Online supplemental material

Data analysis.

Linear mixed-effects modeling was applied to ACC and RT measures to address the clustered nature of the data, with responses (level-1) nested within subjects (level-2). Mixed-effects models allowed us to examine variability at each level as well as effects and interactions within and across levels. The nested structure of the data further implied that observations at level-1 were not independent, as they were related to specific level-2 units. The mixed-effects model captured this dependency by adding a random effect for each level-2 unit, which represents the deviation of each level-2 unit's average from the overall mean (Raudenbush & Bryk, 2002).

Model building followed a bottom up theory-guided approach, starting with level-1 predictors. A model comparison framework was used to contrast alternative models that were progressively more complex. This approach is preferable to significance tests of individual parameters in arriving at correct statistical inferences (e.g., Bliese & Ployhart, 2002). The Likelihood Ratio (LR) test and the Akaike Information Criterion (AIC) were used to compare the fit of competing models.

Positive skewness was often noted for RT and negative skewness for accuracy scores. Modeling with transformed data, however, did not change the patterns of effects and interactions, therefore untransformed scores were used for easier meaningful interpretation of the estimates.

Univariate outliers were trimmed following a two-step procedure: first, on the basis of visual inspection of the data, we established upper and lower criteria and replaced values beyond these criteria with either minimum or maximum values of the distribution. Second, values beyond 3 SD, as computed for each task and each participant, were replaced with the minimum or maximum value of the distribution. Overall, no more than 2.5% of the observations were replaced. Residuals were also inspected during the modeling process. Data were analyzed with Stata software (StataCorp, 2011), using the function xtmixed and Maximum Likelihood (ML) for parameter estimation.

References

Bliese, P. D., & Ployhart, R. E. (2002). Growth modeling using random coefficient models:Model building, testing, and illustrations. *Organizational Research Methods, 5*(4), 362-387.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.

StataCorp. (2011). Stata Statistical Software: Release 12. College Station, TX: StataCorp LP.

Table s-1. Descriptive statistics for Accuracy and Reaction Time data for Baseline and interference conditions by group.

Variable	SLI (n=22)	TLD-A (n=22)	TLD-L (n=22)
Baseline			
Accuracy	0.85(0.09)	0.89(0.05)	0.79(0.12)
Reaction Time (ms)	1441.8(249.2)	1181.1(270.4)	1624(212)
Interference			
Accuracy	0.76(0.11)	0.81(0.07)	0.76(0.12)
Reaction Time (ms)	1453.6(251.3)	1306.1(344.5)	1625.9(194)

Note. Mean(SD).

Table s-2. Descriptive statistics for Accuracy and Reaction Time data by repetition number and group.

Repetition	SLI (n=22)	TLD-A (n=22)	TLD-L (n=22)
Accuracy			
0	0.84(.13)	0.85(.13)	0.77(.23)
1	0.86(.14)	0.91(.09)	0.85(.15)
2	0.85(.14)	0.91(.09)	0.89(.13)
3	0.90(.12)	0.91(.12)	0.81(.19)
Reaction Time (ms)			
0	1262(290)	1142(254)	1234(252)
1	1210(266)	1078(309)	1329(273)
2	1195(237)	986(230)	1163(153)
3	1146(216)	1048(274)	1221(232)

Note. Mean(SD). Repetition 0 represents the item's first appearance, 1 represents its second appearance, 2 represents its third appearance, and 3 represents the item's fourth appearance (third repetition).

Table s-3. Descriptive statistics for Accuracy and Reaction Time data for repeated and post repetition items.

Variable	SLI (n=22)	TLD-A (n=22)	TLD-L (n=22)	Total (n=66)
Original scale				
rep ACC	0.864(0.135)	0.897(0.113)	0.829(0.179)	0.863(0.147)
rep RT	1203.7(253.08)	1063.9(269.83)	1236.9(235.9)	1168.2(263.3)
arep ACC	0.847(0.163)	0.928(0.11)	0.812(0.186)	0.862(0.163)
arep RT	1287.7(345.57)	1117.8(294.43)	1312.1(250.98)	1239.2(310.7)
Rescaled $(x - x[\text{at } r=0])$				
rep_r ACC	0.024(0.132)	0.05(0.118)	0.057(0.171)	0.044(0.142)
rep_r RT	-58.67(250.4)	-78.77(175.98)	2.97(244.82)	-44.82(228.1)
arep_r ACC	-0.099(0.167)	0.058(0.139)	-0.061(0.14)	-0.034(0.163)
arep_r RT	45.82(292.47)	91.62(186.32)	66.13(243.81)	67.86(244.53)

Note. Mean(SD). rep = repeated items; arep = post repetition items. The rescaled variables were computed by subtracting Baseline performance ($r = 0$) from performance at repetitions 1, 2, and 3, with resulting variables indexing the difference between performance on each repetition compared to Baseline.

Table s-4. Interference effect: Results of linear mixed-effects model comparison for performance accuracy. Best fitting model in bold.

Model	Fixed effects	Deviance	df	AIC[b]	Comparison[c]	χ^2(df)	p
						LR test[a]	
ma0[d]		-239.43	3	-233.43	–	–	–
ma1a	Interf	-272.39	4	-264.39	ma0-ma1a	32.96(1)	<.001
ma2a	Interf, nvIQ, Vocab	-283.1	6	-271.1	ma1a-ma2a	10.71(2)	0.004
ma2b	Interf, nvIQ, Vocab, Interf*nvIQ	-284.24	7	-270.24	ma2a-ma2b	1.14(1)	0.285
ma2c	Interf, nvIQ, Vocab, Interf*Vocab	-287.04	7	-273.04	ma2a-ma2c	3.94(1)	0.047
ma3a	Interf, nvIQ, Vocab, Interf*Vocab, SLI, TLD-L	-289.01	9	-271.01	ma2c-ma3a	1.97(2)	0.373
ma3c	**Interf, nvIQ, Vocab, Interf*Vocab, Interf*SLI**	**-293.68**	**8**	**-277.68**	**ma2c-ma3c**	**6.64(1)**	**0.01**
ma3d	Interf, nvIQ, Vocab, Interf*Vocab, Interf*SLI, Interf*TLD-L	-293.72	9	-275.72	ma3c-ma3d	0.04(1)	0.842
ma3e	Interf, nvIQ, Vocab, Interf*Vocab, Interf*SLI, Interf*SLI*Vocab	-294.56	9	-276.56	ma3c-ma3e	0.88(1)	0.348

Note. Participants included as random effects in all models. N=66. Main effects of nonverbal IQ and Vocabulary were kept, regardless of significance level, in order to control for these covariates. [a]Likelihood-ratio test. [b]Akaike Information Criterion. [c]Comparison between current and best model built thus far. [d]Intraclass Correlation Coefficient (ICC) for the null model = .533.
Interf = interference condition; nvIQ = nonverbal IQ (TONI-3); Vocab = Vocabulary (EOWPVT); SLI = children with specific language impairment; TLD-L = language-matched typically developing children.

Table s-5. Interference effect: Results of linear mixed-effects model comparison for Reaction Time data. Best fitting model in bold.

Model	Fixed effects	Deviance	df	AIC[b]	Comparison[c]	χ^2(df)	p
mr0[d]		1797.9	3	1803.9	–	–	–
mr1a	Interf	1792.7	4	1800.7	mr0-mr1a	5.19(1)	0.023
mr2a	Interf, nvIQ, Vocab	1777.1	6	1789.1	mr1a-mr2a	15.62(2)	<.001
mr2b	Interf, nvIQ, Vocab, Interf*nvIQ	1775.3	7	1789.3	mr2a-mr2b	1.79(1)	0.18
mr2c	Interf, nvIQ, Vocab, Interf*Vocab	1771.9	7	1785.9	mr2a-mr2c	5.22(1)	0.022
mr3a	Interf, nvIQ, Vocab, Interf*Vocab, SLI	1770.6	8	1786.6	mr2c-mr3a	1.31(1)	0.252
mr3b	**Interf, nvIQ, Vocab, Interf*Vocab, TLD-L**	**1761.9**	**8**	**1777.9**	**mr2c-mr3b**	**9.94(1)**	**0.002**
mr3c	Interf, nvIQ, Vocab, Interf*Vocab, TLD-L, Interf*SLI	1761.9	9	1779.9	mr3b-mr3c	0.04(1)	0.84
mr3d	Interf, nvIQ, Vocab, Interf*Vocab, TLD-L, Interf*TLD-L	1760.8	9	1778.8	mr3b-mr3d	1.13(1)	0.288
mr3e	Interf, nvIQ, Vocab, Interf*Vocab, TLD-L, Interf*SLI, Interf*SLI*Vocab	1760.5	10	1780.5	mr3b-mr3e	1.46(2)	0.482
mr3f	Interf, nvIQ, Vocab, Interf*Vocab, TLD-L, Interf*TLD-L, Interf*TLD-L*Vocab	1760.6	10	1780.6	mr3b-mr3f	1.33(2)	0.514

Note. Participants included as random effects in all models. N=66. Main effects of nonverbal IQ and Vocabulary were kept, regardless of significance level, in order to control for these covariates. [a]Likelihood-ratio test. [b]Akaike Information Criterion. [c]Comparison between current and best model built thus far. [d]Intraclass Correlation Coefficient (ICC) for the null model = .842.
Interf = interference condition; nvIQ = nonverbal IQ (TONI-3); Vocab = Vocabulary (EOWPVT); SLI = children with specific language impairment; TLD-L = language-matched typically developing children.

Table s-6. Repeated presentation of items effect: Results of linear mixed-effects model comparison for performance accuracy. Best fitting model in bold.

Model	Fixed effects	Deviance	df	AIC[b]	Comparison[c]	$\chi^2(df)$	p
ma0[d]		-317.4	3	-311.4	–	–	–
ma1a	rep	-326.7	4	-318.7	ma0-ma1a	9.33(1)	0.002
ma1b	rep, rep2	-331.5	5	-321.5	ma1a-ma1b	4.85(1)	0.028
ma1c	rep, rep2, Baseline	-336.2	6	-324.2	ma1b-ma1c	4.7(1)	0.03
ma2a	rep, rep2, Baseline, nvIQ, Vocab	-349.3	8	-333.3	ma1c-ma2a	13.08(2)	0.001
ma2b	rep, rep2, Baseline, nvIQ, Vocab, rep*nvIQ	-349.5	9	-331.5	ma2a-ma2b	.16(1)	0.685
ma2c	rep, rep2, Baseline, nvIQ, Vocab, rep2*nvIQ	-349.4	9	-331.4	ma2a-ma2c	.12(1)	0.73
ma2bc	rep, rep2, Baseline, nvIQ, Vocab, rep*nvIQ, rep2*nvIQ	-349.5	10	-329.5	ma2a-ma2bc	.19(2)	0.91
ma2d	rep, rep2, Baseline, nvIQ, Vocab, rep*Vocab	-349.7	9	-331.7	ma2a-ma2d	.37(1)	0.541
ma2e	rep, rep2, Baseline, nvIQ, Vocab, rep2*Vocab	-349.4	9	-331.4	ma2a-ma2e	.1(1)	0.751
ma2de	rep, rep2, Baseline, nvIQ, Vocab, rep*Vocab, rep2*Vocab	-350.6	10	-330.6	ma2a-ma2de	1.26(2)	0.533
ma3a	rep, rep2, Baseline, nvIQ, Vocab, SLI	-350.3	9	-332.3	ma2a-ma3a	.97(1)	0.325
ma3b	rep, rep2, Baseline, nvIQ, Vocab, TLD-L	-349.6	9	-331.6	ma2a-ma3b	.28(1)	0.595
ma3c	rep, rep2, Baseline, nvIQ, Vocab, rep*SLI	-349.8	9	-331.8	ma2a-ma3c	.47(1)	0.49
ma3d	**rep, rep2, Baseline, nvIQ, Vocab, rep*SLI, rep2*SLI**	**-353.4**	**9**	**-335.4**	**ma2a-ma3d**	**4.07(1)**	**0.043**
ma3e	rep, rep2, Baseline, nvIQ, Vocab, rep*SLI, rep2*SLI, rep*TLD-L	-353.5	11	-331.5	ma3d-ma3e	.06(2)	0.97
ma3f	rep, rep2, Baseline, nvIQ, Vocab, rep*SLI, rep2*SLI, rep*TLD-L, rep2*TLD-L	-354.8	12	-330.8	ma3d-ma3f	1.41(3)	0.704

Note. Participants included as random effects in all models. N=66. Main effects of nonverbal IQ and Vocabulary were kept, regardless of significance level, in order to control for these covariates. [a]Likelihood-ratio test. [b]Akaike Information Criterion. [c]Comparison between current and best model built thus far. [d]Intraclass Correlation Coefficient (ICC) for the null model = .419.
rep = repetition number (0-3); rep2 = repetition number squared (0-9); nvIQ = nonverbal IQ (TONI-3); Vocab = Vocabulary (EOWPVT); SLI = children with specific language impairment; Baseline = see text for description.

Table s-7. Repeated presentation of items effect: Results of linear mixed-effects model comparison for Reaction Time data. Best fitting model in bold.

Model	Fixed effects	Deviance	df	AIC[b]	Comparison[c]	χ^2(df)	p
					LR test[a]		
mr0[d]		3604.2	3	3610.2	–	–	–
mr1a	rep	3593.8	4	3601.8	mr0-mr1a	10.32(1)	0.001
mr1b	rep, rep2	3593.3	5	3603.3	mr1a-mr1b	.5(1)	0.478
mr1c	rep, Baseline	3568.3	5	3578.3	mr1a-mr1c	25.51(1)	<.001
mr2a	**rep, Baseline, nvIQ, Vocab**	**3564.3**	**7**	**3578.3**	**mr1c-mr2a**	**4.02(2)**	**0.133**
mr2b	rep, Baseline, nvIQ, Vocab, rep*nvIQ	3564.2	8	3580.2	mr2a-mr2b	.15(1)	0.701
mr2d	rep, Baseline, nvIQ, Vocab, rep*Vocab	3563.2	8	3579.2	mr2a-mr2d	1.09(1)	0.297
mr3a	rep, Baseline, nvIQ, Vocab, SLI	3563.5	8	3579.5	mr2a-mr3a	.8(1)	0.372
mr3b	rep, Baseline, nvIQ, Vocab, TLD-L	3564.2	8	3580.2	mr2a-mr3b	.11(1)	0.736
mr3c	rep, Baseline, nvIQ, Vocab, rep*SLI	3563.4	8	3579.4	mr2a-mr3c	.89(1)	0.346
mr3e	rep, Baseline, nvIQ, Vocab, rep*TLD-L	3563.5	8	3579.5	mr2a-mr3e	.78(1)	0.376

Note. Participants included as random effects in all models. N=66. Main effects of nonverbal IQ and Vocabulary were kept, regardless of significance level, in order to control for these covariates. [a]Likelihood-ratio test. [b]Akaike Information Criterion. [c]Comparison between current and best model built thus far. [d]Intraclass Correlation Coefficient (ICC) for the null model = .53.
rep = repetition number (0-3); rep2 = repetition number squared (0-9); nvIQ = nonverbal IQ (TONI-3); Vocab = Vocabulary (EOWPVT); SLI = children with specific language impairment; Baseline = see text for description.

Table s-8. Effect of repetition benefit on post-repetition items: Results of linear mixed-effects model comparison for performance accuracy. Best fitting model in bold.

Model	Fixed effects	Deviance	*df*	AIC[b]	Comparison[c]	χ^2(*df*)	*p*
ma0[d]		-243.42	3	-237.42	–	–	–
ma1a	Baseline	-243.53	4	-235.53	ma0-ma1a	.11(1)	0.745
ma2a	Baseline, repACC	-243.98	5	-233.98	ma1a-ma2a	.46(1)	0.498
ma2b	Baseline, repACC, repACC2	-267.11	6	-255.11	ma1a-ma2b	23.58(2)	<.001
ma3a	Baseline, repACC, repACC2, nvIQ, Vocab	-272.06	8	-256.06	ma2b-ma3a	4.95(2)	0.084
ma3b	Baseline, repACC, repACC2, nvIQ, Vocab, SLI	-283.26	9	-265.26	ma3a-ma3b	11.2(1)	<.001
ma3c	Baseline, repACC, repACC2, nvIQ, Vocab, SLI, TLD-L	-294.44	10	-274.44	ma3b-ma3c	11.18(1)	<.001
ma3d	Baseline, repACC, repACC2, nvIQ, Vocab, SLI, TLD-L, SLI*repACC	-296.22	11	-274.22	ma3c-ma3d	1.78(1)	0.181
ma3e	**Baseline, repACC, repACC2, nvIQ, Vocab, SLI, TLD-L, SLI* repACC2**	**-300.09**	**11**	**-278.09**	**ma3c-ma3e**	**5.65(1)**	**0.017**
ma3f	Baseline, repACC, repACC2, nvIQ, Vocab, SLI, TLD-L, SLI*repACC, SLI* repACC2	-301.89	12	-277.89	ma3e-ma3f	1.81(1)	0.178
ma4a	Baseline, repACC, repACC2, nvIQ, Vocab, SLI, TLD-L, SLI* repACC2, TLD-L*repACC	-300.13	12	-276.13	ma3e-ma4a	.04(1)	0.836
ma4b	Baseline, repACC, repACC2, nvIQ, Vocab, SLI, TLD-L, SLI* repACC2, TLD-L* repACC2	-303.22	12	-279.22	ma3e-ma4b	3.13(1)	0.077
ma4c	Baseline, repACC, repACC2, nvIQ, Vocab, SLI, TLD-L, SLI* repACC2, TLD-L*repACC, TLD-L* repACC2	-303.51	13	-277.51	ma3e-ma4c	3.43(2)	0.18

Note. Participants included as random effects in all models. N=66. Main effects of nonverbal IQ and Vocabulary were kept, regardless of significance level, in order to control for these covariates. [a]Likelihood-ratio test. [b]Akaike Information Criterion. [c]Comparison between current and best model built thus far. [d]Intraclass Correlation Coefficient (ICC) for the null model = .332.
repACC = repetition benefit; repACC2 = repetition benefit squared; nvIQ = nonverbal IQ (TONI-3); Vocab = Vocabulary (EOWPVT); SLI = children with specific language impairment; TLD-L = language-matched typically developing children; Baseline = see text for description.

Table s-9. Effect of repetition benefit on post-repetition items: Results of linear mixed-effects model comparison for Reaction Time data. Best fitting model in bold.

Model	Fixed effects	Deviance	*df*	AIC[b]	Comparison[c]	$\chi^2(df)$	*p*
					LR test[a]		
mr0[d]		3635.3	3	3641.3	–	–	–
mr1	Baseline	3633.6	4	3641.6	mr0-mr1	1.7(1)	0.191
mr2a	Baseline, repRT	3624.3	5	3634.3	mr1-mr2a	9.27(1)	0.002
mr2b	Baseline, repRT, repRT2	3624	6	3636	mr2a-mr2b	.25(1)	0.617
mr3a	Baseline, repRT, repRT2, nvIQ, Vocab	3623.6	8	3639.6	mr2b-mr3a	.41(2)	0.815
mr3b	Baseline, repRT, repRT2, nvIQ, Vocab, SLI	3621.9	9	3639.9	mr3a-mr3b	1.75(1)	0.186
mr3c	Baseline, repRT, repRT2, nvIQ, Vocab, TLD-L	3623.5	9	3641.5	mr3a-mr3c	.08(1)	0.774
mr3d	Baseline, repRT, repRT2, nvIQ, Vocab, SLI*repRT	3622.4	9	3640.4	mr3a-mr3d	1.25(1)	0.263
mr3d2	Baseline, repRT, repRT2, nvIQ, Vocab, SLI*repRT2	3615.9	9	3633.9	mr3a-mr3d2	7.77(1)	0.005
mr3d3	Baseline, repRT, repRT2, nvIQ, Vocab, SLI*repRT, SLI*repRT2	3613.6	10	3633.6	mr3a-mr3d3	9.99(2)	0.006
mr3e	Baseline, repRT, repRT2, nvIQ, Vocab, SLI*repRT, SLI*repRT2, TLD-L*repRT	3613.6	11	3635.6	mr3d3-mr3e	.03(1)	0.854
mr3e2	**Baseline, repRT, repRT2, nvIQ, Vocab, SLI*repRT, SLI*repRT2, TLD-L*repRT2**	**3609.2**	**11**	**3631.2**	**mr3d3-mr3e2**	**4.49(1)**	**0.034**
mr3e3	Baseline, repRT, repRT2, nvIQ, Vocab, SLI*repRT, SLI*repRT2, TLD-L*repRT, TLD-L*repRT2	3606.5	12	3630.5	mr3e2-mr3e3	2.61(1)	0.106

Note. Participants included as random effects in all models. N=66. Main effects of nonverbal IQ and Vocabulary were kept, regardless of significance level, in order to control for these covariates. [a]Likelihood-ratio test. [b]Akaike Information Criterion. [c]Comparison between current and best model built thus far. [d]Intraclass Correlation Coefficient (ICC) for the null model = .226.
repRT = repetition benefit; repRT2 = repetition benefit squared; nvIQ = nonverbal IQ (TONI-3); Vocab = Vocabulary (EOWPVT); SLI = children with specific language impairment; TLD-L = language-matched typically developing children; Baseline = see text for description.