markers, or a borrowed word.
4. Analyze 6 min	Learners test two hard cases, the idiom "Break a leg!" and the word "privacy", switching between Spanish, Hindi and Japanese to see which language maps, breaks or borrows each.	• What does "Break a leg!" become word for word in Hindi? • Why do Spanish and Japanese use a different phrase instead? • Check our guesses on the board. Were we right?	Word for word, Hindi gets "break your leg", which sounds like a threat. Spanish and Japanese use their own phrase.
5. Evaluate 4 min	"Privacy matters." goes into Japanese. "Privacy" has no native word, so it is borrowed, and the meter reads 1 of 2. Learners decide: ship it, or send it to a human translator.	• What does 1 of 2 tell you about this translation? • Which needs a human translator: a game chat, a menu, or a medicine label? Why?	One unit did not truly translate, so meaning may be lost. When fidelity is low, a human translator checks before it ships.

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

Deepen the thinking

Critical thinking	A translation app gives you a smooth sentence in a language you cannot read. How could you check whether the meaning survived?
AI thinking	The bench uses a small, hand-built phrase bank. A real translation AI learns from many example translations. Would it still break "Break a leg!"? Why or why not?
Ethics and fairness	Some languages have far less written text online than English. How might that affect translation apps, and who could be left out at a hospital?
Connect to life	Think of a saying your family uses. How would you explain it to someone from another country so the meaning, not just the words, arrives?

Misconceptions to listen for

"Translation is swapping each word for its foreign partner."
Understand shows the units reorder: in Japanese and Hindi the verb moves to the end. Translation carries meaning units, not single words.

"Any phrase can be translated word for word."
Analyze shows "Break a leg!" turning into a threat in Hindi. Idioms need an equivalent phrase, or they break.

"Every idea has a word in every language."
Japanese has no native everyday word for "privacy", so it borrows the English one. Some words have no slot to land in.

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

- Translation apps don't swap word for word: in Hindi or Japanese the verb moves to the end.
- Signs in several languages at hospitals and stations need careful translation, because a wrong word there really matters.
- Auto-translated captions can turn an idiom like "break a leg" into nonsense, or even a threat.

Try it this week. Ask someone who speaks another language how they'd say "it's raining cats and dogs". Is it word for word, or a different idiom?

Watch out for. Translation isn't swapping words. Idioms and culture words can break or have no match, so a person should check anything important.

Talk about it. Is there a word in a language you know that has no good English match?

Exit ticket

1. Where does the verb go in a Japanese or Hindi sentence?

2. Why did "Break a leg!" break in Hindi?

3. The meter reads 1 of 2. Ship it, or ask a human? Why?

Going further

For homework, use the "Try it this week" activity. Learners can also ask a family member to translate "The girl gave a book to her friend" into another language and mark where the verb goes.

EdSutra

Teaching guide

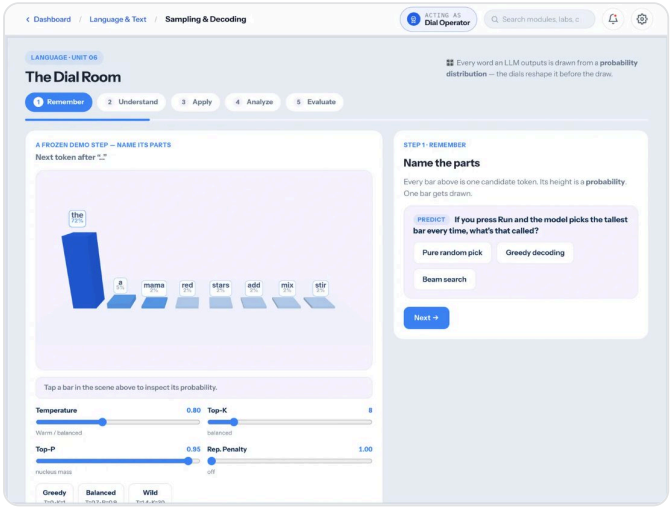
Module 03 · Language & Text · Unit 06 · Stage 2, Capabilities

Sampling & Decoding

The Dial Room

Grades 8–12 | 40 minutes | Role: **Dial Operator**

Grade 8: About age 13–14 · Year 9 in the UK · Class 8 in India



The big idea

A language model does not know the one right next word. It gives every possible next word a probability, and a sampling rule picks one. Dials like temperature, Top-K, Top-P and a repetition penalty reshape those odds before the pick. Careful settings give safe, repetitive text; loose settings give lively, riskier text. The job decides which is right.

The learner's job

Learners are the **Dial Operator** at Bureau of Plain Language · Dial Room. This machine can be careful or surprising. Somebody has to choose which, and for what. Their task: learn what each dial does. Match the setting to the job.

By the end, learners can say

- I can explain how language models generate text token by token
- I can describe what temperature controls and its effect on output diversity
- I can compare top-K and top-P (nucleus) cuts, and use a repetition penalty to break loops
- I can design sampling configurations for different use cases (creative vs. factual)

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 first two steps yourself, so you know where the dials are and what the Run button does.
- Draw a bar chart on the board with five words and their chances. You will flatten and sharpen it by hand during Understand.
- Best after Language & AI Intro and The Meaning Detector.
- 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: One piece of text the model writes at a time; here, a whole word.

Probability distribution: The chance the model gives each possible next word; they add up to 100%.

Greedy decoding: Always picking the most likely next word. Same input, same output.

Temperature: A dial that sharpens (low) or flattens (high) the odds before the pick.

Top-K: Keep only the K most likely words, then pick among them.

Top-P: Keep the smallest group of top words whose chances add up to P.

Repetition penalty: Makes recently used words less likely, so the text stops looping.

Running the lesson 40 minutes

Open 4 min

"Ask a chatbot the same question twice and you can get two different answers. If it is the same machine, where does the difference come from?"

Before anyone clicks, ask: if we press Run three times with the same settings, will the machine write the same sentence each time? Take a vote, write it on the board, and come back to it in Apply.

Step	What learners do	Ask the class	Listen for
1. Remember 4 min	A frozen demo shows the candidate next words as bars; each bar's height is its probability. Learners tap bars to read them, then name the rule that always takes the tallest bar.	• What does the height of each bar tell you? • If the model always takes the tallest bar, what happens when you run it twice?	Each bar is a chance, not a certainty. Always taking the tallest is greedy decoding: same input, same output.
2. Understand 6 min	Learners drag the Temperature dial from low to high, watch the bars reshape live, then say what happens to them above 1.0.	• At low temperature, which word wins almost every time? • At high temperature, why do rare words start getting picked?	Low temperature sharpens the odds toward one word; high temperature flattens them, so unlikely words get a real chance.
3. Apply 8 min	Learners pick a mission and set the dials. Cold Start: three runs in a row come out exactly the same. Hot Take: three runs share no three-word phrase.	• Which dial did you move first, and why? • Look at our vote on the board. Does the machine always write the same sentence?	Only very sharp settings, like temperature near 0, repeat exactly. Otherwise each word is drawn, so runs differ.
4. Analyze 9 min	Learners predict which cut keeps a fixed number of words, Top-K or Top-P. Then, with Temperature at 0.4 or below, they raise the Repetition Penalty until no three-word phrase repeats.	• Why does Top-P keep fewer words when the model is confident? • Why not just raise the temperature to stop the loop?	Top-K keeps a fixed count; Top-P adapts. The penalty fixes loops without adding the randomness a hotter temperature brings.
5. Evaluate 9 min	Learners choose settings for a factual assistant, then hit the Goldilocks target: at least 10 words, 6 different words and no word twice in a row. A three-question quiz closes the Unit.	• Why does a factual assistant want low temperature and a small Top-P? • What would your Goldilocks settings be bad at?	A choice tied to the job and its cost: safe settings can be dull; lively settings can be wrong.

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

Deepen the thinking

Critical thinking	Two people ask a chatbot the same question and get different answers. What would you need to know before deciding which answer to trust?
AI thinking	The Dial Room's model learned from 30 short sentences. What would change if it learned from millions, and what would stay the same?
Ethics and fairness	A city uses an AI with a high temperature to write its safety notices, so they sound lively. Who could be harmed, and who should choose the setting?
Connect to life	Think of a music playlist on shuffle versus repeat. Where else do you choose between surprise and the same thing every time?

Misconceptions to listen for

"A higher temperature makes the AI more creative or smarter."

Temperature only reshapes the odds. In Understand the bars flatten; the model knows nothing new, so rare words, right or wrong, get picked more.

"Top-K and Top-P do the same thing."

Top-K always keeps exactly K words. Top-P keeps fewer when the model is confident and more when it is unsure, as Analyze shows.

"When the model repeats itself, it is broken."

Loops come naturally from sharp, low-temperature picks. In Analyze the Repetition Penalty makes recent words less likely and breaks the loop without adding randomness.

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

- Ask a chatbot the same question twice and answers may differ: each word is drawn from likely options.
- Some AI writing tools offer "creative" or "precise" modes, which can work a bit like a temperature dial.
- At a school raffle, a name with more tickets wins more often. AI picks next words much the same way.

Try it this week. Put word cards in a bag, with extra copies of likely words. Draw to build a silly sentence, then remove the rare words and draw again.

Watch out for. A higher temperature doesn't make AI smarter or more truthful; it only makes word picks less predictable. The right setting depends on the job.

Talk about it. For homework facts, would you want a predictable AI or a surprising one? What about story ideas?

Exit ticket

1. What does turning the temperature up do to the bars, and to the text?

2. Top-K and Top-P both cut words. Which one keeps a fixed number?

3. Choose settings for a factual helper and name one thing they cost.

Going further

In the AI Story Canvas Lab, learners build a choose-your-own-adventure story where each scene offers a few next steps. There the reader picks the path; in the Dial Room, weighted chance does. For homework, try the In real life card's word-bag activity.

EdSutra

Teaching guide

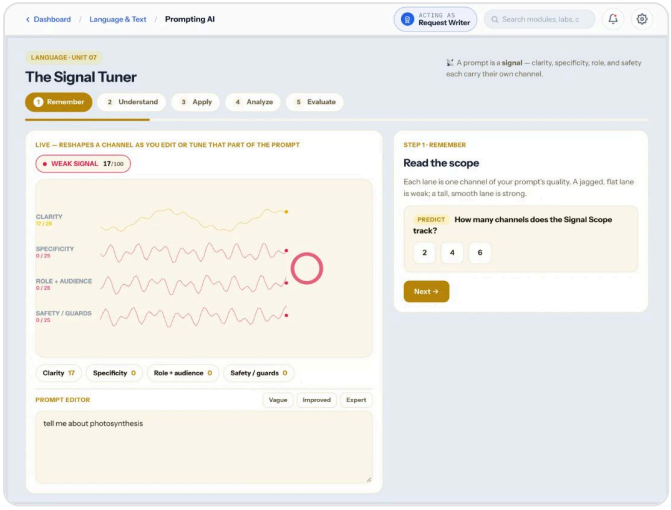
Module 03 · Language & Text · Unit 07 · Stage 2, Capabilities

Prompting AI

The Signal Tuner

Grades 6–12 | 40 minutes | Role: **Request Writer**

Grade 6: About age 11–12 · Year 7 in the UK · Class 6 in India



The big idea

A prompt's quality is not a feeling. The Signal Tuner splits it into four channels (clarity, specificity, role and audience, safety), each fixed by a specific kind of edit. Removing structure hurts far more than typos. The score is a heuristic: it matches words, so it cannot tell whether an answer would be true or useful.

The learner's job

Learners are the **Request Writer** at Bureau of Plain Language · Front Counter. Someone at the counter needs an answer. Their question will not get them one. Their task: turn a vague wish into a request that works. Show what changed it.

By the end, learners can say

- I can name the four features that make a prompt clear: clarity, specificity, role and audience, guardrails
- I can rewrite a vague prompt until it reaches an expert score
- I can predict and test which kinds of noise hurt a prompt most
- I can back your best change with evidence, and explain what a heuristic prompt score can't see

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 first two steps yourself, so you know where the prompt editor and the channel chips are.
- Write the starter prompt, "tell me about photosynthesis", on the board. The class will mark each upgrade on it during Apply.
- Best after Language & AI Intro and The Meaning Detector.
- 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

Prompt: The instruction or question you give an AI tool.

Clarity: Saying exactly what you want, with no hedges like "maybe" or vague verbs.

Specificity: Naming the format and length you want, such as "3 short paragraphs".

Role and audience: Who the AI should act as, and who the answer is for.

Guardrails (Safety / guards): Rules about what to avoid, and what to do if the AI is unsure.

Heuristic: A rule of thumb: quick pattern checks, not real understanding.

Noise: Damage to a prompt, such as typos, hedging or a missing format.

Running the lesson 40 minutes

Open 4 min

"People often know what they want but cannot say it. When an AI gives a useless answer, is it the machine's fault or the question's?"

Before anyone clicks, ask: which would hurt a good prompt more, a few typos or deleting its role and format lines? Take a vote, write it on the board, and come back to it in Analyze.

Step	What learners do	Ask the class	Listen for
1. Remember 3 min	Learners read the Signal Scope and answer how many channels it tracks: four, worth 25 points each. Jagged lanes are weak, smooth ones strong. Tapping a channel chip lists its issues.	• Which channels are weakest for "tell me about photosynthesis"? • What does the issue list say is missing?	It has no format, length, role, audience or guardrails, so three lanes sit at zero; only clarity scores.
2. Understand 5 min	Learners add a role, such as "You are a science teacher", to the starter prompt, predict which channel it fixes, and watch the Role + audience lane clean up.	• Which lane changed most, and which words did it react to? • Would adding "please" change the score? Try it.	A specific edit moves a specific channel. The role words matched what the Role + audience channel looks for.
3. Apply 10 min	Starting from "tell me about photosynthesis", learners type their own upgrades until the prompt reaches 70, then 90 (expert). Loading the Expert example does not count; only their own edits do.	• Which single edit bought the most points? • Compare yours with the Expert example. What does it have that yours does not?	A format and an audience lift the score fast; expert prompts also add length, guardrails and a plan for when unsure.
4. Analyze 8 min	Learners reset to the expert prompt, predict which noise hurts most, then break it on purpose: typos, hedging, a contradiction, no role, no format. They try at least three kinds.	• Which noise dropped the score most? Which barely moved it? • Look at our vote on the board. Were we right?	Typos are surface noise. Removing the role or format collapses whole channels, so structure matters more than polish.
5. Evaluate 10 min	Learners clear every high-severity issue, then defend their tuning by choosing: the channel that mattered most, the evidence for it, and what the scorer cannot see.	• Does a 100/100 prompt guarantee a good answer? Why not? • Why does the evidence have to match the channel you picked?	The scorer only spots surface features in the wording. It runs no AI and cannot check facts or answer quality.

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

Deepen the thinking

Critical thinking	The scorer rewards the words "do not". How could someone get a high score with a prompt that is still unhelpful?
AI thinking	Why might a real AI model give a better answer to the expert prompt than to "tell me about photosynthesis"? Use the four channels in your answer.
Ethics and fairness	If a bad answer is usually a bad question, who is to blame when someone still learning English gets a poor AI answer?
Connect to life	Think of a time a teacher, coach or parent misunderstood what you asked for. Which channel was missing from your request?

Misconceptions to listen for

"Typos are the worst thing that can happen to a prompt."
In Analyze, typos barely dent the score. Deleting the role or format lines collapses whole channels: structure matters more than spelling.

"A 100/100 prompt guarantees a perfect answer."
The score is a heuristic that checks wording only. It runs no AI, so it cannot check facts. Evaluate asks learners to name this limit.

"Whether a prompt is good is just a feeling."
The Scope splits quality into four named channels, each with words it looks for. Learners can see exactly which edit lifted which lane.

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

- Chatbots and homework helpers answer more usefully when a prompt names a format, like "three bullet points".
- Hedges like "maybe" or "kind of" in a request to an AI tool make the answer harder to aim.
- A good worksheet or recipe card is like a good prompt: a clear task, a format and limits.

Try it this week. With a parent or teacher, ask an AI tool one question twice: once vague, once with a role, audience and format. Compare.

Watch out for. A longer prompt isn't a better one. A missing format hurts more than a typo, and one lucky answer doesn't prove a prompt works.

Talk about it. Can a score tell you whether a piece of writing is good? What might it miss?

Exit ticket

1. Name the four channels the Signal Scope measures.

   _____

2. Which hurt the score more: typos, or removing the role and format? Why?

   _____

3. Name one thing the prompt scorer cannot see.

   _____

Going further

For homework, learners rewrite a request they made this week, to a person or an app, so it names a role, an audience, a format and one guardrail. Which channel changed most? Early finishers can build a branching story in the AI Story Canvas Lab.

EdSutra

Teaching guide

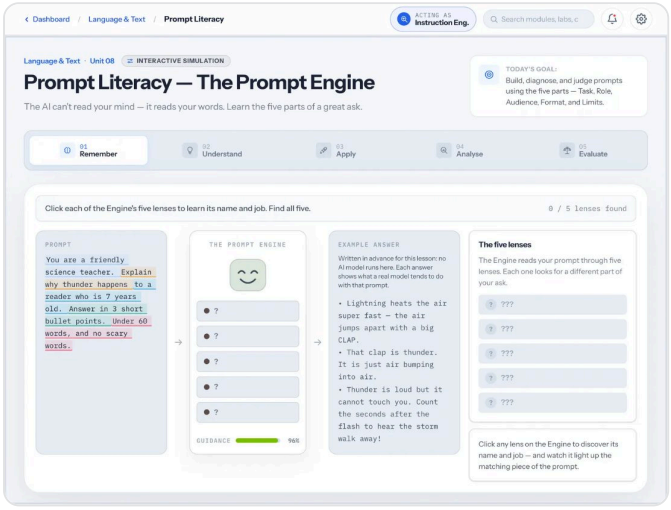
Module 03 · Language & Text · Unit 08 · Stage 2, Capabilities

Prompt Literacy

The Prompt Engine

Grades 7–11 | 45 minutes | Role: **Instruction Engineer**

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



The big idea

The AI cannot read your mind; it reads your words. A strong prompt has five parts: Task, Role, Audience, Format and Limits, and each one steers a different side of the answer. When an answer disappoints, the cause is usually a part that was vague, missing or fighting another.

The learner's job

Learners are the **Instruction Engineer** at Bureau of Plain Language · Engine Room. A good question you wrote once helps one person. A good instruction helps everyone who uses it after you. Their task: build an instruction that still works when you are not here.

By the end, learners can say

- I can name the five parts of a strong prompt and what each one controls
- I can explain how adding each part changes the same AI answer
- I can build complete prompts for real goals by choosing the right part for each slot
- I can diagnose why a prompt produced a bad answer — vague, missing, or conflicting parts
- I can judge competing prompts against a rubric and choose which one to ship

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 first two steps yourself, so you know how the lenses light up and what the Guidance meter shows.
- Tell learners up front: the Engine's answers were written in advance. No AI model runs; they show what real models tend to do.
- Best after Language & AI Intro and Prompting AI.
- 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

Task: The action you want, such as explain, list or compare.

Role: Who the AI should act as; it sets the voice.

Audience: Who the answer is for; it sets how hard the words are.

Format: The shape of the answer: a list, steps, a table, a story.

Limits: Rules the answer must follow: how long, what to include or leave out.

Guidance meter: The Engine's gauge of how clearly a prompt steers the answer.

Rubric: A checklist for judging things the same way each time.

Running the lesson 45 minutes

Open 5 min

"If an AI cannot read your mind, what exactly does it need from your words to give you the answer you pictured?"

Before anyone clicks, write "Tell me about black holes." on the board. Ask: what will the answer be like, and what would you add first to improve it? Come back to the guesses in Understand.

Step	What learners do	Ask the class	Listen for
1. Remember 4 min	Learners click each of the Engine's five lenses to learn its name and job, and find all five. Each lens lights up its part of a sample prompt about thunder.	• Which part of the thunder prompt sets who the answer is for? • What would the answer lose without the Limits line?	Each part has its own job: Task is the action, Role the voice, Audience the reader, Format the shape, Limits the rules.
2. Understand 7 min	Learners run the vague prompt "Tell me about black holes.", then add one part at a time: Task, Audience, Format, then Role and Limits. The answer and the Guidance meter change each time.	• Which upgrade changed the answer most, and how? • Look at our board guesses. Did the part we chose first help most?	Same topic, same AI; only the words changed. Each part steered one thing: aim, difficulty, shape, voice, length.
3. Apply 10 min	Two missions: calm a scared 7-year-old about thunder, and find five fast facts for a honeybee poster. Learners pick one chip per slot, run the Engine, and need a score of 4 out of 5.	• Why is "You are an AI assistant" a weak role? • A chip can be specific and still wrong for the job. Which one fooled you?	The best chip fits this goal and this reader. Specific is not enough: a lecture voice is specific but wrong here.
4. Analyze 10 min	Four broken prompts and their bad answers: an essay rewrite, a vaccine talk too hard for a 6-year-old, a pancake wall of text, an impossible summary. Learners name the part that failed.	• In "Make my essay better", better how? • The vaccine prompt had a doctor role. Why did the answer still fail? • Can any AI follow "three words" and "every chapter in detail" at once?	The fault is in the ask: a vague task, a missing audience, a missing format, or limits that fight.
5. Evaluate 9 min	The Prompt-Off: three rival prompts for one job, a volcano study sheet for tomorrow's quiz. Learners tick the rubric for each prompt, run them to see the answers, then ship the best one.	• Prompt B is friendly and clear. Why is it still not the one to ship? • What does Prompt C have that Prompt B is missing?	C fills all five parts. B lacks a format and limits, so it reads like a chat, not a study sheet.

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

Deepen the thinking

Critical thinking	The Engine's answers were written in advance to show what real models tend to do. How could you test whether a real AI behaves the same way?
AI thinking	The AI answered the vaccine question for an average adult because nobody said otherwise. Why would a model fall back on an "average" reader?
Ethics and fairness	Your instruction will be used by people you never meet. What could go wrong for a younger or nervous reader if you leave out the Audience?
Connect to life	Pick a set of instructions you follow often, like a recipe, game rules or a homework sheet. Which of the five parts does it use, and which is missing?

Misconceptions to listen for

"The AI gave a bad answer, so the AI is bad."

In the Autopsy, every bad answer traces back to the ask: a vague task, a missing audience or format, or limits that fight.

"The AI knows what I mean."

The Bedtime Backfire case never mentions the 6-year-old cousin, so the AI writes for an adult. It cannot use information you never gave it.

"Prompting is about typing tricks or magic words."

The five parts are plain, structured communication, the same skill as briefing a person well. The Prompt-Off rewards a complete prompt, not a clever one.

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

- Asking a voice assistant or chatbot for help works best with all five parts: task, role, audience, format, limits.
- Leaving notes for a babysitter or substitute teacher is prompting too: task, audience and limits all matter.
- In a group chat, "make slides" gets guesses, but "five slides for our class by Friday" gets what you need.

Try it this week. Write instructions for making a sandwich using all five parts. Have someone follow them exactly as written. What went wrong?

Watch out for. A bad answer doesn't always mean a bad AI. Check your ask first: it can't use information you never gave it.

Talk about it. Who is the best explainer you know? Which of the five parts do they always include?

Exit ticket

1. Name the five parts of a strong prompt.

2. "Make my essay better." Which part is weak, and how would you fix it?

3. What goes wrong when two limits contradict each other?

Going further

In the AI Story Canvas Lab, learners build a choose-your-own-adventure story and play every path, checking that a stranger could follow each choice. For homework, use the "Try it this week" activity from the In real life card.

EdSutra

Teaching guide

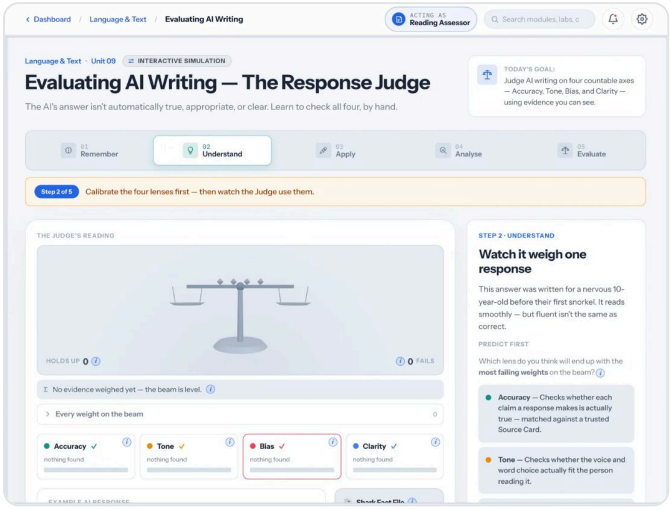
Module 03 · Language & Text · Unit 09 · Stage 2, Capabilities

Evaluating AI Writing

The Response Judge

Grades 7–12 | 25 minutes | Role: **Reading Assessor**

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



The big idea

Fluent writing can still be wrong. Any AI answer can be checked through four lenses: Accuracy (is each claim backed by a source?), Tone (does it fit the reader?), Bias (do loaded words slip in an opinion?) and Clarity (can the reader use it?). The counts support a ship, revise or reject call, but the decision stays human.

The learner's job

Learners are the **Reading Assessor** at Bureau of Plain Language · Review Panel. Two answers. One question. Somebody has to choose, in public. Their task: say which answer is better. Then show the evidence that decided it.

By the end, learners can say

- I can name the four judging lenses and the failure each one catches
- I can explain how a response's verdict is counted from evidence, not guessed
- I can tag claims, tone cues, loaded words, and clarity signals in a fresh response
- I can diagnose which lens a broken response fails, and cite the exact span that proves it
- I can weigh two rival responses and defend a ship / revise / reject call with evidence

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 first two steps yourself, so you know how a lens is calibrated and how the beam tips.
- Tell learners the answers were written in advance; no AI model runs. Wrong picks cost nothing and only bring a clearer hint.
- Best after Language & AI Intro and Prompt Literacy.
- 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

Accuracy: Whether each claim is true, checked against a trusted Source Card.

Tone: Whether the voice and words fit the person reading.

Bias (loaded words): Words that slip in an opinion as if it were a fact.

Clarity: How readable it is: short sentences and no unexplained jargon (specialist words).

Hallucination: A confident claim an AI makes up, with no source behind it.

Source Card: The trusted fact list each claim is checked against.

Verdict: The final call on a response: ship, revise or reject.

Running the lesson 30 minutes

Open 4 min

"An AI answer sounds confident, friendly and smart. How would you prove whether it is actually good, with evidence instead of a gut feeling?"

Before anyone clicks, read aloud: "Great whites can smell a single drop of blood from ten miles away." Ask: true or false, and how would you check? Take a vote and come back to it in Understand.

Step	What learners do	Ask the class	Listen for
1. Remember 3 min	Learners pick a failure card, then click the lens that catches it: Accuracy, Tone, Bias or Clarity. Four correct matches bring the balance scale online.	• Which lens catches an invented fact? • Which lens catches a correct answer that is too hard to read?	Each lens catches a different failure: wrong facts, wrong voice, loaded words, or text too dense to use.
2. Understand 4 min	Learners predict which lens will get the most failing weights on a shark answer for a nervous 10-year-old, run the scan, then open each lens to see its evidence alone.	• Look at our vote: was the blood-smell claim on the Source Card? • How is the beam's tilt worked out from the weights?	Every weight points at a phrase. The blood claim has no source, so it fails Accuracy; the tilt is counted, not guessed.
3. Apply 7 min	Learners tag phrases in at least two of four case files, such as the Great Wall facts. Their tags tip their own beam; then the real one swings in beside it.	• How far was your beam from the real one, and why? • Did you tag a phrase that was actually fine? What fooled you?	They point at exact phrases, such as the Great Wall's "spot it from space" myth, and name the lens each one fails.
4. Analyze 6 min	The Writing Autopsy: three broken answers. For each, learners name the failing lens and click the exact phrase that proves it. Two must be solved; wrong picks only bring a hint.	• In the Ada Lovelace answer, the fact is true. So what failed? • Why must you point at a phrase, not just name a lens?	A true fact can carry loaded words. A diagnosis counts only when it points at the exact phrase that proves it.
5. Evaluate 6 min	The Verdict: two water-cycle answers for a 5th-grade handout. Learners weigh both, fill the rubric, choose ship, revise or reject, and pick a defense that rests on evidence.	• Response A is accurate. Why might you still not ship it? • Someone chose "revise B". Can that be a fair call? Why?	A call holds when its reason rests on the counts. "B sounds friendlier" is not evidence; A's jargon count is.

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

Deepen the thinking

Critical thinking	The Judge counts evidence, but a person chose what goes on the Source Card. What could go wrong if the Source Card itself were wrong?
AI thinking	Why might an AI write a fluent, confident sentence that no source backs up? Think back to how it picks each next word.
Ethics and fairness	"Lazy students" and "clinging to outdated thinking" appeared in a summary of both sides. Who is hurt when a summary takes a side while sounding neutral?
Connect to life	Before you forward a message or share a post, which of the four lenses do you already use, and which do you skip?

Misconceptions to listen for

"If it reads confidently, it is true."
In Understand, the smooth shark answer still has a claim no source backs. Fluency and truth are different things; check the source.

"A biased answer always sounds obviously biased."
The Lovelace case wraps loaded words, "an unmatched genius" and "too small-minded", around a true fact in a calm, reasonable sentence.

"A correct answer is automatically a good answer."
In the Verdict, Response A is accurate and unbiased but fails tone and clarity for 5th graders. Correct is not the same as usable.

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

- AI-written text now shows up in search results, homework helpers and shop product pages.
- Newspapers, ads and leaflets can use loaded words like "obviously" that slip an opinion in as fact.
- A chatbot answer can be accurate but pitched wrong for its reader, or clear but unfair.

Try it this week. Pick a paragraph from an ad or leaflet. Check it for accuracy, tone, loaded words and clarity. Would you ship, revise or reject it?

Watch out for. Confident writing isn't proof it's true, and a correct answer can still fail on tone or clarity. Check the evidence, not the polish.

Talk about it. Have you read something that sounded certain but turned out wrong? How could you have checked?

Exit ticket

1. Name the four lenses and the failure each one catches.

2. What makes a word "loaded"? Give an example from the Unit.

3. Why is a correct answer not always ready to ship?

Going further

In the AI Story Canvas Lab, learners build a choose-your-own-adventure story. Partners then swap seats and judge each other's scenes through the four lenses, pointing at the exact phrase for each call. For homework, use the In real life card's leaflet check.