**Appendix C**

This appendix reports the results of the mixed-effects model for reading accuracy, in which Japanese L2 proficiency (advanced/intermediate), lexical frequency, reading consistency, and their interaction were fixed factors. The results contain model comparisons, summary of the optimal model, continuous-variable models and pair-wise comparisons.

**Table C1**

*Model Comparison for the Mixed-Effects Model Analysis on Reading Accuracy*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | npar | AIC | BIC | Deviance | χ2 | df | *p* |
| RA1 | 9 | 4,841.7 | 4,903.4 | 4,823.7 | 1.79 | 2 | 0.40 |
| RA2 | 11 | 4,843.9 | 4,919.4 | 4,821.9 |
| RA3 | 12 | 4844.2 | 4,926.5 | 4,820.2 | 1.73 | 1 | 0.18 |
| RA4 | 14 | 4836.9 | 4,932.9 | 4,808.9 | 11.30 | 2 | < .05 |
| **RA5** | **19** | **4835.5** | **4955.8** | **4,787.5** | **21.41** | **5** | **< .001** |
| RA6 | 21 | 4838.4 | 4982.5 | 4,796.4 | 0.00 | 2 | 1.00 |

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

*RA1*: Correct/Incorrect ~ Frequency + Consistency + Proficiency + Frequency × Consistency + (1 | participant) + (1 | word)

*RA2*: Correct/Incorrect ~ Frequency + Consistency + Proficiency + Frequency × Consistency + (1 + Consistency | participant) + (1 | word)

*RA3:* Correct/Incorrect~ Frequency × Consistency + Proficiency × Frequency + Proficiency × Consistency + (1+ | participant) + (1 | word)

*RA4:* Correct/Incorrect~ Frequency × Consistency × Proficiency + (1 | participant) + (1 | word).

***RA5:* Correct/Incorrect~ Frequency × Consistency × Proficiency + (1+ Consistency | participant) + (1 | word).**

*RA6:* Correct/Incorrect~ Frequency × Consistency + Proficiency × Frequency + L2 Proficiency × Consistency + (1+ Frequency + Consistency | participant) + (1 | word)

Model *RA5* has the lowest AIC and deviance value with a significant improvement, indicating the best overall fit among the tested models.

**Table C2**

*Summary of the Mixed-Effects Model on Reading Accuracy (Errors)*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Fixed Effects | *β* | SE | *z* | *p* (>|*t*|) |
| **Intercept** | **-2.98** | **0.24** | **-12.54** | **< .001 \*\*\*** |
| **Frequency** | **-1.15** | **0.48** | **-7.50** | **< .001 \*\*\*** |
| **Consistency1 (atyp vs GM)** | **1.26** | **0.22** | **5.65** | **< .001 \*\*\*** |
| **Consistency2 (con vs GM)** | **-1.36** | **0.24** | **-5.64** | **< .001 \*\*\*** |
| **Proficiency** | **1.34** | **0.26** | **5.08** | **< .001 \*\*\*** |
| Frequency × Consistency1 | -0.36 | 0.21 | -1.72 | 0.09 |
| Frequency × Consistency2 | 0.34 | 0.23 | 1.49 | 0.14 |
| Frequency × Proficiency | 0.17 | 0.09 | 1.74 | 0.08 |
| Consistency1 × Proficiency | 0.03 | 0.16 | 0.16 | 0.87 |
| Consistency2 × Proficiency | -0.04 | 0.17 | -0.23 | 0.82 |
| **Frequency × Consistency1 × Proficiency** | **-0.38** | **0.12** | **-3.15** | **< .01 \*\*** |
| Frequency × Consistency2 × Proficiency | 0.10 | 0.15 | 0.65 | 0.51 |
|  Random Effects | Variance | Standard Deviation |  |  |
|  word (intercept) | 1.89 | 1.38 |  |  |
| participant (intercept) | 0.87 | 0.93 |  |  |
|  (Consistency1) | 0.14 | 0.38 |  |  |
|  (Consistency2) | 0.08 | 0.28 |  |  |

*Note.* Formula: Correct/Incorrect ~ Frequency × Consistency × Proficiency + (1 | participant) + (1 | word); Consistency and frequency was coded using contr.sum, with comparisons against the log-odds of grand mean (GM) under each variable; Proficiency = Japanese L2 proficiency; SE = standard error; \**p* < .05; \*\**p* < .01; \*\*\**p* < .001. *p*-values below .05 and statistically significant fixed effects are shown in bold. The number of observations is 7038.

**Table C3**

*Summary of the Continuous-variable Model for Reading Accuracy*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Fixed Effects | *β* | SE | *z* | *p* (>|*t*|) |
| **Intercept** | **-2.96** | **9.23** | **-13.00** |  **< .001\*\*\*** |
| **Frequency** | **-1.36** | **0.15** | **-8.86** | **< .001\*\*\*** |
| **Consistency** | **-0.80** | **0.16** | **-5.06** | **< .001\*\*\*** |
| **Proficiency** | **1.30** | **0.25** | **5.28** | **< .001\*\*\*** |
| Frequency × Consistency | 0.17 | 0.15 | 1.15 | 0.25 |
| Frequency × Proficiency | 0.17 | 0.10 | 1.68 | 0.09 |
| Consistency × Proficiency | 0.03 | 0.11 | 0.25 | 0.80 |
| **Frequency × Consistency× Proficiency** | **0.31** | **0.10** | **3.02** | **< .01\*\*** |
| Random Effects | Variance | Standard Deviation |  |  |
| word (intercept) | 1.75 | 1.32 |  |  |
| participant (intercept) | 0.84 | 0.92 |  |  |
| (Consistency1) | 0.14 | 0.38 |  |  |
| (Consistency1) | 0.05 | 0.23 |  |  |

*Note.* Formula: Correct/Incorrect~ Frequency × Consistency × Proficiency + (1|participant) + (1|word); Consistency and frequency are numeric factors; Proficiency = Japanese L2 proficiency; SE = standard error; \**p* < .05; \*\**p* < .01; \*\*\**p* < .001. *p*-values below .05 and statistically significant fixed effects are shown in bold.

**Table C4**

*Results of the Pairwise Comparisons on Reading Accuracy (Errors)*

|  |
| --- |
| Frequency Effect ( HF- LF) within Consistency Levels |
|  | Consistency | Estimate | SE | *z* | *p* | 95% CI |
| Advanced | **con** | **-1.62** | **0.58** | **-2.77** | **< .01** | **[-2.76, -0.47]** |
| **typ** | **-2.26** | **0.51** | **-4.42** | **< .001** | **[-3.26, -1.26]** |
| **atyp** | **-3.01** | **0.49** | **-6.17** | **< .001** | **[-3.96, -2.05]** |
| Intermediate | **con** | **-1.09** | **0.50** | **-2.20** | **< .05** | **[-2.06, -0.12]** |
| **typ** | **-1.37** | **0.47** | **-2.92** | **< .01** | **[-2.29, -0.45]** |
| **atyp** | **-3.43** | **0.48** | **-7.21** | **< .001** | **[-4.36, -2.50]** |
| Consistency Effect within Frequency Levels |
| Proficiency = Advanced |
| HF |  **atyp-con** | **1.92** | **0.60** | **3.20** | **< .05** | **[0.51, 3.33]** |
| atyp-typ | 0.78 | 0.54 | 1.44 | 0.32 | [-0.49, 2.05] |
| con-typ | -1.14 | 0.61 | -1.87 | 0.15 | [-2.57, 0.29] |
| LF |  **atyp-con** | **3.31** | **0.51** | **6.49** | **< .001** | **[2.12, 4.51]** |
|  **atyp-typ** | **1.52** | **0.48** | **3.19** | **< .01** | **[0.40, 2.65]** |
|  **con-typ** | **-1.79** | **0.50** | **-3.55** | **< .01** | **[-2.97, -0.61]** |
| Proficiency = Intermediate |
| HF |  **atyp-con** | **1.51** | **0.52** | **2.93** | **< .01** | **[0.30, 2.72]** |
| atyp-typ | 0.14 | 0.49 | 0.28 | 0.96 | [-1.02, 1.29] |
|  | **con-typ** | **-1.38** | **0.50** | **-2.71** | **< .05** | **[-2.97, -0.61]** |
| LF | **atyp-con** | **3.86** | **0.49** | **7.85** | **< .001** | **[2.71, 5.01]** |
| **atyp-typ** | **2.19** | **0.47** | **4.63** | **< .001** | **[1.08, 3.31]** |
| **con-typ** | **-1.66** | **0.48** | **-3.47** | **< .01** | **[-2.79, -0.54]** |
| Difference between levels of L2 Proficiency (AD-IM) |  | -1.14 | 0.61 | -1.87 | 0.15 | [-2.57, 0.29] |
| HF | **con** | **-1.56** | **0.47** | **-3.33** | **< .001** | **[-2.48, -0.64]** |
| **typ** | **-1.80** | **0.33** | **-5.51** | **< .001** | **[-2.44, -1.16]** |
| **atyp** | **-1.15** | **0.32** | **-3.58** | **< .001** | **[-1.78, -0.52]** |
| LF | **con** | **-1.03** | **0.37** | **-2.76** | **< .01** | **[-1.77, -0.30]** |
| **typ** | **-0.91** | **0.26** | **-3.48** | **< .001** | **[-1.42, -0.40]** |
| **atyp** | **-1.58** | **0.30** | **-5.29** | **< .001** | **[-2.16, -0.99]** |

*Note.* Pair-wise comparisons followed the estimation of the mixed-effects model: Correct/Incorrect~ Frequency × Consistency × Proficiency + (1 + Consistency |participant) + (1|word); con = consistent, typ/typical = inconsistent-typical, atyp/atypical = inconsistent-atypical; AD/Advanced = advanced learner group, IM/Intermediate = intermediate learner group; HF and LF represent high- and low-frequency band, respectively; CI = Confidient Intervel [lower confidence limited, upper confidence limited]; *p*-values below .05 are shown in bold. The *p-value* adjustment used the “Turkey” correction.