Legend: Direct, COT, Override

Charts (left to right): Conv End (F1), Lexical, Syntactic, Semantic, Style

Y-axis categories (top to bottom): English, Chinese, Russian

X-axis: Human/Simulator Similarity (0.0 to 0.2)