ROUNDUP

Best AI Text-to-Speech Tools 2026: Pricing and Picks

Six text-to-speech tools compared on pricing, models and API fit, from ElevenLabs and Cartesia to the OpenAI, Google, AWS and Azure APIs.

Introduction

Text-to-speech now splits into two kinds of product. Voice platforms like ElevenLabs and Cartesia sell expressive voices, cloning and real-time agents on credit-based plans. Cloud APIs from OpenAI, Google, AWS and Microsoft sell speech by the million characters, with prices that drop sharply if you can live with plainer voices. Voice platforms suit content people will listen to closely. Cloud APIs suit a product that handles high volumes.

This roundup covers six tools, using each vendor's own pricing and documentation as of October 2026. scored.tools rates tools on five criteria: quality of output, ease of use, pricing value, API and integration quality, and how well the tool solves the problem it claims to solve. None of these six has a published scored.tools rating yet, so no scores appear below. The entries are in alphabetical order, which is not a ranking.

The tools

1. Amazon Polly

Score: Not yet rated

Amazon Polly is the AWS speech API, and its main appeal is price at scale. AWS lists four voice types. Standard costs $4.00 per million characters, Neural $16.00, Generative $30.00 and Long-Form $100.00. The free tier is large: 5 million Standard characters a month, plus 1 million Neural characters a month for the first 12 months. Polly fits best when your stack already runs on AWS and you need high-volume speech for IVR menus, notifications or accessibility features. The Standard voices are the cheapest option in this list. If expressive delivery matters, you'll have to move up to the Generative or Long-Form tiers, and Long-Form's $100 rate costs more than most dedicated voice platforms.

Best for: AWS teams that need cheap, high-volume speech.

Pricing: Pay as you go. $4 (Standard), $16 (Neural), $30 (Generative) or $100 (Long-Form) per million characters. The free tier covers 5M Standard characters a month.

2. Azure AI Speech

Score: Not yet rated

Azure Ai Speech is Microsoft's speech service. Its pricing page lists standard neural voices, Neural HD and HD Flash voices, MAI-Voice, Professional custom voices and Personal voice. The free F0 tier gives 0.5 million neural characters a month. For heavy users, Microsoft sells monthly commitment bundles of 80M, 400M or 2,000M neural characters. It also sells disconnected containers for on-premises deployment, which is rare in this category and matters for regulated industries. One drawback for buyers: the public pricing page doesn't print the per-character rates and sends you to the Azure calculator instead. That makes quick comparisons harder than with AWS or OpenAI. Custom voice also carries training and hosting charges on top of synthesis.

Best for: Enterprises on Microsoft infrastructure, or teams that need offline or containerised speech.

Pricing: Free tier of 0.5M neural characters a month. Paid use is billed per million characters, with monthly commitment tiers. Exact rates are in the Azure pricing calculator.

3. Cartesia

Score: Not yet rated

Cartesia is built for real-time voice agents. Its current model, Sonic 3.6, supports 44 languages. Per the docs, it adjusts pacing, intonation and pause length to fit the emotion of the text, with no SSML tags needed. Cartesia calls it its fastest model but doesn't publish a latency figure on that page. Developers will like the pinned snapshot IDs (for example `sonic-3.6-2026-08-27`), which keep a production voice from changing under you. Plans are cheap at the low end. Pro costs $5 a month for 100K credits and includes instant cloning and a commercial licence. Each plan also comes with prepaid voice-agent credit. The free plan has no commercial licence.

Best for: Developers building voice agents and phone bots where response time matters.

Pricing: Free (20K credits), Pro $5/mo (100K), Startup $49/mo (1.25M, 2 professional clone slots), Scale $299/mo (8M, 4 professional clone slots). Enterprise is custom.

4. ElevenLabs

Score: Not yet rated

Elevenlabs has the widest feature set here. It offers speech synthesis, instant and professional voice cloning, a dubbing studio, voice design and a long-form Studio editor, all in one account. The model docs list Eleven v3 for expressive output in 70+ languages and Multilingual v2 for 29 languages. Flash v2.5 and the newer Eleven v4 Turbo are the real-time options, with about 100ms median model

latency according to ElevenLabs (excluding network time). Speech costs 1 credit per character, and the Flash and Turbo models cost 0.5 to 1 credit per character over the API. The pricing page needs a careful read. The plan cards show first-month discounts, while the FAQ gives the regular rates as $6 for Starter and $22 for Creator. The jump from Creator to Pro at $99 is steep.

Best for: Creators and studios making narration, audiobooks, dubbing or cloned-voice content.

Pricing: Free (10K credits), Starter $6/mo (30K), Creator $22/mo (121K, professional cloning), Pro $99/mo (600K), Scale $299/mo (1.8M, 3 seats), Business $990/mo (6M, 10 seats). Regular rates are taken from the FAQ.

5. Google Cloud Text-to-Speech

Score: Not yet rated

Google Cloud Tts offers several voice tiers at different prices. Google's pricing page lists Standard at $4 per million characters and Neural2 at $16. Chirp 3: HD costs $30 per million after the first million free each month. Studio voices cost $160, and instant custom voice costs $60. The generous free allowance on Chirp 3: HD makes it easy to prototype with Google's better voices before paying anything. Google also has token-priced Gemini TTS models, which complicates the pricing. Check which billing model applies to the voice you pick before you set a budget.

Best for: Teams on Google Cloud that need many languages and a choice of price points.

Pricing: Pay as you go. $4 (Standard), $16 (Neural2), $30 (Chirp 3: HD, first 1M characters free monthly) or $160 (Studio) per million characters.

6. OpenAI TTS

Score: Not yet rated

Openai Tts is the easiest choice if your app already calls OpenAI models. tts-1 costs $15 per million characters and tts-1-hd costs $30. The newer gpt-4o-mini-tts model is billed by tokens instead: $0.60 per million text input tokens and $12 per million audio output tokens. That means you can't directly compare its cost with character-priced rivals without running your own estimates. The voice catalogue is small next to ElevenLabs or Azure, and OpenAI offers no self-serve voice cloning. For assistants and in-app narration where one decent voice is enough, it's simple to add to an existing stack.

Best for: Developers already building on the OpenAI API.

Pricing: tts-1 $15 and tts-1-hd $30 per million characters. gpt-4o-mini-tts is token-priced ($0.60/1M input, $12/1M audio output).

Comparison table

Tool	Score	Pricing model	Entry price	Free tier	Voice cloning
Amazon Polly	Not yet rated	Per character	$4 per 1M (Standard)	5M Standard chars/mo	No self-serve cloning listed
Azure Ai Speech	Not yet rated	Per character, commitment tiers	See Azure calculator	0.5M neural chars/mo	Professional and Personal voice
Cartesia	Not yet rated	Credit plans	$5/mo (Pro)	20K credits	Instant from Pro, professional from Startup
Elevenlabs	Not yet rated	Credit plans	$6/mo (Starter)	10K credits	Instant from Starter, professional from Creator
Google Cloud Tts	Not yet rated	Per character (some token-priced)	$4 per 1M (Standard)	1M Chirp 3: HD chars/mo	Instant custom voice ($60 per 1M)
Openai Tts	Not yet rated	Per character or per token	$15 per 1M (tts-1)	None listed	No self-serve cloning

Final picks

These picks follow from pricing and documentation, not from scores, since none of the six has a rating yet.

- **Narration, dubbing and cloned voices:** Elevenlabs. It has the most complete creator toolset, and professional cloning starts at the $22 Creator plan.
- **Real-time voice agents:** Cartesia. Sonic 3.6 is built for live conversation, pinned snapshots keep voices stable, and a $5 plan is enough to test it with a commercial licence.
- **Cheapest high-volume speech:** Amazon Polly or Google Cloud Tts, both at $4 per million Standard characters. Pick whichever cloud you already pay for.
- **Already on OpenAI:** Openai Tts. One API key and simple per-character pricing on tts-1.
- **Regulated or offline deployments:** Azure Ai Speech, because of its disconnected containers.

Prices in this category change often, and ElevenLabs in particular runs first-month promotions on its pricing page. Check each vendor's page on the day you buy.
