STATISTICAL ANALYSIS

Pearson's r correlation between text length and MT errors:

Text Length - DL MQM errors: r(50) = -0.128, p = 0.366

Text Length - GT MQM errors: r(50) = -0.116, p = 0.411

Text Length - SY MQM errors: r(50) = -0.081, p = 0.568

--> No statistically significant correlation between text length and MT errors (evaluated according to the MQM framework) for any of the three MT

Pearson's r correlation between readability and MT errors:

Readability - DL MQM errors: r(50) = 0.359, p = **0.009**

Readability - GT MQM errors: r(50) = 0.331, p = **0.016**

Readability - SY MQM errors: r(50) = 0.306, p = **0.027**

--> Statistically significant correlation between readability and MT errors (evaluated according to the MQM framework) for any of the three MT

DeepL (r(50) =0.359, p = 0.009)

Google Translator r(50) = 0.306, p = 0.027),

Systran (r(50) = 0.306, p =0.027)

T-Test for statistical significance of the difference (superiority) of DL over GT and SY

DL - GT MQM errors: Two sample t-test (left-tailed) p = 0.2647

DL - SY MQM errors: Two sample t-test (left-tailed) p = 0.2634

--> No statistically significant difference: the sample average of DL is smaller than the sample average of both GT and SY, but not small enough to be statistically sig engines

engines

nificant