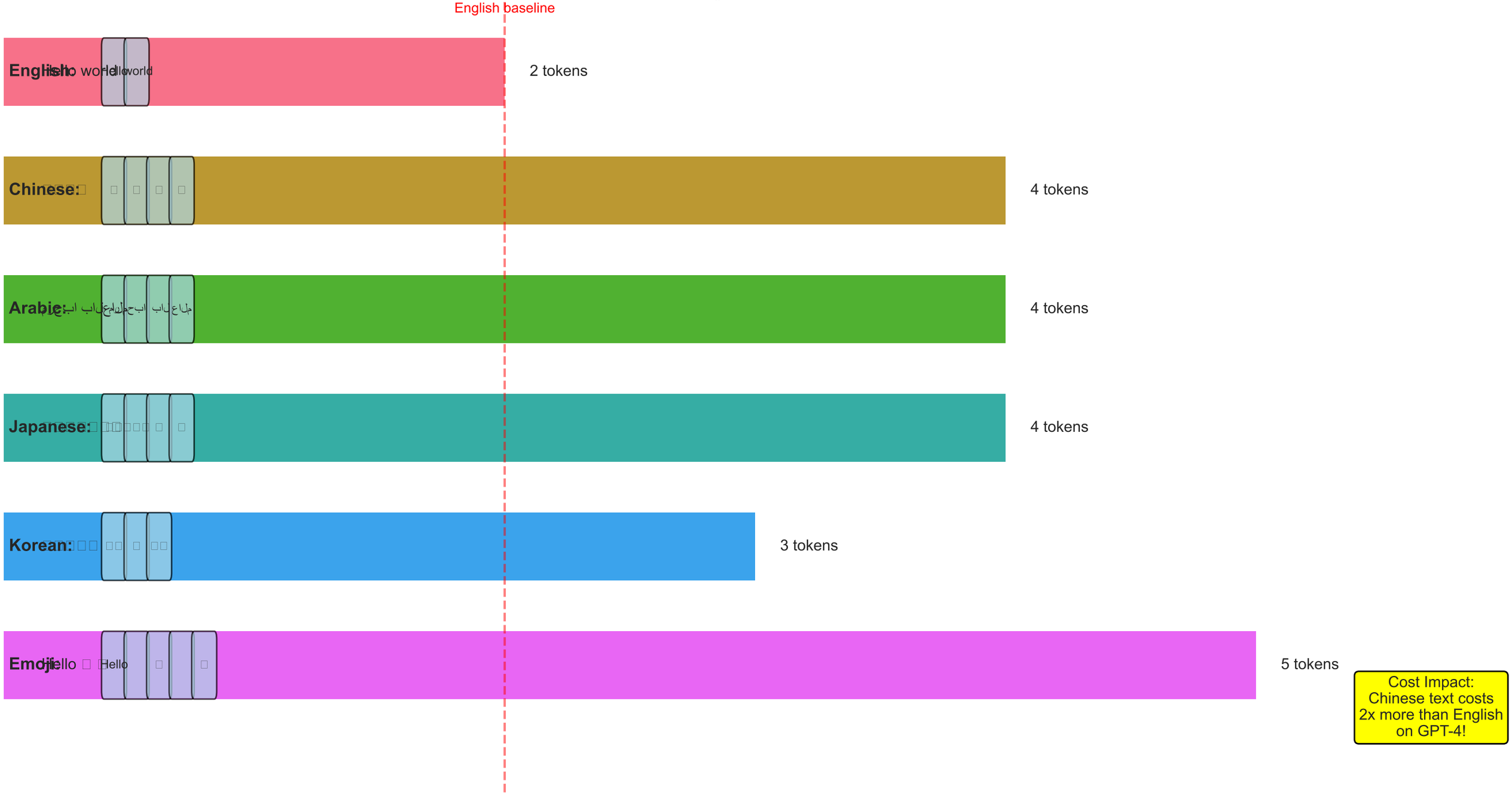
Tokenization Across Languages

Same meaning, different token counts