

Figure 1: Accuracy scores, t-tests, and inter-task correlations for Tone Project behavioral study

		mean	std dev	t-stat	p-value	corr	p-value	n
<b>Chinese:</b>	Dyads	91.01	10.11					
	Words	97.67	4.42	4.047*	.000	-0.015	0.921	46
<b>English:</b>	Dyads	94.34	7.45					
	Words	86.17	14.78	4.270*	.000	.464*	0.001	47
<b>Chinese:</b>	Melodies	88.14	9.47					
	Sentences	95.87	4.79	5.169*	.000	0.107	0.479	46
<b>English:</b>	Melodies	83.38	10.07					
	Sentences	76.87	15.20	2.854*	.006	.289*	0.049	47

Figure 2: Performance scores on the *different* tasks by cluster group

<b>English:</b>	<b>Group 1 (<i>n</i>=11)</b>	<b>Group 2 (<i>n</i>=30)</b>	<b>Group 3 (<i>n</i>=6)</b>
Dyads	96.7 (4.9)	93.5 (8.0)	72.4 (12.3)*
Melodies	82.5 (14.4)	89.7 (11.9)	71.3 (15.4)*
Words	85.9 (18.0)	81.1 (19.3)	29.2 (25.8)*
Sentences	65.8 (20.5)	70.1 (20.4)	16.3 (16.9)*
<b>Chinese:</b>	<b>Group 1 (<i>n</i>=15)</b>	<b>Group 2 (<i>n</i>=21)</b>	<b>Group 3 (<i>n</i>=10)</b>
Notes	81.6 (18.3)	96.5 (5.5)	55.7 (9.1)*
Melodies	95.6 (8.2)	97.4 (5.9)	75.1 (14.2)*
Words	100 (0.0)	98.8 (3.7)	100 (0.0)
Sentences	97.8 (4.5)	95.1 (5.3)	92.8 (8.4)

% correct score on *different* tasks (std. dev);  
items with an asterisk are significant from the other within-language groups at  $p < .05$

Figure 3: Melody and working memory scores for High and Low cluster groups

	<b>English</b>		<b>Chinese</b>	
	<b>High (<i>n</i>=31)</b>	<b>Low (<i>n</i>=16)</b>	<b>High (<i>n</i>=35)</b>	<b>Low (<i>n</i>=11)</b>
<b>Music melody:</b>				
Melodies (% total)	89.1 (6.3)	71.1 (7.0)*	92.6 (5.0)	74.8 (8.0)*
Melodies (% same)	87.1 (10.6)	68.0 (14.7)*	87.9 (10.2)	74.9 (12.1)*
Melodies (% different)	91.2 (10.2)	74.9 (14.9)*	97.3 (5.4)	74.9 (13.8)*
Melodies ( <i>d'</i> )	3.39 (1.27)	1.42 (0.76)*	4.16 (1.02)	1.56 (0.71)*
<b>Working memory:</b>				
Verbal	86.4 (12.9)	85.3 (11.8)	82.3 (11.7)	82.5 (8.2)
Spatial (length)	77.6 (13.5)	76.3 (11.6)	84.1 (9.9)	83.8 (14.4)
Spatial (direction)	66.4 (17.4)	64.5 (12.9)	75.3 (13.3)	77.3 (14.5)

\* Difference between English high/low and Chinese high/low significant for each score at  $p < .001$ .  
There are no significant differences between groups for any working memory or language task scores.

Figure 4: Intertask relationships

***Controlling for effects of Spatial (length) working memory***

	<b>Music</b>	<b>Working memory</b>	<b>Partial correlation</b>
<b>English Low:</b>	Melodies (d')	Spatial (direction)	.576*
<b>Chinese Low:</b>	Dyads (d')	Spatial (direction)	.640*

***Controlling for effects of Spatial (direction) working memory***

	<b>Language</b>	<b>Working memory</b>	<b>Partial correlation</b>
<b>Chinese Low:</b>	Sentences (d')	Verbal	.815*
<b>Chinese Low:</b>	Sentences (d')	Spatial (length)	.678*

\*  $p < .05$