

PRODUCT · OCTOBER 2026

Languages

Two different lists: what Ovos Wasel can **hear** (recognise as the source of speech) and what it can **speak** (deliver to listeners as interpreted audio).

Hears – 15 languages

Language	Native name	Notes
English	English	
Arabic	العربية	Recognition is strongest on Modern Standard Arabic; Gulf, Egyptian, Levantine and Maghrebi dialects are recognised less accurately
Urdu	اردو	
Persian (Farsi)	فارسی	
Turkish	Türkçe	
French	Français	
German	Deutsch	
Spanish	Español	
Russian	Русский	
Italian	Italiano	
Chinese (Mandarin)	中文	Captions in simplified or traditional script (operator's choice)
Chinese (Cantonese)	粵語	
Hindi	हिन्दी	
Korean	한국어	
Japanese	日本語	

How the source language is handled

- **Automatic detection (default)**: each sentence's language is detected, so speakers can switch languages mid-meeting.
- **Expected languages**: the operator can tick the languages they know will be spoken. This removes whole classes of misrecognition; a sentence detected in a language that is not on the list is dropped rather than forced into the wrong one.
- **Fixed source**: detection can be turned off and one language fixed.

Speaks – 10 languages, any 4 at once

Language	Native name	Voice	Caption direction
Arabic (Modern Standard)	العربية	Neural voice by default; a standard voice as fallback	Right to left
English	English	Standard voice; neural voice selectable	Left to right
German	Deutsch	Standard voice; neural voice selectable	Left to right
French	Français	Standard voice	Left to right
Spanish	Español	Standard voice	Left to right
Italian	Italiano	Standard voice	Left to right
Russian	Русский	Standard voice	Left to right
Turkish	Türkçe	Standard voice	Left to right
Persian (Farsi)	فارسی	Standard voice	Right to left
Urdu	اردو	Standard voice	Right to left

"Neural" voices are the more natural-sounding option; "standard" voices are lighter voices – clear and intelligible, recognisably synthetic. The operator chooses per language under Settings → Models → Voices.

Why four at once: every active output language costs one translation and one voice on every sentence. Four keeps the room close to live on a single PC. The limit is enforced everywhere languages are chosen, with an explanation.

Why Chinese is heard but not spoken: no on-premise voice is offered for Mandarin or Cantonese yet, so neither is offered as an output – rather than offering it and producing silence. Speak Mandarin or Cantonese and the room hears it in their own languages. (The optional cloud profile can speak them.)

Arabic

- Output is Modern Standard Arabic, understood by educated speakers across dialects.
- Arabic is spoken slightly faster by default. Arabic renderings run a little longer than the original, and this keeps listeners close to live.
- Arabic, Urdu and Persian captions render right to left, in Noto Sans Arabic, on phones, tablets and the room display.

Glossary

A per-language glossary fixes names, ranks, units and designations so they are never paraphrased and never drift between meetings. One line per term:

Northern Line = ar:الخط الشمالي, fr:Ligne Nord, de:Nordlinie
lines starting with # are notes