

Supplemental Material S1. Intensity, pitch fundamental frequency (f_0) of the T2 and T4 starting points and ending points, and tone duration averaged across the four syllables (“*ba*,” “*pa*,” “*da*,” and “*ta*”) used in the experiment for the high-variability (H+) stimuli and the low-variability (H–) stimuli.

	T2 Intensity (dB)	T4 Intensity (dB)	T2 starting pitch f_0 (Hz)	T2 ending pitch f_0 (Hz)	T4 starting pitch f_0 (Hz)	T4 ending pitch f_0 (Hz)	T2 duration (ms)	T4 duration (ms)
High-variability (H+) stimuli: Produced by six different speakers, one time for each speaker								
Speaker 1	78	78	140	176	224	209	552	378
Speaker 2	78	78	129	167	194	161	433	288
Speaker 3	76	77	84	113	124	99	428	346
Speaker 4	79	79	208	227	323	188	697	507
Speaker 5	75	76	206	242	296	197	462	374
Speaker 6	76	77	188	254	276	126	740	673
Mean	77	78	159	197	240	163	552	428
Low-variability (H–) stimuli: Produced by only Speaker 1, six times for this speaker								
Time 1	78	78	140	176	224	209	552	378
Time 2	79	78	140	173	226	216	485	331
Time 3	78	77	145	179	239	219	560	345
Time 4	78	77	143	186	242	189	546	370
Time 5	78	78	140	168	226	212	528	346
Time 6	78	77	150	180	252	224	558	348
Mean	78	77	143	177	235	212	538	353

Note. The value of average intensity was measured on speakers’ pronunciations while recording. All measures were recorded using Praat Version 6.1.53. One thing to explain is that, the intensity or the duration of the T2 and the T4 can be a cue for the participants when identifying these tone pairs. Nevertheless, we did not modify them in order to allow the stimuli to be as natural as possible, given that we controlled for the differences in intensity and duration across all experimental conditions.