Table S1. *Participant characteristics by group.*

	Group				
Measure	***Trilinguals*** *(N=45, 39 females)*			***Bilinguals*** *(N = 47, 34 females)*	
Demographic Variables					
Education (years)	14.900 (1.829)			14.031 (2.073)	
Maternal Education (years)	13.466 (3.532)			14.542 (3.291)	
Cognitive Variables					
Backward Color Span	4.688 (.668)			4.851 (.721)	
Numerical Double-Stroop Test	40.400 (13.723)			39.893 (10.899)	
Language Experience Variables	***Arabic (L1)***	***Hebrew (L2)***	***English (L3)***	***Hebrew (L1)***	***English (L2)***
AoA (years)	Birth	7.555 (.867)	8.511 (.815)	Birth	7.436 (1.227)
Self-rated Proficiency (0-10)	9.750 (.398)	8.155 (1.018)	6.366 (1.843)	9.718 (.459)	7.425 (1.294)
Childhood Exposure (%)	80.605 (6.826)	11.538 (5.644)	7.855 (3.105)	82.404 (9.701)	16.335 (8.057)
Current Exposure (%)	61.160 (12.408)	32.084 (10.567)	6.622 (5.020)	80.629 (9.271)	17.340 (7.942)
Current Use (%)	51.329 (18.553)	32.477 (14.384)	15.822 (13.604)	64.283 (15.735)	34.370 (15.281)
Semantic Fluency (number of items)	22.600 (4.677)	16.844 (4.776)	12.266 (5.314)	32.936 (6.831)	21.702 (5.763)
Mint Sprint Test (range 0-80)	67.733 (6.304)	37.600 (12.034)	31.111 (8.901)	73.212 (4.032)	50.276 (10.761)

Note. SDs in parentheses. Percentages for childhood and current exposure and language use among Hebrew-English bilinguals may not add up to 100% (~98%) due to four participants reporting minimal exposure to an additional language, which was not integrated into their daily routines. More information about L1 and L2 of trilinguals can be found in Bsharat-Maalouf et al. 2024.

Table S2. *Summary model for pupil mean.*

	Single Words				Sentences			
Fixed effects	***b***	***SE***	***t***	***p***	***b***	***SE***	***t***	***p***
Intercept	.025	.003	7.177	**< .001**	.051	.005	9.497	**< .001**
Condition (Noise)	.818	.001	13.386	**< .001**	.0168	.001	12.318	**< .001**
Overall Language Experience	-.007	.001	-8.682	**< .001**	-.018	.001	-19.769	**< .001**
Context (High)	-	-	-	-	.005	.001	4.682	**< .001**
Condition (Noise): Overall Language Experience	-.001	.001	-1.012	.311	.007	.02	5.857	**< .001**
Context (High): Overall Language Experience	-	-	-	-	-.003	.001	-2.902	**.004**
Condition (Noise): Context (High): Overall Language Experience	-	-	-	-	.003	.001	1.929	**.045**
Maternal Education	-.002	.003	-.739	.462	-.003	.005	-.596	.553
Backward Color Span	.001	.003	.022	.983	.001	.005	.022	.983
Numerical Double-Stroop	.001	.003	.231	.818	.001	.005	.004	.997
Trial Order	-.007	.001	-12.794	**< .001**	-.012	.001	-28.114	**< .001**
Perceptual Accuracy	-.005	.001	-8.762	**< .001**	-.001	.001	-.676	.499
Random effects	***Var.***	***SD***			***Var.***	***SD***		
Item (Intercept)	.001	.005			.001	.005		
Participant (Intercept)	.001	.033			.002	.050		

Note. The observed results pattern closely aligns with the findings described for peak amplitude in the results section. In the sentence model, the context effect (high versus low predictability sentences) along with the interaction of this variable with listening conditions and overall language experience, is reported

Table S3. *Correlations between individual control variables.*

Variable	***1***	***2***	***3***	***4***
1 *Maternal Education*	-			
2 *Backward Color Span*	.027	-		
3 *Numerical Double-Stroop*	.010	.136	-	
4 *Overall Language Experience*	.055	.032	-.008	-

Note. All correlations were not significant $p > .05$.

Table S4. *Summary of the results.*

	Single words		**Sentences**	
Effect/interaction	***Perception***	***Pupillometry***	***Perception***	***Pupillometry***
Condition	Quiet > Noise	Quiet > Noise	Quiet > Noise	Quiet > Noise
Overall Language Experience	Greater language experience → Higher accuracy	Greater language experience → Smaller peak amplitudes and earlier peak latencies	Greater language experience → Higher accuracy	Greater language experience → Smaller peak amplitudes and earlier peak latencies
Condition X Overall Language Experience	n.s.	*Peak latency:* Effect of language experience Quiet < Noise	Effect of language experience Quiet < Noise	Effect of language experience Quiet > Noise
Context X Overall Language Experience	-	-	Effect of language experience High > Low	Effect of language experience High > Low
Condition X Context X Overall Language Experience	-	-	***Quiet:*** Effect of language experience High = Low ***Noise:*** Effect of language experience High > Low	*Peak amplitude:* ***Quiet:*** Effect of language experience High > Low ***Noise:*** Effect of language experience High = Low

Note. In perception, "x > y" means higher perception in x than y, while in pupillometry, it reflects greater effort. "n.s." indicates a non-significant interaction. "High" and "Low" refer to high- and low-predictability sentences, respectively. If no specific pupillometry measure is mentioned, significance applies to both peak amplitude and latency; otherwise, non-significance is implied for the other measure.

Table S5. *Pairwise comparisons for significant interactions using Bonferroni corrections.*

Single Words - Peak Latency				
Listening Condition x Overall Language Experience				
	value	*df*	χ^2	*p*
Quiet	-17.242	1	1.927	.330
Noise	-52.521	1	18.226	< .001
Sentences - Perceptual Accuracy				
Listening Condition x Context x Overall Language Experience				
	value	*df*	χ^2	*p*
Quiet: Low-High	*-.013*	*1*	*4.591*	.064
Noise: Low-High	*-.074*	*1*	*112.201*	< .001
Sentences – Peak Amplitude				
Listening Condition x Context x Overall Language Experience				
	value	*df*	χ^2	*p*
Quiet: Low-High	*.005*	*1*	*10.951*	.0018
Noise: Low-High	*.003*	*1*	*.056*	1.000
Sentences – Peak Latency				
Listening Condition x Overall Language Experience				
	value	*df*	χ^2	*p*
Quiet	*-236.830*	*1*	*303.643*	< .001
Noise	*-109.12*	*1*	*55.065*	< .001

Note. Pairwise comparisons were conducted using the *phia* package in R. The table presents the significant higher-level interaction observed. For a breakdown of simpler effects, refer to *Table S4*.