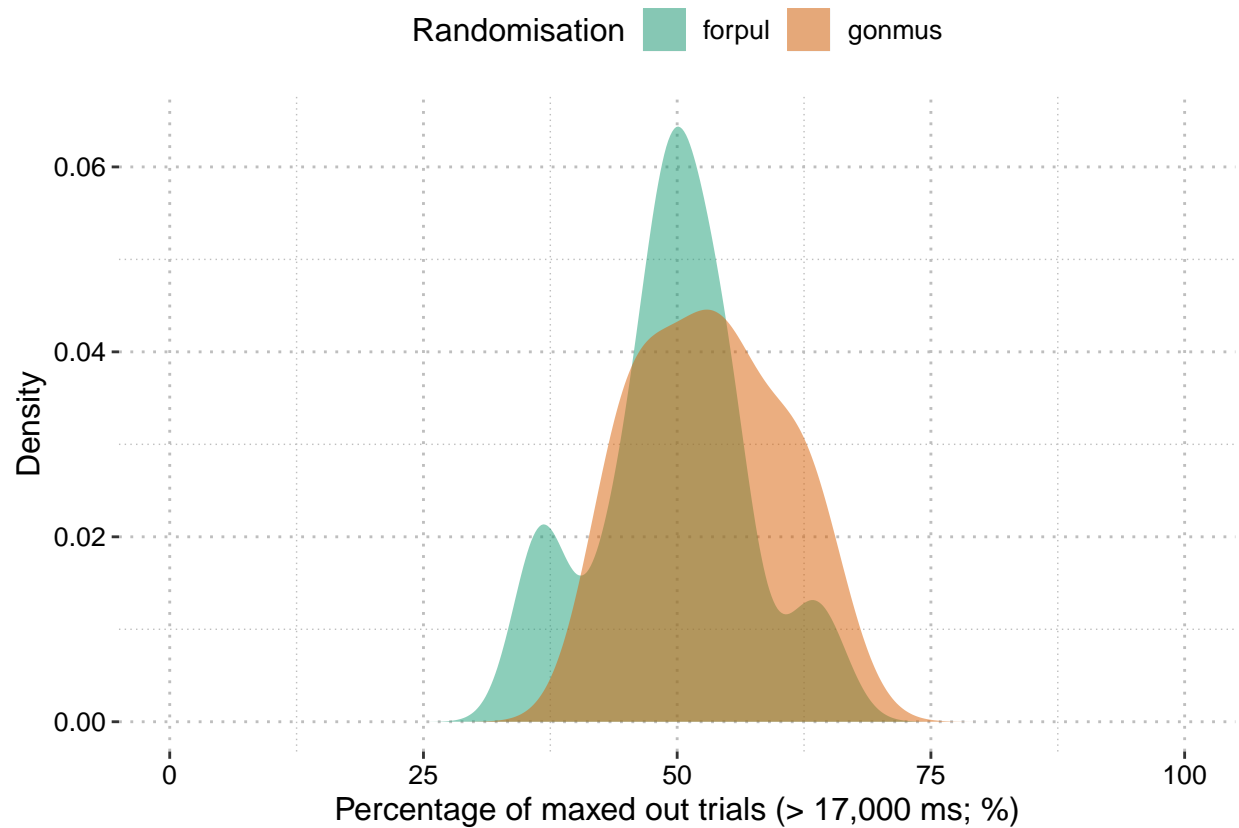


Appendix 3: Results without maxed out trials

During data collection, one of the concerns raised by one of the testers was that many infants reached the maximum looking time by trial very frequently. If this was the case, it would be possible that the lack of statistically significant differences was not due to the absence of preference toward one of the stimuli types, but rather because of a ceiling effect. To rule out this possibility we first calculated the proportion of trials in which infants looked more than 17,000 ms, then discarded those trials, and re-run the same analysis as with the original data.

Language	Randomisation	N	Familiar (ms)	Novel (ms)
cat	forpul	14	8529.48	8595.15
cat	gonmus	9	8033.05	9268.69
spa	forpul	7	9026.73	8838.25
spa	gonmus	6	9654.79	10161.54



Randomisation	Word	Maxed out	Total	%	Type
forpul	for	4	62	6.45	familiar
forpul	gon	9	63	14.29	novel
forpul	mus	10	63	15.87	novel
forpul	pul	6	63	9.52	familiar
gonmus	for	3	45	6.67	novel
gonmus	gon	6	45	13.33	familiar
gonmus	mus	7	44	15.91	familiar
gonmus	pul	4	44	9.09	novel

Looking times were Filtered ($\leq 17,000$ ms) and aggregated by participant.

Trial type	N	Mean	Median	SEM	95% CI
familiar	36	8634.21	8466.12	426.85	[7797.6, 9470.81]
novel	36	9126.68	9450.42	410.97	[8321.19, 9932.18]

Familiar	Unfamiliar	T-value	DF	95% CI	p-value	Cohen's D
8634.206	9126.683	-1.186	35	[-1335.29, 350.34]	0.244	0.198