

Figure 2: t-SNE visualization of the embedding space. The plot shows the distribution of human and AI-generated data points for Russian, Chinese, Japanese, and European languages. The x-axis represents t-SNE Dimension 1, and the y-axis represents t-SNE Dimension 2. The legend identifies the data series: Russian (human) (blue circle), Russian (ai) (blue triangle), Chinese (human) (brown square), Chinese (ai) (brown diamond), Japanese (human) (dark green diamond), Japanese (ai) (dark green inverted triangle), European (human) (purple plus), and European (ai) (purple cross). Dashed lines indicate clusters for each language. A shaded gray region highlights the overlap between human and AI data points for the European language.

