

Simba 3.5

English and multilingual evaluations: listener preference, intelligibility and speed

Simba 3.5 is SpeechifyAI's text-to-speech model for 48 languages. This side covers its English evaluation, judged by US and UK listeners; the multilingual results are overleaf.

Preferred by listeners, faster under load. In blind listening tests, listeners prefer Simba 3.5 to Simba 3.2 in 63% of comparisons, and it rates ahead in all 10 content domains we tested.

63%¹

win rate against Simba 3.2 in blind side-by-side listening, all content (95% CI 60–67%)

1.8x²

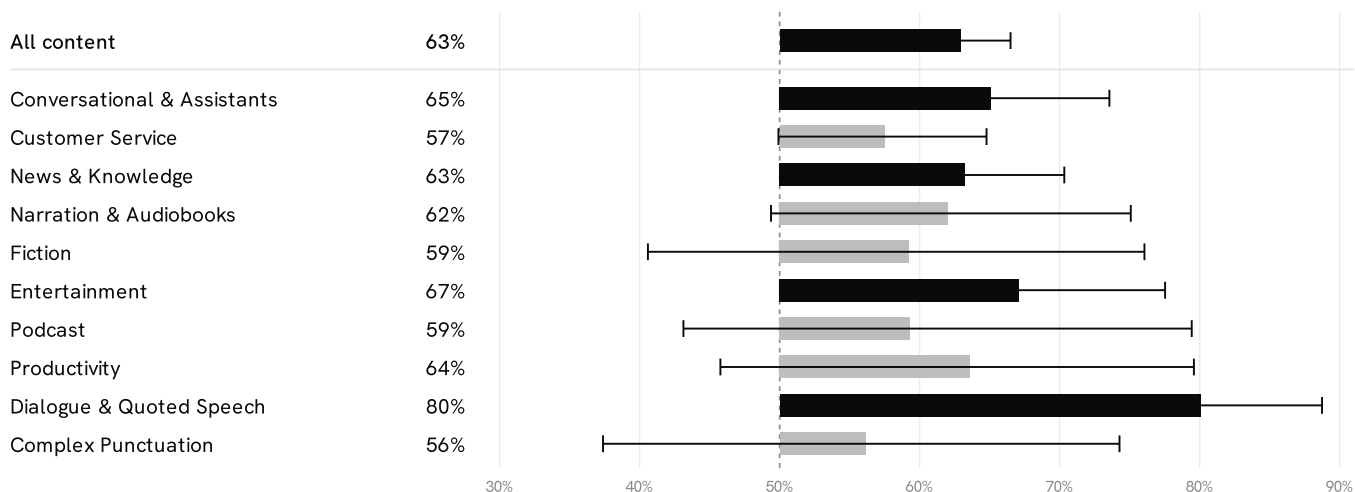
Simba 3.2's speech throughput per serving instance

72ms³

time to first audio for 95% of streaming requests, under production load

Listener preference by content domain

Simba 3.5 win rate against Simba 3.2



Bars show Simba 3.5's win rate against Simba 3.2, starting from the 50% parity line (dashed); whiskers are 95% confidence intervals. Dark bars: the interval lies entirely above 50% (4 of 10 domains and all content).

AVAILABILITY

Available now in the SpeechifyAI API as `simba-3.5`.

Hear it at speechify.ai/launch/simba-3-5.

Docs at docs.speechify.ai/build/guides/concepts/models.

For volume and enterprise: speechify.ai/talk-to-sales.

Try Simba 3.5

Talk to sales

¹ Blind pairwise listening tests: listeners hear the same text in the same voice from two systems, in random order, and choose the one they prefer. Nearly 65,000 judgments from nearly 6,000 US and UK English listeners across 10 content domains. Win rates are derived from Bradley-Terry (Elo-style) ratings fitted to all judgments, excluding ties, with 95% confidence intervals from resampling listeners.

³ 95th-percentile time from request to first audio while streaming, measured under the same conditions as note 2.

² Speech synthesized per second by one serving instance, with both models served in their production configuration (word timing marks on) on the same hardware and the same production text mix.

Preferred by listeners beyond English. Simba 3.5 covers 48 languages in one model, against 6 for Simba 3.0 Multilingual. In blind listening tests in the six locales both models speak, native listeners prefer Simba 3.5 to Simba 3.0 Multilingual in 62% of comparisons, and it rates ahead in all six, by a clear margin in German and Spanish (Spain). English results are overleaf.

62%⁴

win rate against Simba 3.0 Multilingual in blind side-by-side listening, six locales (95% CI 56–67%)

1.4x⁵

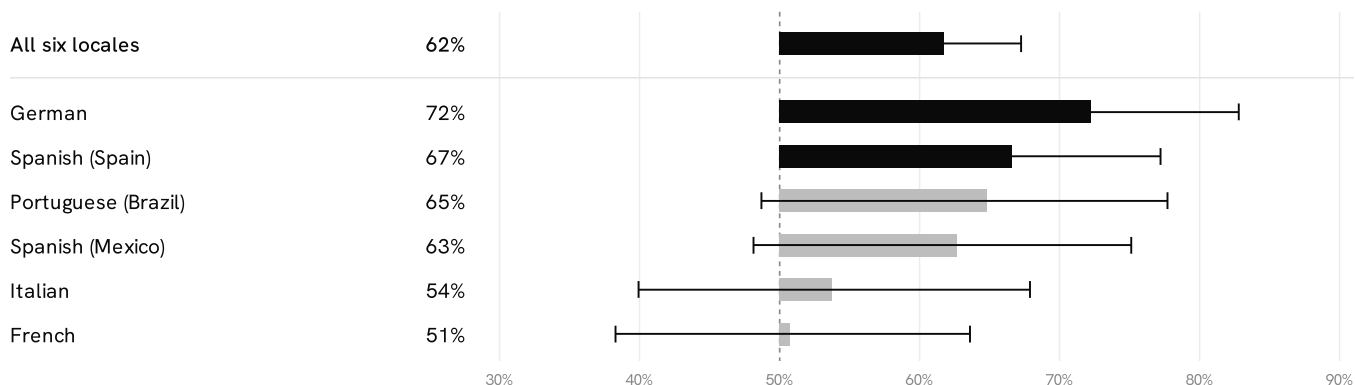
Simba 3.0 Multilingual’s speech throughput per serving instance, on average

48

languages in one model, against 6 for Simba 3.0 Multilingual

Listener preference by locale

Simba 3.5 win rate against Simba 3.0 Multilingual



Bars show Simba 3.5’s win rate against Simba 3.0 Multilingual, starting from the 50% parity line (dashed); whiskers are 95% confidence intervals. Dark bars: the interval lies entirely above 50%.

By content domain

Pooled over the six locales, listeners prefer Simba 3.5 by a clear margin in Complex Punctuation (70%), Fiction (69%), News & Knowledge (69%), Dialogue & Quoted Speech (65%) and Narration & Audiobooks (65%). In the other 4 content domains the two models are statistically on par.

Speed

Served side by side on the same hardware, each Simba 3.5 serving instance delivers on average 1.4x the speech throughput of a Simba 3.0 Multilingual instance.

Coverage

One model for 48 languages. Simba 3.0 Multilingual serves six languages: English, Spanish, Portuguese, French, German and Italian.

⁴ Blind pairwise listening tests in German, Spanish (Spain and Mexico), French, Italian and Portuguese (Brazil), judged by native-speaker listeners: nearly 24,000 judgments from over 2,000 listeners. Same method as note 1; per-locale win rates from a separate fit in each locale.

⁵ Speech synthesized per second by one serving instance, with both models served side by side on the same hardware in their production configuration (word timing marks on), with the same prompts; averaged over a standard prompt set.