

ORIGINAL RESEARCH ARTICLE

Cross-lingual data science methods to enhance cognitive embodied intelligence in English and Uzbek robotic communication

Supplementary File

1. Experimental configuration

1.1. Hyperparameters and training details

The models were trained using the AdamW optimizer. The full grid of hyperparameters searched, and the final selected values are provided in Table S1 below.

The final selection was based on performance on a held-out validation set (10% of the training data).

1.2. Participant demographics and bias analysis

The 120 participants in our user studies were recruited to ensure diversity. The demographics of the participants included gender (62% male, 38% female); age (20–30: 35%, 31–40: 40%, 41–50: 20%, 50+: 5%); and region (60% Tashkent, 40% other Uzbek regions). A Fisher’s exact test confirmed no significant association ($p>0.05$) between these demographic factors and task success rates, indicating no sampling bias in our evaluation cohorts.

Table S1. Hyperparameter search space and final values

Hyperparameter	Search space	Final value
Learning rate	$\{1 \times 10^{-5}, 3 \times 10^{-5}, 5 \times 10^{-5}, 1 \times 10^{-4}\}$	3×10^{-5}
Batch size	{16, 32, 64}	32
Warm-up steps	{500, 1,000, 2,000}	1,000
Training epochs	{10, 20, 30}	20
Dropout rate	{0.1, 0.2, 0.3}	0.1