# Supplemental Analyses

# Analysis of Training Stimuli

Phonetic analyses for on the training stimuli included six metrics: Maximum Pitch (Peak F0), Pitch Excursion, Duration, Minimum Intensity (MinDB), Maximum Intensity (MaxDB), and Average Intensity (MeanDB). Table S1 presents the means and standard deviations for these metrics by segment (NP1 and NP2).

Table S1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Acoustic Properties of Training Stimuli (standard deviations in parentheses) | | | | |
|  | NP1 |  | NP2 |  |
| Peak F0 | SVO | OVS | SVO | OVS |
| Contrastive | 247.30 (28.74) | 329.52 (18.63) | 197.28 (29.18) | 178.30 (10.33) |
| Neutral | 236.85 (17.16) | 230.02 (13.30) | 197.71 (27.48) | 200.11 (17.34) |
|  |  |  |  |  |
| Pitch Excursion | SVO | OVS | SVO | OVS |
| Contrastive | 61.05 (29.28) | 119.30 (32.34) | -42.08 (32.24) | -16.32 (9.83) |
| Neutral | 31.72 (26.41) | 31.24 (18.60) | -22.63 (10.46) | -26.80 (20.36) |
|  |  |  |  |  |
| Duration | SVO | OVS | SVO | OVS |
| Contrastive | 608.15 (120.17) | 707.27 (136.08) | 662.48 (247.58) | 642.2 (167.77) |
| Neutral | 637.53 (129.3) | 663.28 (124.28) | 701.37 (252.01) | 687.44 (233.61) |
|  |  |  |  |  |
| MaxDB | SVO | OVS | SVO | OVS |
| Contrastive | 76.41 (1.78) | 77.19 (1.82) | 74.99 (0.7) | 74.17 (0.52) |
| Neutral | 74.57 (0.9) | 74.39 (0.74) | 73.84 (0.36) | 73.75 (0.44) |
|  |  |  |  |  |
| MinDB | SVO | OVS | SVO | OVS |
| Contrastive | 73.04 (0.24) | 73.06 (0.18) | 72.97 (0.06) | 72.97 (0.09) |
| Neutral | 73.06 (0.17) | 73 (0.1) | 72.94 (0.05) | 72.95 (0.05) |
|  |  |  |  |  |
| MaxDB | SVO | OVS | SVO | OVS |
| Contrastive | 76.41 (1.78) | 77.19 (1.82) | 74.99 (0.7) | 74.17 (0.52) |
| Neutral | 74.57 (0.9) | 74.39 (0.74) | 73.84 (0.36) | 73.75 (0.44) |

These stimuli were the same as in Henry, Jackson, and DiMidio (2017). Their analysis of Peak F0 and Pitch Excursion is provided below:

Following O’Brien, Jackson, and Gardner (2014), pitch excursion was determined by calculating the difference between the minimum and maximum pitch at the beginning or end of a rise or fall. On syllables without a distinct rise or fall, the beginning of the accented vowel and end of the stressed syllable were used for analysis. Means and standard deviations for peak F0 and pitch excursion on NP1 and NP2 are presented [in Table S1]. On NP1, OVS sentences in the focused prosody condition had higher peak F0 (*t*(48) = 11.58, *p* < .001) and larger pitch excursion values (*t*(48) = 5.56, *p* < .001) than the SVO sentences. On NP2, SVO sentences in the focused prosody condition had higher peak F0 values (*t*(48) = 3.44, *p* < .001) and larger pitch excursion values (*t*(48) = 4.40, *p* < .001) than the OVS sentences. In the monotone prosody condition, SVO and OVS sentences were nearly identical in peak F0 (NP1: *t*(48) = 1.45, *p* = .155; NP2: *t* < 1) and pitch excursion (NP1: *t <* 1; NP2: *t* < 1). (p. 6)

Duration was calculated by marking the beginning and end of each constituent. As expected, analyses on NP1 confirmed that OVS sentences in the focused prosody condition were longer than the SVO sentences (*t*(48) = -2.26, *p* = .029). There were no differences between OVS and SVO sentences in the monotone prosody condition (*t*(48) = -0.62, *p* = .538). There were no differences between SVO and OVS sentences in the focused (*t*(48) = 0.323, *p* = .747) or monotone prosody (*t*(48) = 0.18, *p* = .860) conditions.

Maximum Intensity was the highest value of intensity (measured in DB) for a given constituent. Minimum Intensity was the lowest value of intensity (measured in DB) for a given constituent. For both maximum and minimum intensity, there were no differences between SVO and OVS sentences for the focused or monotone prosody conditions on either NP1 or NP2 (all p > .05).

Mean Intensity was the averaged intensity (measured in DB) throughout the constituent. On NP1, there was no difference between SVO and OVS sentences for the focused (*t*(48) = ‑0.89, *p* = .378) or monotone prosody (*t*(48) = 1.09, *p* = .281) conditions. On NP2 there was a difference between SVO and OVS sentences for the focused prosody condition (*t*(48) = 2.08, *p* = .043), but not for the monotone prosody condition (*t*(48) = 0.26, *p* = .800).

# Analysis of SVO accuracy on sentences and nominative case markers

## Sentence Comprehension

The descriptive statistics for the interpretation task are presented in Table S1. The final GLMM model yielded no effect for Group (χ2(1) = 0.64, *p* < .422), no effect of Time (χ2(1) = 0.71, *p* < .400) and no Time x Group interaction (χ2(1) = 0.71, *p* = .400).

Table S2

Descriptive statistics for sentence interpretation task (maximum score of six)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | SVO Sentences | | | |  | OVS Sentences | | | |
|  | *M(SD)* | *95% CI* | *Mdn* | *IQR* |  | *M(SD)* | *95% CI* | *Mdn* | *IQR* |
| PI |  |  |  |  |  |  |  |  |  |
| Pretest | 3.54 (0.8) | 3.21, 3.87 | 4 | 1 |  | 0.67 (0.62) | 0.42, 0.93 | 1 | 1 |
| Posttest | 3.68 (0.69) | 3.4, 3.96 | 4 | 0.5 |  | 3.4 (0.91) | 3.02, 3.78 | 4 | 1 |
|  |  |  |  |  |  |  |  |  |  |
| PI+P |  |  |  |  |  |  |  |  |  |
| Pretest | 3.39 (0.88) | 3.05, 3.73 | 4 | 1 |  | 0.89 (0.79) | 0.59, 1.2 | 1 | 1.75 |
| Posttest | 3.68 (0.72) | 3.4, 3.96 | 4 | 0 |  | 2.96 (1.2) | 2.5, 3.43 | 3 | 1 |

*Note:* IQR = Interquartile Range. Values for the 95% confidence interval are bootstrapped using one thousand samples

## Written Production

The group means and standard deviations for the production task are shown in S2. The final GLMM model yielded a main effect for Time (χ2(1) = 11.80, *p* < .001), but not for Group (χ2(1) = 2.57, *p* = .109) or for the Time x Group interaction (χ2(1) = 2.99, *p* = .084).

Table S3  
Descriptive Statistics for written production task (ratio of correct to obligatory occasions)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Nominative Forms | | | |  | Accusative Forms | | | |
|  | *M(SD)* | *95% CI* | *Mdn* | *IQR* |  | *M(SD)* | *95% CI* | *Mdn* | *IQR* |
| PI |  |  |  |  |  |  |  |  |  |
| Pretest | 0.95 (0.12) | 0.9, 1.00 | 1.00 | 0.00 |  | 0.46 (0.4) | 0.3, 0.63 | 0.5 | 0.88 |
| Posttest | 0.99 (0.05) | 0.97, 1.01 | 1.00 | 0.00 |  | 0.85 (0.32) | 0.72, 0.98 | 1.00 | 0.00 |
|  |  |  |  |  |  |  |  |  |  |
| PI+P |  |  |  |  |  |  |  |  |  |
| Pretest | 0.96 (0.19) | 0.88, 1.03 | 1.00 | 0.00 |  | 0.43 (0.46) | 0.25, 0.6 | 0.13 | 1.00 |
| Posttest | 1.00 (0.00) | 1.00, 1.00 | 1.00 | 0.00 |  | 0.85 (0.31) | 0.73, 0.97 | 1.00 | 0.13 |
|  |  |  |  |  |  |  |  |  |  |

*Note:* Values for the 95% confidence interval are bootstrapped using one thousand samples

## SPR Comprehension

The means and standard deviations for comprehension accuracy in the SPR task are given in Table S3. The final GLMM model yielded no effects for Time (χ2(1) = 0.10, *p* = .747), Group (χ2(1) = 0.03, *p* < .861), or the Time x Group interaction (χ2(1) = 0.77, *p* = .380).

Table S4

Descriptive Statistics for SPR comprehension questions (percentage of correct answers)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Time | Pretest | | | |  | Posttest | | | |
|  | *M(SD)* | *95% CI* | *Mdn* | *IQR* |  | *M(SD)* | *95% CI* | *Mdn* | *IQR* |
| PI |  |  |  |  |  |  |  |  |  |
| Total | 72.99 (4.73) | 70.83, 75.14 | 73.10 | 6.70 |  | 76 (5.11) | 73.67, 78.32 | 76.90 | 6.65 |
| SVO | 0.78 (0.21) | 0.68, 0.87 | 0.83 | 0.42 |  | 0.8 (0.15) | 0.74, 0.87 | 0.83 | 0.25 |
| OVS | 0.26 (0.20) | 0.17, 0.35 | 0.17 | 0.25 |  | 0.44 (0.24) | 0.33, 0.55 | 0.5 | 0.33 |
|  |  |  |  |  |  |  |  |  |  |
| PI+P |  |  |  |  |  |  |  |  |  |
| Total | 0.74 (0.05) | 0.72, 0.76 | 0.73 | 0.08 |  | 0.76 (0.05) | 0.74, 0.78 | 0.77 | 0.06 |
| SVO | 0.80 (0.16) | 0.73, 0.86 | 0.83 | 0.33 |  | 0.76 (0.19) | 0.69, 0.83 | 0.83 | 0.17 |
| OVS | 0.27 (0.23) | 0.18, 0.36 | 0.17 | 0.17 |  | 0.37 (0.24) | 0.28, 0.46 | 0.33 | 0.33 |

# Descriptive Statistics for SPR task – Reading Times

Table S5

Descriptive Statistics by group and condition for SPR task, Masculine-First items (Log RTs)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Segment | NP1 |  |  |  | Verb |  |  |  | NP2 |  |  |  | Prep |  |  |  | Final |  |  |  |
| PI | Mean | SD | Mdn | IQR | Mean | SD | Mdn | IQR | Mean | SD | Mdn | IQR | Mean | SD | Mdn | IQR | Mean | SD | Mdn | IQR |
| SVO, Pretest | 2.95 | 0.23 | 2.89 | 0.32 | 2.87 | 0.25 | 2.84 | 0.36 | 2.9 | 0.23 | 2.85 | 0.32 | 2.71 | 0.16 | 2.67 | 0.21 | 2.89 | 0.22 | 2.88 | 0.31 |
| OVS, Pretest | 2.98 | 0.26 | 2.92 | 0.4 | 2.88 | 0.25 | 2.8 | 0.35 | 2.93 | 0.23 | 2.91 | 0.26 | 2.69 | 0.17 | 2.65 | 0.16 | 2.87 | 0.23 | 2.82 | 0.37 |
| SVO, Posttest | 2.96 | 0.26 | 2.88 | 0.3 | 2.84 | 0.25 | 2.81 | 0.33 | 2.94 | 0.28 | 2.9 | 0.37 | 2.68 | 0.16 | 2.65 | 0.21 | 2.86 | 0.25 | 2.82 | 0.35 |
| OVS, Posttest | 3.03 | 0.25 | 3.01 | 0.34 | 2.87 | 0.26 | 2.84 | 0.36 | 3.01 | 0.29 | 3.02 | 0.53 | 2.68 | 0.14 | 2.67 | 0.18 | 2.87 | 0.25 | 2.84 | 0.36 |
| PI+P |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| SVO, Pretest | 2.97 | 0.19 | 2.96 | 0.2 | 2.91 | 0.23 | 2.88 | 0.34 | 2.93 | 0.2 | 2.91 | 0.25 | 2.7 | 0.15 | 2.68 | 0.18 | 2.88 | 0.19 | 2.86 | 0.21 |
| OVS, Pretest | 2.97 | 0.18 | 2.94 | 0.27 | 2.92 | 0.22 | 2.91 | 0.3 | 2.92 | 0.21 | 2.89 | 0.29 | 2.68 | 0.14 | 2.64 | 0.16 | 2.91 | 0.21 | 2.87 | 0.27 |
| SVO, Posttest | 2.94 | 0.19 | 2.93 | 0.19 | 2.87 | 0.21 | 2.85 | 0.3 | 2.92 | 0.22 | 2.86 | 0.29 | 2.69 | 0.16 | 2.66 | 0.15 | 2.84 | 0.2 | 2.81 | 0.28 |
| OVS, Posttest | 2.98 | 0.22 | 2.92 | 0.36 | 2.85 | 0.19 | 2.85 | 0.25 | 2.93 | 0.25 | 2.88 | 0.31 | 2.69 | 0.15 | 2.66 | 0.18 | 2.89 | 0.23 | 2.86 | 0.28 |

Table S6

Descriptive Statistics by group and condition for SPR task, Masculine-First items (Log RTs)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Segment | NP1 |  |  |  | Verb |  |  |  | NP2 |  |  |  | Prep |  |  |  | Final |  |  |  |
| PI | Mean | SD | Mdn | IQR | Mean | SD | Mdn | IQR | Mean | SD | Mdn | IQR | Mean | SD | Mdn | IQR | Mean | SD | Mdn | IQR |
| SVO, Pretest | 3.00 | 0.23 | 2.98 | 0.30 | 2.91 | 0.27 | 2.88 | 0.36 | 2.91 | 0.23 | 2.87 | 0.33 | 2.69 | 0.18 | 2.67 | 0.18 | 2.89 | 0.23 | 2.83 | 0.34 |
| OVS, Pretest | 3.01 | 0.24 | 2.94 | 0.35 | 2.88 | 0.29 | 2.81 | 0.48 | 2.90 | 0.23 | 2.82 | 0.30 | 2.68 | 0.17 | 2.63 | 0.22 | 2.90 | 0.24 | 2.90 | 0.30 |
| SVO, Posttest | 2.97 | 0.24 | 2.91 | 0.36 | 2.83 | 0.24 | 2.80 | 0.39 | 2.93 | 0.27 | 2.91 | 0.41 | 2.70 | 0.18 | 2.66 | 0.19 | 2.84 | 0.21 | 2.83 | 0.27 |
| OVS, Posttest | 2.97 | 0.23 | 2.96 | 0.35 | 2.83 | 0.23 | 2.79 | 0.36 | 2.88 | 0.27 | 2.81 | 0.30 | 2.67 | 0.17 | 2.64 | 0.19 | 2.87 | 0.26 | 2.87 | 0.32 |
| PI+P |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| SVO, Pretest | 3.03 | 0.20 | 3.00 | 0.28 | 2.90 | 0.24 | 2.86 | 0.37 | 2.94 | 0.22 | 2.91 | 0.25 | 2.69 | 0.12 | 2.66 | 0.12 | 2.91 | 0.20 | 2.88 | 0.23 |
| OVS, Pretest | 2.99 | 0.20 | 2.97 | 0.26 | 2.89 | 0.23 | 2.83 | 0.30 | 2.89 | 0.21 | 2.86 | 0.26 | 2.71 | 0.15 | 2.68 | 0.19 | 2.89 | 0.22 | 2.90 | 0.26 |
| SVO, Posttest | 2.96 | 0.18 | 2.95 | 0.23 | 2.86 | 0.20 | 2.84 | 0.31 | 2.93 | 0.22 | 2.92 | 0.28 | 2.67 | 0.16 | 2.64 | 0.20 | 2.87 | 0.22 | 2.85 | 0.29 |
| OVS, Posttest | 2.99 | 0.21 | 2.95 | 0.25 | 2.87 | 0.23 | 2.83 | 0.26 | 2.90 | 0.25 | 2.83 | 0.32 | 2.72 | 0.23 | 2.65 | 0.23 | 2.89 | 0.24 | 2.87 | 0.30 |

Table S7

Descriptive Statistics by group and condition for SPR task, Masculine-First items (Log RTs)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Segment | NP1 |  |  |  | Verb |  |  |  | NP2 |  |  |  | Prep |  |  |  | Final |  |  |  |
| PI | Mean | SD | Mdn | IQR | Mean | SD | Mdn | IQR | Mean | SD | Mdn | IQR | Mean | SD | Mdn | IQR | Mean | SD | Mdn | IQR |
| SVO, Pretest | 1062 | 535 | 851 | 536 | 898 | 436 | 694 | 617 | 957 | 371 | 929 | 670 | 554 | 181 | 523 | 118 | 918 | 377 | 523 | 419 |
| OVS, Pretest | 1172 | 632 | 1011 | 900 | 930 | 508 | 758 | 732 | 1022 | 473 | 913 | 611 | 540 | 210 | 476 | 150 | 888 | 385 | 476 | 557 |
| SVO, Posttest | 1126 | 682 | 922 | 815 | 825 | 402 | 705 | 581 | 1055 | 509 | 988 | 876 | 517 | 156 | 457 | 145 | 861 | 387 | 457 | 397 |
| OVS, Posttest | 1253 | 658 | 1029 | 461 | 872 | 399 | 867 | 631 | 1234 | 662 | 1014 | 1206 | 500 | 110 | 506 | 117 | 872 | 343 | 506 | 511 |
| PI+P |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| SVO, Pretest | 998 | 286 | 918 | 501 | 941 | 317 | 895 | 312 | 914 | 254 | 919 | 374 | 541 | 147 | 522 | 137 | 825 | 245 | 522 | 275 |
| OVS, Pretest | 1054 | 319 | 1008 | 647 | 984 | 313 | 944 | 523 | 962 | 318 | 897 | 463 | 522 | 117 | 517 | 183 | 935 | 316 | 517 | 390 |
| SVO, Posttest | 959 | 289 | 1021 | 537 | 840 | 277 | 859 | 413 | 958 | 403 | 821 | 406 | 528 | 129 | 498 | 156 | 790 | 304 | 498 | 326 |
| OVS, Posttest | 1081 | 437 | 955 | 489 | 781 | 209 | 761 | 357 | 1016 | 597 | 934 | 468 | 532 | 127 | 503 | 146 | 892 | 362 | 503 | 545 |

Table S8

Descriptive Statistics by group and condition for SPR task, Masculine-First items (Log RTs)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Segment | NP1 |  |  |  | Verb |  |  |  | NP2 |  |  |  | Prep |  |  |  | Final |  |  |  |
| PI | Mean | SD | Mdn | IQR | Mean | SD | Mdn | IQR | Mean | SD | Mdn | IQR | Mean | SD | Mdn | IQR | Mean | SD | Mdn | IQR |
| SVO, Pretest | 1153 | 484 | 1005 | 696 | 988 | 552 | 917 | 564 | 959 | 412 | 835 | 568 | 541 | 171 | 470 | 193 | 895 | 333 | 470 | 373 |
| OVS, Pretest | 1220 | 588 | 1054 | 839 | 977 | 592 | 953 | 673 | 922 | 502 | 872 | 516 | 522 | 166 | 430 | 287 | 932 | 404 | 430 | 498 |
| SVO, Posttest | 1117 | 584 | 933 | 836 | 813 | 403 | 710 | 416 | 1028 | 484 | 963 | 840 | 548 | 199 | 465 | 216 | 790 | 297 | 465 | 314 |
| OVS, Posttest | 1103 | 534 | 959 | 645 | 778 | 421 | 653 | 482 | 968 | 670 | 672 | 564 | 511 | 189 | 439 | 227 | 901 | 446 | 439 | 323 |
| PI+P |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| SVO, Pretest | 1194 | 398 | 1114 | 528 | 938 | 379 | 899 | 318 | 986 | 351 | 950 | 374 | 510 | 104 | 494 | 146 | 912 | 265 | 494 | 387 |
| OVS, Pretest | 1100 | 415 | 1008 | 524 | 968 | 469 | 813 | 474 | 881 | 297 | 818 | 231 | 555 | 131 | 533 | 192 | 910 | 387 | 533 | 354 |
| SVO, Posttest | 990 | 217 | 990 | 471 | 819 | 289 | 808 | 301 | 998 | 465 | 857 | 493 | 503 | 127 | 459 | 137 | 860 | 332 | 459 | 297 |
| OVS, Posttest | 1065 | 443 | 942 | 721 | 856 | 410 | 772 | 348 | 907 | 432 | 769 | 424 | 682 | 569 | 520 | 294 | 947 | 453 | 520 | 440 |

# Supplemental analysis of results excluding two participants from the PI+P group with high reported time abroad

## Sentence Comprehension

The descriptive statistics for the interpretation task are presented in Table S9 The maximal GLMM model yielded no effect for Group (χ2(1) = 0.05, *p* = .816), but did show a main effect of Time (χ2(1) = 165.15, *p* < .001) and a significant Time x Group interaction (χ2(1) = 4.41, *p* = .036). Follow up pairwise comparisons indicated that both groups made significant gains from pretest to posttest (both *p* < .001), and that there were no significant differences between the groups either at pretest (*p* = .221) or at posttest (*p* = .120).

Table S9

Descriptive statistics for sentence interpretation task (maximum score of six)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | SVO Sentences | | | |  | OVS Sentences | | | |
|  | *M(SD)* | *95% CI* | *Mdn* | *IQR* |  | *M(SD)* | *95% CI* | *Mdn* | *IQR* |
| PI |  |  |  |  |  |  |  |  |  |
| Pretest | 3.68 (0.48) | 3.50, 3.87 | 4 | 1 |  | 0.64 (0.64) | 0.39, 0.91 | 1 | 1 |
| Posttest | 3.68 (0.69) | 2.98, 3.74 | 4 | 1 |  | 3.4 (0.91) | 2.98, 3.74 | 4 | 1 |
|  |  |  |  |  |  |  |  |  |  |
| PI+P |  |  |  |  |  |  |  |  |  |
| Pretest | 3.46 (0.76) | 3.14, 3.75 | 4 | 1 |  | 0.92 (0.8) | 0.61, 1.23 | 1 | 1.75 |
| Posttest | 3.65 (0.75) | 3.30, 3.92 | 4 | 1 |  | 3.04 (1.08) | 2.59, 3.44 | 3 | 1 |

*Note:* IQR = Interquartile Range. Values for the 95% confidence interval are bootstrapped using one thousand samples

## Written Production

The group means and standard deviations for the production task are shown in Table S10. The maximal GLMM model yielded a main effect for Time (χ2(1) = 32.64, *p* < .001), but not for Group (χ2(1) = 0.06, *p* = .806) or for the Time x Group interaction (χ2(1) = 0.15, *p* = .698). Pairwise comparisons indicating that both groups improved from pretest to posttest (*p* < .001).

Table S10  
Descriptive Statistics for written production task (ratio of correct to obligatory occasions)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Nominative Forms | | | |  | Accusative Forms | | | |
|  | *M(SD)* | *95% CI* | *Mdn* | *IQR* |  | *M(SD)* | *95% CI* | *Mdn* | *IQR* |
| PI |  |  |  |  |  |  |  |  |  |
| Pretest | 0.95 (0.12) | 0.90,0.99 | 1.00 | 0.00 |  | 0.46 (0.4) | 0.30, 0.62 | 0.5 | 0.75 |
| Posttest | 0.99 (0.05) | 0.99, 1.05 | 1.00 | 0.00 |  | 0.85 (0.32) | 0.69, 0.97 | 1.00 | 0.00 |
|  |  |  |  |  |  |  |  |  |  |
| PI+P |  |  |  |  |  |  |  |  |  |
| Pretest | 0.95 (0.20) | 0.87, 1.01 | 1.00 | 0.00 |  | 0.42 (0.46) | 0.24, 0.60 | 0.13 | 0.96 |
| Posttest | 1.00 (0.00) | 1.00, 1.00 | 1.00 | 0.00 |  | 0.84 (0.32) | 0.70, 0.96 | 1.00 | 0.00 |
|  |  |  |  |  |  |  |  |  |  |

*Note:* Values for the 95% confidence interval are bootstrapped using one thousand samples

## SPR Comprehension

The means and standard deviations for comprehension accuracy in the SPR task are given in Table S11. The maximal GLMM model yielded a main effect for Time (χ2(1) = 7.87, *p* = .005). There was no effect for Group (χ2(1) = 0.16, *p* = .690) or the Time x Group interaction (χ2(1) = 0.70, *p* = .402). Pairwise comparisons for Time indicated that the participants improved from pretest to posttest (*p*= .004).

Although the PI group outscored the PI+P group on the posttest, both groups failed to reach 50% accuracy on the OVS items in the post-test. In order to assess whether individual participants had abandoned a strict subject-first strategy, the number of participants reaching 50% accuracy on OVS sentences was evaluated for each group following AUTHOR (under review). In the PI group, this proportion rose from 6 (24%) on the pretest to 15 (60%) on the posttest. In the PI+P group, this figure rose from 6 (23.1%) to 12 participants (46.2%).

Table S11

Descriptive Statistics for SPR comprehension questions (percentage of correct answers)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Time | Pretest | | | |  | Posttest | | | |
|  | *M(SD)* | *95% CI* | *Mdn* | *IQR* |  | *M(SD)* | *95% CI* | *Mdn* | *IQR* |
| PI |  |  |  |  |  |  |  |  |  |
| Total | 0.73 (0.05) | 0.71, 0.74 | 0.73 | 0.06 |  | 0.76 (0.05) | 0.74, 0.78 | 0.77 | 0.06 |
| SVO | 0.78 (0.21) | 0.68, 0.87 | 0.83 | 0.33 |  | 0.8 (0.15) | 0.74, 0.86 | 0.83 | 0.17 |
| OVS | 0.26 (0.2) | 0.18, 0.36 | 0.17 | 0.17 |  | 0.44 (0.24) | 0.33, 0.53 | 0.50 | 0.33 |
|  |  |  |  |  |  |  |  |  |  |
| PI+P |  |  |  |  |  |  |  |  |  |
| Total | 0.74 (0.05) | 0.72, 0.74 | 0.73 | 0.06 |  | 0.76 (0.05) | 0.74, 0.78 | 0.77 | 0.06 |
| SVO | 0.8 (0.17) | 0.73, 0.86 | 0.83 | 0.33 |  | 0.75 (0.19) | 0.68, 0.83 | 0.83 | 0.17 |
| OVS | 0.28 (0.23) | 0.19, 0.38 | 0.33 | 0.17 |  | 0.37 (0.24) | 0.28, 0.47 | 0.33 | 0.33 |

*Note:* Values for the 95% confidence interval are bootstrapped using one thousand samples; values for SVO and OVS sentences reflect only those items in which grammatical role assignment was tested in the comprehension question.

## SPR Reading Times

The analysis of RTs was conducted separately for Masculine-First and Masculine-Second sentences and for each region. For Masculine-First sentences, the critical and spillover segments were Segments 1 and 2. For Masculine-Second sentences, they were Segments 3 and 4.

### RTs for Masculine First Sentences

**.** The mean RTs and standard deviations for Masculine-First sentences at each segment are displayed by group and condition in Table S12.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Table S12 Mean reading times (SDs) by group and condition for SPR task, Masculine-First items | | | | | |
| Segment | NP1 | V | NP2 | Prep | Final |
| Processing Instruction |  |  |  |  |  |
| SVO, Pretest | 1062 (535) | 898 (436) | 957 (371) | 554 (181) | 918 (377) |
| OVS, Pretest | 1172 (632) | 930 (508) | 1022 (473) | 540 (210) | 888 (385) |
| SVO, Posttest | 1126 (682) | 825 (402) | 1055 (509) | 517 (156) | 861 (387) |
| OVS, Posttest | 1253 (658) | 872 (399) | 1234 (662) | 500 (110) | 872 (343) |
|  |  |  |  |  |  |
| Processing Instruction with Prosody | |  |  |  |  |
| SVO, Pretest | 986 (290) | 945 (330) | 906 (258) | 544 (153) | 805 (241) |
| OVS, Pretest | 1066 (328) | 993 (324) | 975 (327) | 523 (121) | 918 (301) |
| SVO, Posttest | 931 (272) | 858 (281) | 976 (413) | 532 (133) | 782 (312) |
| OVS, Posttest | 1095 (451) | 776 (213) | 1020 (613) | 536 (130) | 897 (375) |

For the critical segment, Segment 1 (the first NP), the maximal model did not converge. The final model[[1]](#footnote-1) (Table S13) yielded a significant main effect for Word Order. Pairwise comparisons for the SVO-OVS contrast showed that the OVS sentences had higher RTs than SVO sentences (*p* = .023). The estimated marginal means indicate that this effect was driven by larger differences between SVO and OVS sentences in the posttest than in the pretest (Table S14).

Table S13  
Model results for Segments 1 and 2, Masculine-First Sentences

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Segment 1 | | | Segment 2 | | |
| Effect | *Df* | *F* | *p* | *Df* | *F* | *p* |
| Time | 1, 42.15 | 0.04 | .834 | 1, 32.07 | 5.21 | .029\* |
| Word Order | 1, 19.98 | 5.89 | .025\* | 1, 690.97 | 0.02 | .884 |
| Group | 1, 45.27 | 0.37 | .547 | 1, 42.03 | 0.35 | .56 |
| Time x Word Order | 1, 43.24 | 1.72 | .196 | 1, 679.09 | 0.4 | .53 |
| Time x Group | 1, 42.37 | 0.84 | .364 | 1, 45.13 | 0.15 | .698 |
| Word Order x Group | 1, 44.09 | 0.12 | .727 | 1, 675.4 | 0.43 | .513 |
| Time x Word Order x Group | 1, 42.68 | 0.19 | .661 | 1, 674.22 | 1.74 | .188 |

Table S14

Estimated marginal means for Word Order by Time on Region 1 in Masculine First Sentences

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Word Order | Time | EM Mean | SE | *df* | Lower CL | Upper CL |
| SVO | Pretest | 2.96 | 0.03 | 86.4 | 2.91 | 3.01 |
| OVS | Pretest | 2.99 | 0.03 | 87.8 | 2.93 | 3.05 |
| SVO | Posttest | 2.95 | 0.03 | 87.4 | 2.89 | 3.00 |
| OVS | Posttest | 3.01 | 0.03 | 82.9 | 2.95 | 3.07 |

For the spillover segment, Segment 2 (the verb), the maximal model did not converge. The final model[[2]](#footnote-2) (Table S13) yielded only one significant effect, a main effect for Time, which indicated that participants read Segment 2 faster in the posttest than in the pretest (*p* = .020).

### RTs for Masculine Second Sentences

**.** The mean RTs and standard deviations for Masculine-Second sentences at each segment are displayed by group and condition in Table S15

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Table S15 Mean reading times (SDs) by group and condition for SPR task, Masculine-Second items | | | | | |
| Segment | NP1 | V | NP2 | Prep | Final |
| Processing Instruction |  |  |  |  |  |
| SVO, Pretest | 1153 (484) | 988 (552) | 959 (412) | 541 (171) | 895 (333) |
| OVS, Pretest | 1220 (588) | 977 (592) | 922 (502) | 522 (166) | 932 (404) |
| SVO, Posttest | 1117 (584) | 813 (403) | 1028 (484) | 548 (199) | 790 (297) |
| OVS, Posttest | 1103 (534) | 778 (421) | 968 (670) | 511 (189) | 901 (446) |
|  |  |  |  |  |  |
| Processing Instruction with Prosody | |  |  |  |  |
| SVO, Pretest | 1204 (412) | 940 (393) | 1000 (362) | 511 (108) | 892 (256) |
| OVS, Pretest | 1113 (427) | 958 (476) | 886 (308) | 561 (133) | 882 (365) |
| SVO, Posttest | 993 (220) | 816 (298) | 1011 (481) | 507 (131) | 835 (315) |
| OVS, Posttest | 1077 (458) | 876 (420) | 929 (441) | 679 (592) | 913 (409) |

For the critical segment, Segment 3 (the second NP), the maximal model did not converge. The final model[[3]](#footnote-3) (Table S16) yielded no significant effects.

For the spillover segment, Segment 4 (the preposition), the maximal model did not converge. The final model[[4]](#footnote-4) (Table S16), yielded a Word Order x Group interaction. Follow-up pairwise comparisons indicated that participants in the PI+P group had higher reading times in OVS sentences than SVO sentences, whereas the PI group did not (*p* = .005). Estimated marginal means (Table S17) suggest that this effect was driven by the greater differences between SVO and OVS sentences for the PI+P group on the posttest.

Table S15  
Model results for Segments 3 and 4, Masculine-Second Sentences

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Segment 3 | | | Segment 4 | | |
| Effect | *Df* | *F* | *p* | *Df* | *F* | *p* |
| Time | 1, 38.37 | 0.02 | .898 | 1, 45.39 | 0.10 | .758 |
| Word Order | 1, 17.49 | 2.53 | .130 | 1, 11.79 | 1.61 | .229 |
| Group | 1, 44.25 | 0.1 | .748 | 1, 42.29 | 0.28 | .6 |
| Time x Word Order | 1, 43.65 | 0.14 | .714 | 1, 669.00 | 0.08 | .783 |
| Time x Group | 1, 43.82 | 0.01 | .930 | 1, 45.38 | 0.01 | .94 |
| Word Order x Group | 1, 43.99 | 0 | .960 | 1, 664.63 | 8.34 | .004\*\* |
| Time x Word Order x Group | 1, 42.37 | 0.29 | .594 | 1, 669.16 | 0.81 | .37 |

Table S16

Estimated Marginal Means for Segment 4, Masculine-Second Sentences

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Word Order | Time | EM Mean | SE | *df* | Lower CL | Upper CL |
| Processing Instruction | | | | | | | |
|  | SVO | Pre | 2.70 | 0.02 | 71.2 | 2.65 | 2.75 |
|  | OVS | Pre | 2.69 | 0.03 | 86.4 | 2.63 | 2.74 |
|  | SVO | Post | 2.70 | 0.03 | 69.7 | 2.65 | 2.75 |
|  | OVS | Post | 2.68 | 0.03 | 79.4 | 2.62 | 2.73 |
|  |  |  |  |  |  |  |  |
| Processing Instruction with Prosody | | | | | | | |
|  | SVO | Pre | 2.69 | 0.02 | 72.0 | 2.64 | 2.73 |
|  | OVS | Pre | 2.72 | 0.03 | 77.5 | 2.67 | 2.77 |
|  | SVO | Post | 2.67 | 0.02 | 67.8 | 2.63 | 2.72 |
|  | OVS | Post | 2.73 | 0.03 | 80.8 | 2.68 | 2.78 |

1. RT.log ~ Time \* Word Order \* Group + (Time \* Word Order|| Subject) + (Word Order| Item) [↑](#footnote-ref-1)
2. RT.log ~ Time \* Word Order \* Group + (Time | Subject) + (Time|| Item) [↑](#footnote-ref-2)
3. RT.log ~ Time \* Word Order \* Group + (Time \* Word Order || Subject) + (Time + Word Order|| Item) [↑](#footnote-ref-3)
4. RT.log ~ Time \* Word Order \* Group + (Time | Subject) + (Word Order | Item) [↑](#footnote-ref-4)