**Statistical Analysis Code**

**All the statistical analyses in our study were conducted on Stata 17.0.**

**First, we used the following code to obtain such information as mean, median, standard deviation and 95% CI.**

* mean *variable*

For example:

mean ASR

Mean estimation Number of obs = 31

--------------------------------------------------------------

| Mean Std. err. [95% conf. interval]

-------------+------------------------------------------------

ASR | 2.967742 .1180839 2.726582 3.208901

--------------------------------------------------------------

* sum *variable*, detail

For example:

sum ASR, detail

A-SR

-------------------------------------------------------------

Percentiles Smallest

1% 2 2

5% 2 2

10% 2 2 Obs 31

25% 3 2 Sum of wgt. 31

50% 3 Mean 2.967742

Largest Std. dev. .6574634

75% 3 4

90% 4 4 Variance .4322581

95% 4 4 Skewness .0305208

99% 4 4 Kurtosis 2.387651

**Second, we conducted Friedman tests on SR, translation time, edit, fixation and pause. We used the following code.**

* emh *variable* Group, strata (Participants’ number) anova transformation(rank)

For example:

emh SR Group, strata (Participants’ number) anova transformation(rank)

Extended Mantel-Haenszel (Cochran-Mantel-Haenszel) Stratified Test of Association

ANOVA (Row Mean Scores) Statistic:

Q (3) = 47.3462, P = 0.0000

Transformation: Ranks

**Third, we did Dunn’s tests to make pairwise comparison and further explore the differences. We used the following code.**

* dunntest *variable*, by (Group)

For example:

dunntest SR, by(Group)

Kruskal–Wallis equality-of-populations rank test

+---------------------+

| GN | Obs | Rank sum |

|----+-----+----------|

| 1 | 31 | 1551.50 |

| 2 | 31 | 1214.50 |

| 3 | 31 | 2770.50 |

| 4 | 31 | 2213.50 |

+---------------------+

chi2(3) = 36.007

Prob = 0.0001

chi2(3) with ties = 40.535

Prob = 0.0001

Dunn's Pairwise Comparison of SR by GN

(No adjustment)

Col Mean-|

Row Mean | 1 2 3

---------+---------------------------------

2 | 1.263505

| 0.1032

|

3 | -4.570364 -5.833868

| 0.0000 0.0000

|

4 | -2.482019 -3.745524 2.088345

| 0.0065 0.0001 0.0184

**Fourth, we conducted Kruskal-Wallis equality-of-populations rank tests on translation entropy. We used the following code.**

* kwallis *variable*, by(Group)

For example:

kwallis Phrasalentropy, by(Group)

Kruskal–Wallis equality-of-populations rank test

+---------------------+

| GP | Obs | Rank sum |

|----+-----+----------|

| 1 | 11 | 155.00 |

| 2 | 10 | 108.00 |

| 3 | 6 | 189.00 |

| 4 | 8 | 178.00 |

+---------------------+

chi2(3) = 18.328

Prob = 0.0004

chi2(3) with ties = 18.336

Prob = 0.0004

**Finally, we adopted linear mixed models (LMMs) analyses to study the impact of translation entropy on effort. We built models on SR, TTw, TEw, FDw, FCw, PDw and PCw. The fixed effect is the translation entropy of four texts. The random effect is participants’ English proficiency. We used the following code.**

* mixed *variable* fixedfactor || randomfactor:

For example:

mixed SR PhraseEntropy || EP:

Performing EM optimization:

Performing gradient-based optimization:

Iteration 0: log likelihood = -148.86533

Iteration 1: log likelihood = -148.86531

Computing standard errors:

Mixed-effects ML regression Number of obs = 124

Group variable: Subjects Number of groups = 31

Obs per group:

min = 4

avg = 4.0

max = 4

Wald chi2(1) = 37.44

Log likelihood = -148.86531 Prob > chi2 = 0.0000

------------------------------------------------------------------------------

SR | Coef. Std. Err. z P>|z| [95% Conf. Interval]

-------------+----------------------------------------------------------------

PhraseEntropy | .5915492 .096676 6.12 0.000 .4020676 .7810307

\_cons | .9962055 .3890607 2.56 0.010 .2336604 1.758751

------------------------------------------------------------------------------

------------------------------------------------------------------------------

Random-effects Parameters | Estimate Std. Err. [95% Conf. Interval]

-----------------------------+------------------------------------------------

EP: |

var(\_cons) | .1733036 .0795903 .0704515 .4263093

-----------------------------+------------------------------------------------

var(Residual) | .5232263 .0767296 .3925214 .6974544

------------------------------------------------------------------------------

LR test vs. linear model: chibar2(01) = 9.32 Prob >= chibar2 = 0.0011

**We started with a simple model from the fixed effect and obtained the best model with the lowest BIC. We used the following code.**

**Model comparison analysis code:**

* estat ic

For example:

estat ic

Akaike's information criterion and Bayesian information criterion

-----------------------------------------------------------------------------

Model | N ll(null) ll(model) df AIC BIC

-------------+---------------------------------------------------------------

. | 124 . -143.7967 5 297.5934 311.6948

-----------------------------------------------------------------------------

Note: BIC uses N = number of observations. See [R] BIC note.

**We did model checking by examining the distribution of residuals. We used the following code.**

**Model checking analysis code:**

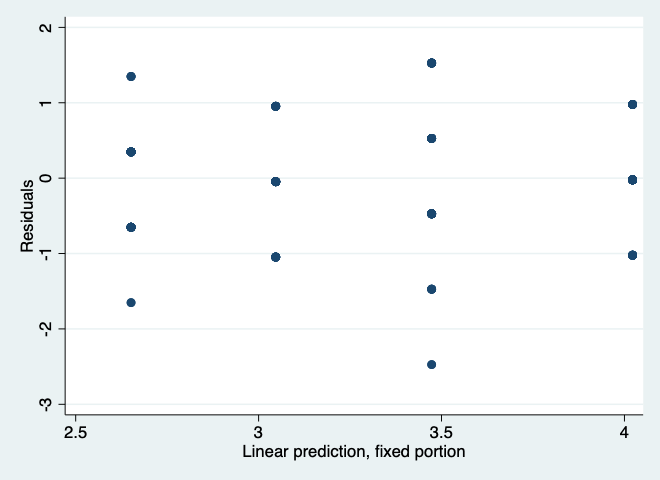
* predict fitted, xb
* predict residuals, res
* twoway scatter residuals fitted
* pnorm residuals

For example:

predict fitted, xb

predict residuals, res

twoway scatter residuals fitted



pnorm residuals

