

Agentic AI Report

Table of contents

Cleaning	1
Baseline model: frequency and strokes	2
Frequency effect on recognition speed	2

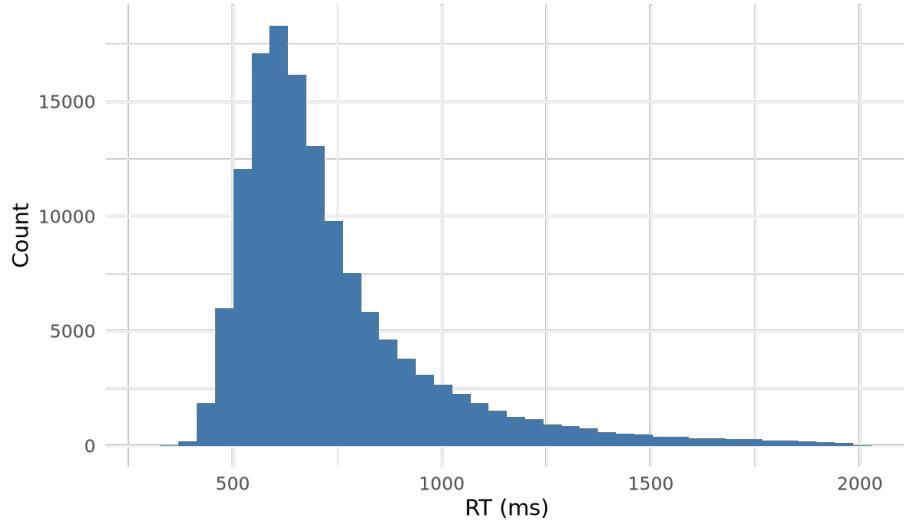
here() starts at /root/repo

Cleaning

The pipeline kept 137133 of 235016 trials (dropped 97883). Settings: correct-only = TRUE, RT range = 200–2000 ms.

	setting	value
1	correct_only	TRUE
2	rt_min_ms	200
3	rt_max_ms	2000
4	total_trials	235016
5	kept_trials	137133
6	dropped_trials	97883

RT histogram (kept 137133/235016 trials)



Baseline model: frequency and strokes

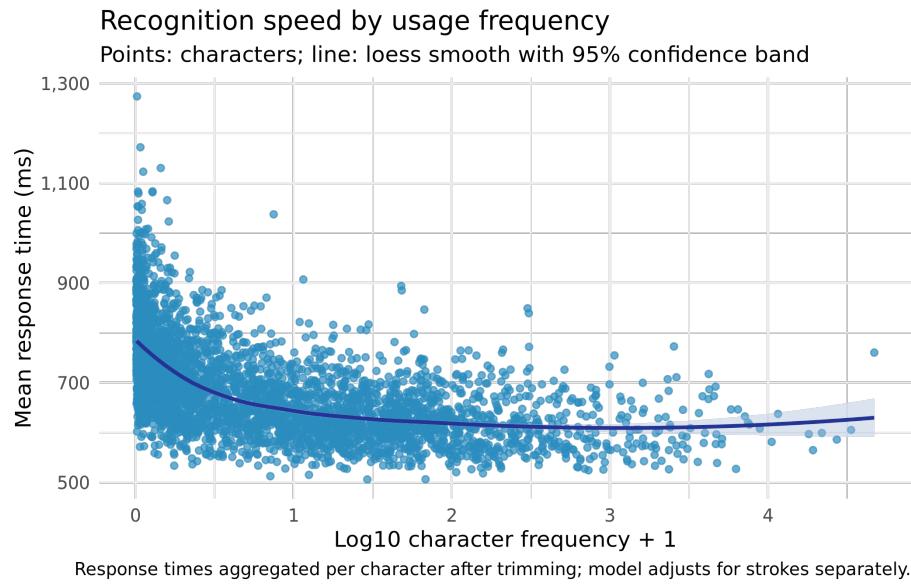
	term	estimate
1	intercept	6.452355
2	log_freq	-0.070823
3	strokes	0.013355

R² 0.434; adjusted R² 0.433; residual sigma 0.099. AIC -6851.160, BIC -6826.134.

Frequency effect on recognition speed

	term	estimate	conf_low	conf_high
(Intercept)	(Intercept)	6.452355	6.441170	6.463540
log_freq	log_freq	-0.070823	-0.074644	-0.067002
strokes	strokes	0.013355	0.012417	0.014294

A 1-unit increase in log10 frequency (roughly a tenfold usage jump) is associated with an estimated 6.837% faster recognition (95% CI 6.481% to 7.193%) when strokes are held at their median (10). This corresponds to predicted mean response times of approximately 723.216 ms (10th percentile frequency), 694.402 ms (median), and 624.486 ms (90th percentile). The marginal gains are largest through common characters—going from the median to the 90th percentile yields 69.916 ms faster responses—but they taper for the most extreme items: the 10th to 50th percentile shift saves about 28.814 ms, and the jump from the 90th to 95th percentile trims only 18.729 ms.



Response times aggregated per character after trimming; model adjusts for strokes separately.

A loess smooth over the character-level means reinforces the flattening: after \log_{10} frequency around 2.104, the curve levels off with uncertainty bands that overlap tightly, suggesting diminishing recognition-speed returns for the very most frequent characters.