**Supplementary materials**

**Table S1**

*Participant demographics (means, standard deviations and ranges)*

|  |  |  |  |
| --- | --- | --- | --- |
|  | L2ers  (*N*=23*, Male=*10*)* | Heritage learners (HLs)  (*N*=18*, Male=*4*)* |  |
| Age | 28.5 (SD= 14.2)  (18-76) | 24.2 (SD= 12.8)  (18-73) | *t*(39) =1.005  *p=*0.32 |
| Length of stay in a Mandarin-speaking country/environment (in months) | 23.1 (SD = 43.8)  (0-180) | 38.6 (SD = 68.6)  (0-276) | *t*(39) = -0.876  *p=*0.39 |
| LexTALE\_CH score  (Corrected Accuracy, range:  -60–60) | 13.7 (SD = 12.0)  (-12-36) | 17.9 (SD = 13.6)  (-4-37) | *t*(39) = -1.0503  *p=*0.3 |
| Self-rating of overall Mandarin language ability (1-10) | 5.3 (SD = 1.5)  (3-8) | 5.6 (SD = 1.3)  (3-8) | *t*(39) =-0.594  *p=*0.56 |

**Table S2**

*Numbers, means, and standard deviations for constructions produced by group pre- and post-priming*

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | TELL | | | | | | MAKE | | | | | | GIVE | | | | | |
| Baseline | | | Post-priming | | | Baseline | | | Post-priming | | | Baseline | | | Post-priming | | |
| N | mean | SD | N | mean | SD | N | mean | SD | N | mean | SD | N | mean | SD | N | mean | SD |
| L1 | DO | 73 | .97 | .16 | 69 | .92 | .27 | 0 | 0 | 0 | 0 | 0 | 0 | 26 | .35 | .48 | 30 | .4 | .49 |
| GO | 1 | .01 | .12 | 0 | 0 | 0 | 52 | .69 | .46 | 68 | .91 | .29 | 44 | .59 | .50 | 41 | .55 | .50 |
| Other | 1 | .01 | .12 | 6 | .08 | .27 | 23 | .31 | .46 | 7 | .09 | .29 | 5 | .07 | .25 | 4 | .05 | .23 |
| CL | DO | 114 | .93 | .26 | 110 | .89 | .31 | 11 | .09 | .29 | 2 | .02 | .13 | 46 | .37 | .49 | 34 | .28 | .45 |
| GO | 2 | .02 | .13 | 7 | .06 | .23 | 57 | .46 | .50 | 109 | .89 | .32 | 40 | .33 | .47 | 80 | .65 | .48 |
| Other | 7 | .06 | .23 | 6 | .05 | .22 | 55 | .45 | .50 | 12 | .10 | .30 | 37 | .30 | .46 | 9 | .07 | .26 |

**Figure S1**

*Proportion of DO, GO and other constructions produced by group pre-and post-priming*

**A graph of different colored bars

Description automatically generated with medium confidence***A graph of different colored bars

Description automatically generated with medium confidence*

*Panel A: with 6 primed verbs*

*Panel B: with 3 unprimed verbs*

**Table S3**

*Model output for productions with TELL verbs by the CLs to explore proficiency effects (n = 41)*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Predictors* | *Estimate* | *SE* | *z* | *p* |
| **LexTALE score** |  |  |  |  |
| (Intercept) | -5.11 | 1.43 | -3.58 | **<.001** |
| LexTALE score | -0.13 | 0.60 | -0.22 | 0.83 |
| Task phase | 1.38 | 0.98 | 1.41 | 0.16 |
| LexTALE score \* Task phase | -1.26 | 0.96 | -1.32 | 0.19 |
| **Overall self-rating** |  |  |  |  |
| (Intercept) | -5.10 | 1.45 | -3.52 | **<.001** |
| Overall self-rating | -0.27 | 0.60 | -0.46 | 0.65 |
| Task phase | 1.10 | 1.01 | 1.08 | 0.28 |
| Overall self-rating \* Task phase | -1.76 | 1.04 | -1.70 | 0.09 |

Formula: glmer(isGO ~ scale(LexTALE score/Overall self-rating) \* Task phase + (1| participant))

**Table S4**

*Model output for productions with MAKE verbs by the CLs to explore proficiency effects (n = 41)*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Predictors* | *Estimate* | *SE* | *z* | *p* |
| **LexTALE score** |  |  |  |  |
| (Intercept) | 3.97 | 0.98 | 4.04 | **<0.001** |
| LexTALE score | 0.49 | 0.56 | 0.87 | 0.38 |
| Task phase | 8.35 | 1.95 | 4.29 | **<0.001** |
| LexTALE score \* Task phase | -0.81 | 1.08 | -0.75 | 0.45 |
| **Overall self-rating** |  |  |  |  |
| (Intercept) | 3.92 | 1.05 | 3.73 | **<0.001** |
| Overall self-rating | -0.39 | 0.63 | -0.63 | 0.53 |
| Task phase | 8.24 | 2.07 | 3.99 | **<0.001** |
| Overall self-rating \* Task phase | -0.77 | 1.19 | -0.65 | 0.52 |

Formula for LexTALE score: glmer(isGO ~ scale(LexTALE score) \* Task phase + (1 + Task phase | participant) + (1 + scale(LexTALE score) | item)); Formula for Overall self-rating: glmer(isGO ~ scale(Overall self-rating) \* Task phase + (1 + Task phase | participant) + (1 | item)).

**Table S5**

*Model output for productions with GIVE verbs by the CLs to explore proficiency effects (n = 41)*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Predictors* | *Estimate* | *SE* | *z* | *p* |
| **LexTALE score** |  |  |  |  |
| (Intercept) | -0.07 | 0.19 | -0.37 | 0.72 |
| LexTALE score | 0.30 | 0.19 | 1.54 | 0.12 |
| Task phase | 1.58 | 0.31 | 5.13 | **<0.001** |
| LexTALE score \* Task phase | -0.29 | 0.30 | -0.99 | 0.32 |
| **Overall self-rating** |  |  |  |  |
| (Intercept) | -0.06 | 0.20 | -0.32 | 0.75 |
| Overall self-rating | -0.15 | 0.20 | -0.75 | 0.46 |
| Task phase | 1.58 | 0.31 | 5.13 | **<0.001** |
| Overall self-rating \* Task phase | 0.11 | 0.29 | 0.37 | 0.72 |

Formula: glmer(isGO ~ scale(LexTALE score/ Overall self-rating) \* Task phase + (1 + participant))

**Table S6**

*Model output for productions with TELL verbs by 2 learner groups (L2ers: n=23; HLs: n=18)*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Predictors* | *Estimate* | *SE* | *z* | *p* |
| (Intercept) | -8.52 | 2.11 | -4.03 | **<0.001** |
| Group | -0.09 | 2.03 | -0.05 | 0.96 |
| Task phase | 0.70 | 4.28 | 0.16 | 0.87 |
| Group \* Task phase | -0.66 | 4.17 | -0.16 | 0.86 |

Formula: glmer(isGO~ Group \* Task phase + (1 + Task phase | participant))

*Note.* The overall models with 3 groups (L1, L2, and HL) failed to converge. Thus, we removed the L1 data and ran models with the L2 and HL groups only and presented the results here.

**Table S7**

*Model output for productions with MAKE verbs by 3 groups (L1ers: n=25; L2ers: n=23; HLs: n=18)*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Predictors* | *Estimate* | *SE* | *z* | *p* |
| (Intercept) | 3.17 | 0.97 | 3.26 | **0.001** |
| Group (L2-HL) | 1.05 | 1.08 | 0.97 | 0.33 |
| Group (L2-L1) | 1.06 | 0.96 | 1.10 | 0.27 |
| Task phase | 7.81 | 1.89 | 4.13 | **<0.001** |
| Group (L2-HL) \* Task phase | -0.43 | 2.05 | -0.21 | 0.84 |
| Group (L2-L1) \* Task phase | -1.26 | 1.80 | -0.70 | 0.49 |

Formula: glmer(isGO ~ Group \* Task phase + (1+ Task phase | participant) + (1 | item))

*Note.* The variable Group is dummy coded with “L2” as the reference level.

**Table S8**

*Model output for productions with GIVE verbs by 3 groups (L1ers: n=25; L2ers: n=23; HLs: n=18).*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Predictors* | *Estimate* | *SE* | *z* | *p* |
| (Intercept) | -0.28 | 0.34 | -0.83 | 0.41 |
| Group (L2-HL) | 0.52 | 0.46 | 1.13 | 0.26 |
| Group (L2-L1) | 0.67 | 0.42 | 1.57 | 0.12 |
| Task phase | 2.46 | 0.53 | 4.61 | **<0.001** |
| Group (L2-HL) \* Task phase | -1.42 | 0.63 | -2.26 | **0.02** |
| Group (L2-L1) \* Task phase | -2.61 | 0.60 | -4.38 | **<0.001** |

Formula: glmer(isGO ~ Group \* Task phase + (1 | participant) + (1 | item))

*Note.* The variable Group is dummy coded with “L2” as the reference level. To explore the Group by Task phase interactions, we ran separate models for the 3 groups, see the model output in Table S9.

**Table S9**

*Model output for productions with GIVE verbs for individual groups (L1ers: n=25; L2ers: n=23; HLs: n=18)*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Predictors* | *Estimate* | *SE* | *z* | *p* |
| L1ers | | | | |
| (Intercept) | 0.53 | 0.53 | 1.00 | 0.32 |
| Task phase | -0.04 | 0.83 | -0.04 | 0.97 |
| HLs | | | | |
| (Intercept) | 0.19 | 0.31 | 0.61 | 0.54 |
| Task phase | 0.90 | 0.44 | 2.07 | **0.04** |
| L2ers | | | | |
| (Intercept) | 0.28 | 0.27 | -1.07 | 0.29 |
| Task phase | 2.17 | 0.45 | 4.78 | **<0.001** |

Formula for L1ers: glmer (isGO ~ Task phase + (1 | participant) + (1 | item)); Formula for HLs and L2ers: glmer (isGO ~ Task phase + (1 | participant))

**Figure S2**

*Ratings for 6 individual verb-dative pairings by group pre- and post-priming. Error bars indicate 95% confidence intervals on means by participants.*

*A graph of different colored bars

Description automatically generated*

**Table S10**

*Overall model output for acceptability ratings (L1ers: n = 25; CLs: n=41)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Predictors* | *Estimate* | *SE* | *df* | *t* | *p* |
| (Intercept) | 0.01 | 0.05 | 28.60 | 0.28 | 0.78 |
| Group | -0.06 | 0.04 | 60.54 | -1.46 | 0.15 |
| Task phase | 0.30 | 0.04 | 40.46 | 6.79 | **<0.001** |
| Acceptability | 1.14 | 0.05 | 138.19 | 21.07 | **<0.001** |
| Group \* Task phase | 0.04 | 0.09 | 69.72 | 0.45 | 0.65 |
| Group \* Acceptability | -0.51 | 0.09 | 57.65 | -5.93 | **<0.001** |
| Task phase \* Acceptability | 0.03 | 0.06 | 77.86 | 0.42 | 0.67 |
| Group \* Task phase \* Acceptability | 0.44 | 0.12 | 2050.56 | 3.62 | **<0.001** |

Formula: lmer (zscore ratings ~ Group \* Task phase \* Acceptability + (1 + Task phase + Acceptability | participant)+

(1 + Task phase | item))

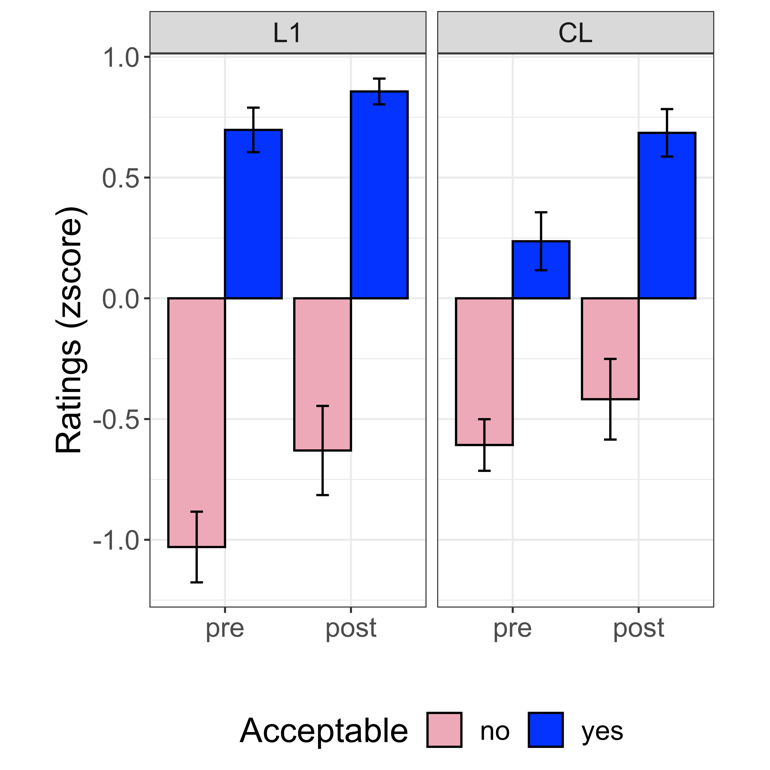
**Figure S3**

*Ratings for acceptable (blue) and unacceptable (pink) verb-dative pairings by group pre- and post-priming. Error bars indicate 95% confidence intervals on means by participants.*

*Panel B: with 3 unprimed verbs*

*Panel A: with 6 primed verbs*

*A graph of different levels of a bar graph

Description automatically generated with medium confidence*

**Table S11**

*Model output for acceptability ratings for the CLs to explore (Overall self-rating) proficiency effects (n=41)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Predictors* | *Estimate* | *SE* | *df* | *t* | *p* |
| (Intercept) | -0.05 | 0.04 | 30.76 | -1.10 | 0.28 |
| Overall self-rating | 0.00 | 0.03 | 38.68 | 0.09 | 0.93 |
| Task phase | 0.32 | 0.06 | 42.49 | 5.66 | **<0.001** |
| Acceptability | 0.90 | 0.08 | 26.55 | 11.38 | **<0.001** |
| Overall self-rating \* Task phase | -0.10 | 0.06 | 44.25 | -1.79 | 0.08 |
| Overall self-rating \* Acceptability | 0.09 | 0.05 | 36.55 | 1.64 | 0.11 |
| Task phase \* Acceptability | 0.25 | 0.08 | 1250.38 | 3.12 | **0.002** |
| Overall self-rating \* Task phase \* Acceptability | -0.03 | 0.08 | 1262.77 | -0.34 | 0.73 |

Formula: lmer (zscore ratings ~ scale(Overall self-rating) \* Task phase \* Acceptability + (1 + Task phase + Acceptability | participant)+(1 + Acceptability | item))

**Table S12**

*Model output for acceptability ratings for the CLs to explore (LexTALE score) proficiency effects (n=41)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Predictors* | *Estimate* | *SE* | *df* | *t* | *p* |
| (Intercept) | -0.05 | 0.04 | 29.91 | -1.10 | 0.28 |
| LexTALE score | 0.04 | 0.03 | 36.95 | 1.39 | 0.17 |
| Task phase | 0.32 | 0.06 | 42.76 | 5.49 | **<0.001** |
| Acceptability | 0.90 | 0.08 | 25.78 | 11.53 | **<0.001** |
| LexTALE score \* Task phase | -0.10 | 0.06 | 43.18 | -1.65 | 0.11 |
| LexTALE score \* Acceptability | 0.12 | 0.05 | 34.81 | 2.26 | **0.03** |
| Task phase \* Acceptability | 0.25 | 0.08 | 1251.10 | 3.13 | **0.002** |
| LexTALE score \* Task phase \* Acceptability | 0.14 | 0.08 | 1250.96 | 1.66 | 0.10 |

Formula: lmer (zscore ratings ~ scale(LexTALE score) \* Task phase \* Acceptability + (1 + Task phase + Acceptability | participant) + (1 + Acceptability | item))

*Note.* To explore the LexTALE score by Acceptability interaction, we split data by Acceptability and presented the model output in Table S13.

**Table 13**

*Model output for acceptability ratings by Acceptability for the CLs to explore (LexTALE score) proficiency effects (n=41)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Predictors* | *Estimate* | *SE* | *df* | *t* | *p* |
| Acceptable sentences | | | | | |
| (Intercept) | 0.40 | 0.07 | 26.48 | 5.99 | **<0.001** |
| LexTALE score | 0.10 | 0.04 | 29.69 | 2.34 | **0.03** |
| Task phase | 0.44 | 0.06 | 38.12 | 7.06 | **<0.001** |
| LexTALE score \* Task phase | -0.02 | 0.06 | 38.80 | -0.36 | 0.72 |
| Unacceptable sentences | | | | | |
| (Intercept) | -0.50 | 0.05 | 10.29 | -10.34 | **<0.001** |
| LexTALE score | -0.02 | 0.05 | 10.67 | 0.31 | 0.76 |
| Task phase | 0.18 | 0.07 | 414.51 | 2.57 | **0.01** |
| LexTALE score \* Task phase | -0.16 | 0.07 | 398.76 | -2.20 | **0.03** |

Formula: lmer (zscore ratings ~ scale(LexTALE score) \* Task phase + (1 + scale(LexTALE score) | item) + (1 + Task phase | participant))

*Note.* The model for Acceptable sentences indicated a significant main effect of the LexTALE score (*b=.10, p=.03*), showing that the CLs’ ratings for the acceptable sentences increased with the increase of their LexTALE scores, regardless of task phase. The model for Unacceptable sentences returned a LexTALE score by Task phase interaction. To explore this interaction, we split data by Task phase. The follow-up models showed that the CLs’ ratings for these unacceptable trials numerically increased with the increase of their LexTALE scores in the pretest (*b=.06, p=.24*), whereas their ratings numerically decreased with the increase of their LexTALE scores in the posttest (*b=-.10, p=.25*), causing the LexTALE score by Task phase interaction.

**Table S14**

*Overall model output for acceptability ratings by 3 groups (L1ers: n=25; L2ers: n=23; HLs: n=18)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Predictors* | *Estimate* | *SE* | *df* | *t* | *p* |
| (Intercept) | -0.08 | 0.05 | 44.42 | -1.42 | 0.16 |
| Group (L2-HL) | 0.14 | 0.05 | 61.73 | 2.69 | **0.009** |
| Group (L2-L1) | 0.12 | 0.05 | 60.72 | 2.63 | **0.01** |
| Task phase | 0.31 | 0.07 | 67.82 | 4.20 | **<0.001** |
| Acceptability | 0.85 | 0.08 | 87.74 | 10.75 | **<0.001** |
| Group (L2-HL) \* Task phase | 0.03 | 0.11 | 72.35 | 0.31 | 0.76 |
| Group (L2-L1) \* Task phase | -0.02 | 0.10 | 69.98 | -0.25 | 0.81 |
| Group (L2-HL) \* Acceptability | 0.08 | 0.11 | 58.79 | 0.77 | 0.44 |
| Group (L2-L1) \* Acceptability | 0.54 | 0.10 | 57.90 | 5.52 | **<0.001** |
| Task phase \* Acceptability | 0.39 | 0.11 | 499.73 | 3.75 | **<0.001** |
| Group (L2-HL) \* Task phase \* Acceptability | -0.33 | 0.16 | 2076.86 | -2.09 | **0.04** |
| Group (L2-L1) \* Task phase \* Acceptability | -0.59 | 0.14 | 2061.74 | -4.19 | **<0.001** |

Formula: lmer (zscore ratings ~ Group \* Task phase \* Acceptability + (1 + Task phase + Acceptability | participant) +

(1 + Task phase | item))

*Note.* The variable Group is dummy coded with “L2” as the reference level. To explore the three-way interactions, we ran separate models for the 3 groups, see the model output in Table S15. We also split data by Acceptability and presented the model output in Table S16.

**Table S15**

*Model output for acceptability ratings for individual groups (L1ers: n=25; L2ers: n=23; HLs: n=18)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Predictors* | *Estimate* | *SE* | *df* | *t* | *p* |
| L1ers | | | | | |
| (Intercept) | 0.06 | 0.08 | 25.90 | 0.75 | 0.48 |
| Task phase | 0.28 | 0.06 | 25.30 | 4.75 | **<0.001** |
| Acceptability | 1.47 | 0.07 | 60.01 | 19.79 | **<0.001** |
| Task phase \* Acceptability | -0.18 | 0.08 | 82.93 | -2.27 | **0.03** |
| HLs | | | | | |
| (Intercept) | 0.04 | 0.06 | 23.40 | 0.65 | 0.52 |
| Task phase | 0.33 | 0.11 | 21.09 | 3.07 | **0.006** |
| Acceptability | 0.97 | 0.08 | 110.42 | 12.55 | **<0.001** |
| Task phase \* Acceptability | 0.06 | 0.14 | 74.38 | 0.42 | 0.67 |
| L2ers | | | | | |
| (Intercept) | -0.09 | 0.06 | 27.21 | -1.64 | 0.11 |
| Task phase | 0.31 | 0.07 | 26.20 | 4.51 | **<0.001** |
| Acceptability | 0.86 | 0.07 | 120.27 | 11.97 | **<0.001** |
| Task phase \* Acceptability | 0.35 | 0.12 | 721.04 | 3.00 | **0.003** |

Formula for L1ers: lmer (zscore ratings ~ + Task phase \* Acceptability + (1 + Task phase |item) + (1 + Task phase + Acceptability | participant); Formula for L2ers: lmer (zscore ratings ~ + Task phase \* Acceptability + (1 |item) + (1 + Task phase | participant); Formula for HLs: lmer (zscore ratings ~ + Task phase \* Acceptability + (1 + Task phase |item) + (1 + Task phase | participant)

*Note.* The L2ers only increased their ratings for acceptable sentences post- vs. pre-priming (*b=.51, p<.001*), while their ratings for unacceptable sentences did not change (*b=.12, p=.18*), hence the Task phase by Acceptability interaction. By contrast, the two-way interaction in the L1 group was caused by their greater rating increase for unacceptable sentences (*b=.38, p<.001*) than that for acceptable ones (*b=.18, p<.001*). On the other hand, the HLs increased their ratings in acceptable vs. unacceptable sentences to a similar degree.

**Table S16**

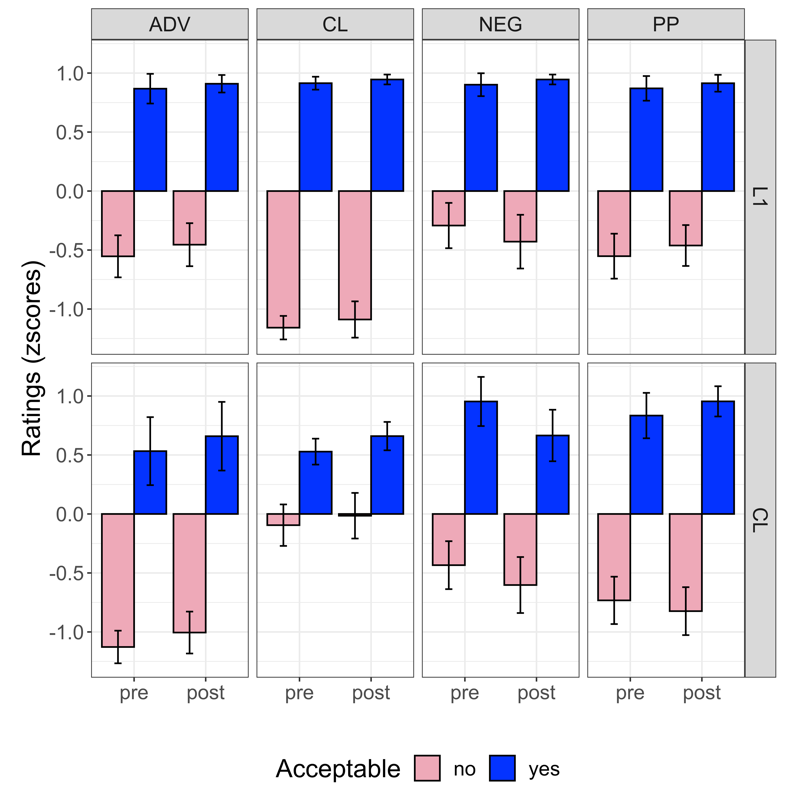
*Model output for acceptability ratings by Acceptability (L1ers: n=25; L2ers: n=23; HLs: n=18)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Predictors* | *Estimate* | *SE* | *df* | *t* | *p* |
| Acceptable sentences | | | | | |
| (Intercept) | 0.32 | 0.07 | 40.22 | 4.71 | **<0.001** |
| Group (L2-HL) | 0.18 | 0.07 | 61.38 | 2.66 | **0.01** |
| Group (L2-L1) | 0.40 | 0.06 | 60.45 | 6.44 | **<0.001** |
| Task phase | 0.52 | 0.08 | 60.12 | 6.43 | **<0.001** |
| Group (L2-HL) \* Task phase | -0.15 | 0.11 | 65.25 | -1.28 | 0.21 |
| Group (L2-L1) \* Task phase | -0.33 | 0.11 | 62.22 | -3.10 | **0.003** |
| Unacceptable sentences | | | | | |
| (Intercept) | -0.52 | 0.10 | 15.66 | -5.24 | **<0.001** |
| Group (L2-HL) | 0.09 | 0.08 | 58.32 | 1.15 | 0.26 |
| Group (L2-L1) | -0.15 | 0.07 | 57.49 | -2.06 | **0.04** |
| Task phase | 0.10 | 0.10 | 694.52 | 1.07 | 0.28 |
| Group (L2-HL) \* Task phase | 0.19 | 0.15 | 694.71 | 1.31 | 0.19 |
| Group (L2-L1) \* Task phase | 0.27 | 0.13 | 689.10 | 2.07 | **0.04** |

Formula for Acceptable sentences: lmer (zscore ratings ~ Group \* Task phase + (1 + Task phase | item) + (1 + Task phase | participant); Formula for Unacceptable sentences: lmer (zscore ratings ~ Group \* Task phase + (1 | item) + (1 | participant).

**Figure S4**

*Ratings for acceptable (blue) and unacceptable (pink) fillers by group pre- and post-priming. Error bars indicate 95% confidence intervals on means by participants.*

**

*Note.* ADV = fillers with correct or incorrect relative orders of adverbial phrases and verbs; CL = fillers with matched or mismatched classifier-noun pairs; NEG = fillers with correct or incorrect negation words; PP = fillers with correct or incorrect relative orders of prepositional phrases and verbs.