**Appendix A**

This appendix contains the results for themodel comparisons and the mixed-effects model analysis on RT and reading accuracy in Experiments 1 and 2. The results of for continuous-variable model analyses, in which frequency and consistency were transformed as numerical variables, were followed to test the robustness of the optimal model.

**Table A1**

*Results of Model Comparison for RT in Experiment 1*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | npar | AIC | BIC | Deviance | χ2 | df | *p* |
| RT1 | 9 | 42,701 | 42,755 | 42,683 |  |  |  |
| **RT2** | **11** | **42,688** | **42,754** | **42,666** | **17.17** | **2** | **< .001** |
| RT3 | 13 | 42,707 | 42,784 | 42,681 | 5.15 | 2 | 1.00 |
| RT4 | 14 | 42,703 | 42,787 | 42,675 | 5.15 | 1 | < .05 |
| RT5 | 15 | 42,710 | 42,799 | 42,680 | 5.15 | 1 | 1.00 |

*Note.* npar = numbers of parameters. The best-fit model is shown in bold. The formulas for each model are as follows:

*RT1*: RT ~ Frequency × Consistency + (1 | participant) + (1 | word)

***RT2*: RT ~ Frequency × Consistency + (1 + Frequency | participant) + (1 | word)**

*RT3:* RT ~ Frequency × Consistency + Similarity × Frequency + Similarity × Consistency + (1 | participant) + (1 | word)

*RT4*: RT ~ Frequency × Consistency + (1 + Consistency | participant) + (1 | word)

*RT5*: RT ~ Frequency × Consistency × Similarity + (1 | participant) + (1 | word)

Model *RT2* has the lowest AIC with a significant improvement over other models and provides a good balance between model complexity and fit.

**Table A2**

*Summary of the Optimal Mixed-Effects Model on RT in Experiment 1*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Fixed effects | *β* | SE | df | *t* | *p* (>|*t*|) |
| Intercept | 1294.67 | 48.21 | 40.39 | 26.85 | <.001 \*\*\* |
| Frequency | -178.37 | 24.42 | 105.88 | -7.30 | <.001 \*\*\* |
| Consistency1 (atypical vs GM) | 134.99 | 32.36 | 110.54 | 4.17 | <.001 \*\*\* |
| Consistency2 (consistent vs GM) | -131.82 | 31.55 | 103.48 | -4.18 | <.001 \*\*\* |
| Frequency × Consistency1 | -110.88 | 32.36 | 110.49 | -3.43 | <.001 \*\*\* |
| Frequency × Consistency2 | 86.07 | 31.55 | 102.46 | 2.73 | <.01 \*\* |
| Random effects | Variance | Standard Deviation |  |  |  |
| word (intercept) | 51,736 | 227.46 |  |  |  |
| participant (intercept) | 40,30 | 200.75 |  |  |  |
| (Frequency) | 9,877 | 99.39 |  |  |  |
| Residual | 180,020 | 424.29 |  |  |  |

*Note.* Formula: RT ~ Frequency × Consistency + (1+ Frequency | participant) + (1|word); Consistency and frequency were coded usingsum contrast*,* with comparisons against the grand mean (GM) under each variable; SE = standard error; \**p* < .05; \*\**p* < .01; \*\*\**p* < .001. The number of observations is 2,832.

**Table A3**

*Summary of the Continuous-variable Model on RT in Experiment 1.*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Fixed effects | *β* | SE | df | | *t* | *p* (>|*t*|) |
| Intercept | 1247.13 | 46.28 | 42.68 | | 26.95 | <.001 \*\*\* |
| Frequency | -184.28 | 22.12 | 117.20 | | -8.33 | <.001 \*\*\* |
| Consistency | -66.09 | 21.10 | 108.75 | | -3.13 | <.01 \*\* |
| Frequency × Consistency | 38.64 | 18.44 | 113.39 | | 2.10 | <.05 \* |
| Random effects | Variance | Standard Deviation | |  |  |  |
| word (intercept) | 45,682 | 213.73 |  | |  |  |
| participant (intercept) | 39,240 | 198.05 |  | |  |  |
| (Frequency) | 9,674 | 98.22 |  | |  |  |
| Residual | 180,174 | 424.47 |  | |  |  |

*Note.* Formula: RT ~ Frequency × Consistency + (1+ Frequency | participant) + (1 | word); Consistency and frequency are numeric factors; SE = standard error; \**p* < .05; \*\**p* < .01; \*\*\**p* < .001.

**Table A4**

*Results of Model Comparison for Reading Accuracy in Experiment 1*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | npar | AIC | BIC | Deviance | χ2 | df | *p* |
| **Baseline** | **6** | **1,969.8** | **2006.5** | **1957.8** | 1.94 | 2 | 0.38 |
| RA\_AD1 | 8 | 1,971.9 | 2020.8 | 1955.9 |
| RA\_AD2 | 12 | 1,976.0 | 2049.3 | 1952.0 | 3.87 | 4 | 0.42 |
| RA\_AD3 | 14 | 1,978.0 | 2063.6 | 1950.0 | 1.93 | 2 | 0.38 |

*Note.* npar = numbers of parameters. The best-fit model is shown in bold. The formulas for each model are as follows:

*Baseline:* Correct\_Incorrect ~ Frequency + Consistency + (1 | participant) + (1 | word)

*RA\_AD1*: Correct\_Incorrect ~ Frequency × Consistency + (1 | participant) + (1 | word)

*RA\_AD2:* Correct\_Incorrect ~ Frequency × Consistency + Similarity × Frequency + Similarity × Consistency + (1 | participant) + (1 | word)

*RA\_AD3:* Correct\_Incorrect ~ Frequency × Consistency × Similarity + (1 | participant) + (1 | word)

The baseline model was included into the comparisons since no significant interaction of consistency and frequency was detected among all converged interaction models. Model *Baseline* proved to have the lowest AIC/BIC thus selected as the optimal.

**Table A5**

*Summary of the Optimal Mixed-Effects Model on Reading Accuracy (Errors) in Experiment 1*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Fixed effects | *β* | SE | *z* | | *p* (>|*z*|) | |
| Intercept | -2.91 | 0.22 | -12.84 | | <.001 \*\*\* | |
| Frequency | -1.16 | 0.16 | -7.31 | | <.001 \*\*\* | |
| Consistency1 (atypical vs GM) | 1.34 | 0.21 | 6.33 | | <.001 \*\*\* | |
| Consistency2 (consistent vs GM) | -1.29 | 0.23 | -5.55 | | <.001 \*\*\* | |
| Random effects | Variance | Standard Deviation | |  |  |  |
| word (intercept) | 1.85 | 1.36 |  | |  |  |
| participant (intercept) | 0.59 | 0.77 |  | |  |  |

*Note.* Formula: Correct\_Incorrect ~ Frequency + Consistency + (1+ |participant) + (1|word); Consistency and frequency were coded using sum contrast, with comparisons against the log-odds of grand mean (GM) under each variable; SE = standard error; \**p* < .05; \*\**p* < .01; \*\*\**p* < .001. The number of observations is 3,332.

**Table A6**

*Summary of the Continuous-variable Model on Reading Accuracy in Experiment 1*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Fixed effects | *β* | SE | *z* | *p* (>|*z*|) | |
| Intercept | -2.93 | 0.23 | -12.95 | <.001 \*\*\* | |
| Frequency | -1.32 | 0.17 | -8.20 | <.001 \*\*\* | |
| Consistency | -0.83 | 0.15 | -5.57 | <.001 \*\*\* | |
| Random effects | Variance | Standard Deviation |  |  |  |
| word (intercept) | 1.75 | 1.32 |  |  |  |
| participant (intercept) | 0.60 | 0.77 |  |  |  |

*Note.* Formula: Correct\_Incorrect ~ Frequency + Consistency + (1 | participant) + (1 | word); Consistency and frequency are numeric factors; SE = standard error; \**p* < .05; \*\**p* < .01; \*\*\**p* < .001.

**Table A7**

*Results of Model Comparison for Reading Accuracy in Experiment 2*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | npar | AIC | BIC | Deviance | χ2 | df | *p* |
| RA\_IM1 | 8 | 2,961.4 | 3011.2 | 2945.4 | 3.14 | 4 | 0.54 |
| RA\_IM2 | 12 | 2,966.3 | 3040.9 | 2942.3 |
| **RA\_IM3** | **13** | **2,956.7** | **3037.5** | **2930.7** | **11.60** | **1** | **<.001** |
| RA\_IM4 | 14 | 2,966.2 | 3053.2 | 2938.2 | 1.93 | 1 | 1.00 |

*Note.* npar = numbers of parameters. The best-fit model is shown in bold. The formulas for each model are as follows:

*RA\_IM1*: Correct\_Incorrect ~ Frequency × Consistency + (1 | participant) + (1 | word)

*RA\_IM2:* Correct\_Incorrect ~ Frequency × Consistency + Similarity × Frequency + Similarity × Consistency + (1 | participant) + (1 | word)

*RA\_IM3****:* Correct\_Incorrect ~ Frequency × Consistency + (1 + Consistency | participant) + (1 | word)**

*RA\_IM4:* Correct\_Incorrect ~ Frequency × Consistency × Similarity + (1 | participant) + (1 | word)

Model *RA\_IM3* has the lowest AIC with a significant improvement over other model.

**Table A8**

*Summary of the Optimal Mixed-Effects Model on Reading Accuracy (Errors) in Experiment 2*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Fixed effects | *β* | SE | *z* | | | *p* (>|*z*|) | |
| Intercept | -1.71 | 0.24 | -7.14 | | | <.001 \*\*\* | |
| Frequency | -0.98 | 0.14 | -6.83 | | | <.001 \*\*\* | |
| Consistency1 (atypical vs GM) | 1.29 | 0.21 | 6.08 | | | <.001 \*\*\* | |
| Consistency2 (consistent vs GM) | -1.48 | 0.23 | -6.47 | | | <.001 \*\*\* | |
| Frequency × Consistency1 | -0.74 | 0.20 | -3.65 | | | <.001 \*\*\* | |
| Frequency × Consistency2 | 0.45 | 0.21 | 2.17 | | | <.05 \* | |
| Random effects | Variance | Standard Deviation | |  |  | |  |
| word (intercept) | 2.02 | 1.42 |  | |  | |  |
| participant (intercept) | 1.09 | 1.04 |  | |  | |  |
| (Consistency1) | 0.12 | 0.35 |  | |  | |  |
| (Consistency2) | 0.17 | 0.41 |  | |  | |  |

*Note.* Formula Correct\_Incorrect ~ Frequency × Consistency + (1 + Consistency | participant) + (1 | word); Consistency and frequency were coded using sum contrast, with comparisons against the log-odds of grand mean (GM) under each variable; SE = standard error; \**p* < .05; \*\**p* < .01; \*\*\**p* < .001. The number of observations is 3,706.

**Table A9**

*Summary of the Continuous-variable Model on Reading Accuracy in Experiment 2*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Fixed effects | *β* | SE | | *z* | | *p* (>|*z*|) | | |
| Intercept | -1.66 | 0.24 | | -7.03 | | <.001 \*\*\* | | |
| Frequency | -1.24 | 0.15 | | -8.30 | | <.001 \*\*\* | | |
| Consistency | -0.77 | 0.15 | | -5.07 | | <.001 \*\*\* | | |
| Frequency × Consistency | 0.49 | 0.14 | | 3.45 | | <.001 \*\*\* | | |
| Random effects | Variance | Standard deviation | | |  | |  |  | |
| word (intercept) | 1.87 | 1.37 |  | | | |  |  | |
| participant (intercept) | 1.05 | 1.03 |  | | | |  |  | |
| (Consistency1) | 0.12 | 0.35 |  | | | |  |  | |
| (Consistency2) | 0.13 | 0.37 |  | | | |  |  | |

*Note.* Formula: Correct\_Incorrect ~ Frequency × Consistency + (1 + Consistency | participant) + (1 | word); Consistency and frequency are numeric factors; SE = standard error; \**p* < .05; \*\**p* < .01; \*\*\**p* < .001.