

**Supplemental Table S6.** Results of the meta-regression analyses on generalization outcomes.

Moderator	Coefficient ( $\beta$ )	95% CI	<i>p</i> value
<b>Participant-related</b>			
Mean age	-0.08	[-0.19; 0.03]	.11
<b>L2 proficiency</b>			
Advanced (reference)	0.43	[-1.32; 2.18]	.57
Intermediate	-0.08	[-1.92; 1.75]	.92
Beginner	-0.06	[-1.88; 1.77]	.94
<b>L1</b>			
Chinese (reference)	0.24	[-0.32; 0.80]	.35
English	0.18	[-0.49; 0.85]	.54
<b>Methodology-related</b>			
<b>Target phone</b>			
Tone (reference)	0.37	[-0.08; 0.82]	.10
Vowel	-0.002	[-0.65; 0.64]	.99
<b>Target language</b>			
Artificial tone language (reference)	0.79	[0.09; 1.50]	.03
Chinese	-0.65	[-1.51; 0.22]	.11
English	-0.55	[-1.44; 0.33]	.17
French	-0.20	[-1.19; 0.80]	.63
<b>Training task type</b>			
Identification (reference)	0.42	[0.06; 0.78]	.03
Passive listening	-0.21	[-0.90; 0.49]	.50
<b>Training stimulus type</b>			
Pseudo word (reference)	0.36	[-0.09; 0.82]	.10
Real word	0.01	[-0.64; 0.65]	.98
<b>Talker number</b>			
Talker number	0.13	[-0.25; 0.50]	.45
<b>Talker presentation</b>			
<b>Blocked (reference)</b>	<b>0.68</b>	<b>[0.30; 1.07]</b>	<b>.004**</b>
<b>Intermixed</b>	<b>-0.50</b>	<b>[-0.99; -0.02]</b>	<b>.045*</b>
Target phone number	0.04	[-0.30; 0.37]	.80
Phonetic context number	0.01	[-0.01; 0.03]	.50

Training session number	-0.03	[-0.15; 0.09]	.57
Training session length	0.01	[-0.01; 0.03]	.19
Token number per session	-0.001	[-0.009; 0.007]	.85
Total token number	-0.00	[-0.001; 0.001]	.64
Training program length	0.08	[-0.06; 0.21]	.21
Training program time span	0.01	[-0.02; 0.04]	.40
Sample size	-0.02	[-0.05; 0.02]	.25
Test stimuli			
Trained + untrained (reference)	0.21	[-0.44; 0.86]	.46
Trained	0.16	[-0.64; 0.96]	.64
Untrained	0.37	[-0.65; 1.39]	.41
Study-related			
Year of publication	-0.02	[-0.08; 0.03]	.36
Source of publication			
Conference (reference)	0.26	[-0.55; 1.07]	.47
Dissertation	0.30	[-0.81; 1.42]	.53
Journal	0.08	[-0.85; 1.00]	.85
Study quality	-0.45	[-6.24; 5.35]	.86