

MARKET RESEARCH

ElevenLabs Productions

IVOC Research

Completed and expanded from Hope's Sticky Market Research Template

Scope and evidence standard

Qualitative synthesis / inference. Raw customer language is quoted from public discussions where participants were not answering a commissioned survey. Official ElevenLabs sources are used only to map product capabilities. The sample is directional and is not statistically representative of the entire market.

Product	ElevenLabs Productions
Research date	July 24, 2026
Raw corpus	54 direct IVOC excerpts from 45 public discussion sources
Output	Seven-category IVOC summary, ranked benefit matrix, message implications, raw quote bank, and source index

Research frame

Dominant conscious desire

[Inference] Take existing high-value content global without sacrificing the voice, intent, accessibility, or professional polish that made it worth localizing in the first place.

Stage of awareness

[Inference] The market is mainly solution-aware and product-aware. Prospects know that AI dubbing, captioning, transcription, and synthetic narration exist. Many have tried tools or seen auto-dubbed content. Their central question is whether a managed AI-plus-human workflow can remove the cleanup work and quality risk that self-serve tools leave behind.

Places looked

- Reddit communities for creators, partnered YouTubers, video editors, filmmakers, localization professionals, instructional designers, podcasters, authors, audiobook listeners, caption users, and ElevenLabs users.
- Public product discussions about self-serve dubbing, voice cloning, transcription, audiobook production, and caption quality.
- YouTube Help for the current multi-language-audio workflow and global-reach context. [A01]
- W3C and FCC guidance for the functional standard behind high-quality captions. [A02, A03]
- ElevenLabs Productions for current service capabilities, workflow, languages, turnaround, and pricing mechanics. [P01-P08]

Market context

Signal	What it means for Productions
Global distribution	YouTube supports multiple language tracks on one video and reports that creators using the feature saw more than 25% of watch time from non-primary-language views. The opportunity is real, but creators still need finished audio tracks. [A01]
Accessibility quality	W3C describes captions as synchronized speech plus speaker identification and meaningful non-speech information. FCC guidance groups quality around accuracy, synchronicity, completeness, and placement. [A02, A03]
Product wedge	Productions combines ElevenLabs audio technology with human editing and project delivery. This directly addresses the market's gap between fast generation and publishable output. [P01, P05, P08]

IVOC Summary

How to read the summary

Items are ordered by qualitative recurrence and emotional intensity within the sampled corpus. Ranking reflects the research sample, not population-level frequency.

Pains and Fears

#	IVOC Summary
1	Localization is operationally hard and slows the cadence of publishing new content.
2	Traditional dubbing and audiobook production can be expensive enough to make ROI uncertain.
3	Self-serve AI often creates a hidden second job: correcting translation, timing, pronunciation, and delivery.
4	Poor dubs, captions, or narration can damage trust, comprehension, retention, and perceived production quality.
5	Multiple speakers, long files, accessibility requirements, and source updates multiply handoffs and rework.

Objections

#	IVOC Summary
1	AI translation cannot reliably handle ambiguity, idioms, humor, dialect, and cultural context without human review.
2	AI voices may sound robotic, emotionless, overly polished, or unlike the original speaker.
3	Buyers doubt that AI-first output is genuinely market-ready rather than merely understandable.
4	A managed service may look expensive beside cheap tools until correction time and failed output are included.
5	AI use can trigger ethical, consent, employment, authenticity, and audience-reaction concerns.

Desires

#	IVOC Summary
1	Reach new language audiences and grow watch time, subscribers, impact, and revenue from existing content.
2	Preserve the original voice identity, emotion, pacing, accent, and communicative intent.
3	Receive accurate, culturally appropriate, professionally finished assets that are ready to publish.

#	IVOC Summary
4	Reduce production time and avoid coordinating translators, voice talent, editors, captioners, and QA separately.
5	Make content accessible through complete, readable, synchronized captions and SDH options.

Firm Beliefs

#	IVOC Summary
1	Localization is adaptation, not a word-for-word language swap.
2	Human judgment remains indispensable at the quality-control and creative interpretation stage.
3	Performance, pronunciation, pauses, and tone are part of the content, not cosmetic details.
4	Captions must be accurate, complete, synchronized, readable, and properly placed.
5	For premium content, a bad localized version can be more damaging than delaying localization.

Shakable Beliefs

#	IVOC Summary
1	AI-generated audio must sound robotic or emotionally flat.
2	Human-grade localization always requires a large traditional studio budget and long lead time.
3	AI can only convey the gist and cannot preserve nuance, emotion, or voice identity.
4	Localization is only financially viable for the largest creators, studios, or publishers.
5	DIY software is automatically cheaper than a managed service once labor and rework are counted.

Other

#	IVOC Summary
1	The market is largely solution-aware and often product-aware, with skepticism shaped by first-hand AI-tool experience.
2	Primary segments include creators and podcasters, media and entertainment teams, enterprise L&D, and publishers/authors.
3	Common buying triggers include multilingual comments, global demand, an expanding back catalog, accessibility needs, and failed DIY output.

#	IVOC Summary
4	Key selection criteria are naturalness, accuracy, human review, turnaround, transparent cost, language coverage, and revision handling.
5	The strongest positioning territory is "market-ready global audio without managing the production mess."

Common Words/Phrases

sound natural	emotion and pacing	original tone	accurate translation	good quality
cost-effective	global audience	multiple languages	human review	cultural nuance
market ready	robotic	soulless	janky	random gibberish
fixing the output	time consuming	per finished hour	pronunciation	lip sync
speaker labels	sound effects	ready to publish	turnaround time	is it worth it?

Summary Rankings

Benefit-mapping rule

The left column is market language synthesized from the IVOC corpus. The right column uses only capabilities stated on the public ElevenLabs Productions page.

#	IVOC Summary	Associated Benefit
1	Take proven content global without building a localization operation.	Fully managed localization, captions, transcripts, dubs, and audiobooks; final assets delivered in the ElevenLabs account. [P01, P05, P06]
2	Human review is the trust bridge between impressive AI and market-ready output.	Productions is powered by vetted and trained linguists and localization professionals. [P08]
3	Naturalness means preserving intent, emotion, pacing, and a believable performance.	Human-edited translations and audio are positioned to make content look, sound, and feel natural. [P01, P02]
4	The hidden cost of DIY AI is correction time, retries, and unpredictable output.	An end-to-end managed service transfers production and QA work to an ElevenLabs Producer. [P05, P06]
5	Accuracy without cultural context is not localization.	Human-edited translation plus optional custom style-guide choices support local relevance and consistency. [P02, P07]
6	Speed matters only when the output is ready to use.	Localization starts from 24 hours; captions and transcripts can start from same-day turnaround. [P02, P03]
7	Buyers want cost predictability before committing.	Pricing is shown upfront in USD per source-audio minute before checkout. [P07]
8	Accessibility requires more than auto-generated dialogue text.	Professionally formatted and synchronized captions/transcripts, with optional closed captions and SDH. [P03]
9	Multiple speakers and long-form narrative raise the quality bar.	Productions covers complex multi-speaker discussions and single- or multicast audiobooks. [P03, P04]
10	Pronunciation, prosody, pauses, and post-processing determine whether an audiobook holds attention.	Those elements are explicitly handled by the audiobook production team; sound design is available as an add-on. [P04]
11	One production surface is more valuable than a chain of disconnected tools and vendors.	Orders, notifications, and final assets live on the ElevenLabs platform. [P01, P06]
12	Language coverage should be framed around quality and expansion, not a vague "global" claim.	Productions currently lists 11 supported languages and says coverage is expanding. [P02]
13	AI stigma drops when human craft is visible and specific.	The offer is framed as "AI audio, crafted by humans" and shows the producer network behind delivery. [P01, P08]
14	Premium buyers need distribution-ready output, not a draft they must rescue.	Enterprise Productions is described as end-to-end and ready for distribution. [P05]
15	The strongest emotional payoff is confidence: the localized version still feels like the original content.	Human-edited localization and expert audio engineering connect the technology to a lower-risk publishing outcome. [P01, P02, P08]

Message Implications

Lead with the job, not the technology

"Take your best content global without hiring and coordinating a localization team."

Name the enemy clearly

The enemy is not "slow translation." It is the pile of retries, reviews, vendors, file handoffs, and quality anxiety between upload and publish.

Make human review concrete

Show who reviews translation, pronunciation, pacing, timing, speaker attribution, and final audio. "Human-edited" needs operational detail.

Prove naturalness with paired examples

Use source and localized clips with comments from fluent speakers. Include difficult material: humor, mixed languages, names, technical terms, and multi-speaker overlap.

Treat DIY AI as the real alternative

The comparison is often Productions vs. a cheap tool plus dozens of hours of cleanup, not Productions vs. a traditional studio alone.

Reduce financial uncertainty

Surface example price ranges, turnaround bands, what is included, and what triggers extra cost or revision work.

Segment the promise

Creators want reach and velocity; L&D teams want consistency and update handling; publishers want performance and cost viability; accessibility teams want completeness and standards.

Avoid overclaiming lip sync or unlimited language coverage

The current Productions page does not clearly promise lip sync, revision rounds, or support beyond the 11 listed languages.

Questions the page still needs to answer

- How many revision rounds are included, and what counts as a correction versus a scope change?
- Can customers supply pronunciation guides, terminology lists, character notes, and brand voice rules?
- How are source-video updates handled without redoing the full project?
- What source lengths, speaker counts, file formats, and output formats are supported?
- Is lip synchronization part of any Productions package, or only audio timing alignment?
- What consent and rights checks apply to cloned/custom voices?
- What security, confidentiality, and data-retention terms apply to unreleased media and manuscripts?
- Are enterprise SLAs, dedicated project management, and volume discounts available?

Raw IVOC Data

Raw-data rule

Quotes are preserved as written aside from occasional shortening with ellipses. Spelling, capitalization, and grammar are intentionally left close to the source. Source IDs resolve to clickable links in the source index.

Pains and Fears

Raw IVOC quote	Segment	Source
"I think most english-speaking youtubers don't do it because its simply hard."	Creator	S01
"It ended up costing me a significant amount, which has impacted my overall profits."	Creator	S02
"It takes me a long time to do the dubbing (1 hour for a 13 minute video)."	Course creator	S04
"Still spending more time fixing the output than filming."	Creator	S08
"I considered localization, but YT auto dubbing doesn't bring results, freelancers seems to be expensive, HeyGen has limits on a paid subscription."	Established creator	S07
"Audiobooks are expensive in time and money to produce."	Author	S13
"Cost and access to readers are both huge hurdles."	Author	S14
"I feel like there is no quality control with audiobooks. It's a cash grab."	Audiobook listener	S19
"Sound effects are missed SO often!... It's so annoying."	Caption user	S20
"When the source video changes, do you end up redoing everything or have a better system in place?"	Localization professional	S24

Objections

Raw IVOC quote	Segment	Source
"You're not going to get "accurate" translations any other way than a human."	Creator/viewer	S02
"Results can range from decent to shit at random in my experience."	ElevenLabs user	S03
"The translation ISN'T market ready."	Bilingual ElevenLabs user	S10
"Tone and humor also come out super janky and often mismatch the original tone and intention."	Bilingual ElevenLabs user	S10
"Why would I watch something emotionless with robotic audio and intonation that isn't even translated well."	Viewer	S05
"I don't think I'd trust them to dub something accurately."	Viewer	S05

Raw IVOC quote	Segment	Source
"If I wanted something read by AI, I'd just copy paste that shit myself. Human or bust."	Audiobook listener	S15
"The German accent is missing."	Voice-cloning user	S11
"I would describe Eleven Labs results as "too perfect" or "too pristine"."	Voice-cloning user	S11
"If VCR drops for dubbed-language viewers, it could end up hurting your channel's overall performance."	Creator	S36

Desires

Raw IVOC quote	Segment	Source
"Ideally, the result should sound natural, with decent emotion and pacing, maybe even similar to my original tone."	Creator	S03
"I'm looking for more cost-effective alternatives that can still provide good quality."	Creator	S02
"Your videos in other languages will increase your: Number of subscribers, Reach, Impact, and Earnings."	Creator	S01
"I'm especially looking for tools that can retain the emotional tone, sync well with lip movement, and sound as natural as possible."	Filmmaker	S27
"The audio quality needs to be excellent... clear pronunciation but also has to sound natural."	Audiobook listener	S17
"AI narration is good for titles that otherwise wouldn't get audio."	Audiobook listener	S30
"Captions that do not include sound effects. Subtitles are not captions, stop trying to cut corners."	Caption user	S20
"I really want readable captions."	Caption user	S22
"I have a bunch of training videos that I want to make available in a bunch of other languages outside of English."	L&D creator	S26
"How did you do it? how much does it cost? Is it worth it?"	Creator	S01

Firm Beliefs

Raw IVOC quote	Segment	Source
"Translation and localization aren't one to one, you have to translate cultural meaning sometimes too between cultures."	Player/translator	S28
"True localization means adapting the message, tone, and references to suit each unique audience."	Localization practitioner	S37

Raw IVOC quote	Segment	Source
"Human narration is preferred."	Audiobook listener	S30
"Narrators are actors that bring a book to life."	Audiobook listener	S16
"The important thing is accuracy."	Caption user	S21
"Nuance is important."	Caption user	S22
"Using AI as support is one thing. Letting it flatten all the human judgment out of the video is what people seem tired of."	Creator	S09
"A finished hour is usually 3-4 hours of work on the back end in reading, recording, listening, re-recording... and editing."	Audiobook producer	S13

Shakable Beliefs

Raw IVOC quote	Segment	Source
"If it would sound perfectly and translation was perfect I would consider using it."	Viewer	S05
"AI voices are handy for on-the-spot, low-budget reads, but when a book needs true emotion, listeners will still reach for a human narrator."	Audiobook listener	S17
"It's better than not having an audio option for accessibility."	Audiobook listener	S17
"If you don't want to spend money, use AI... it's probably not going to be ideal, but it'll be good enough for people to understand."	Creator	S02
"I believe that there are services online that cater affordably for creators who aren't MrBeast."	Creator	S01
"For a one shot, there were some flubbed words, but it was actually pretty shocking seeing it get the translation, original voice, AND the emotion."	Dubbing tester	S31
"AI can't replace human taste or judgment, but it can remove friction from the parts that don't actually require creativity."	Creator	S39
"The use of AI reduces the production costs by 90%, making it more accessible to smaller publishers like my own."	Author	S40

Other

Raw IVOC quote	Segment	Source
"I run a YouTube channel where I create documentary-style videos on various topics."	Creator profile	S02
"I have some Udemy courses where the most common complaint is the audio quality."	Course creator trigger	S04
"I'm working on a multilingual film project."	Filmmaker profile	S27

Raw IVOC quote	Segment	Source
"I own a podcast studio, and I recently recorded an audiobook for someone. It's about 28 hours..."	Audio producer profile	S34
"I want to use AI voice to showcase various features... English is not my first language."	SaaS marketer profile	S33
"Any chance we'd get something like speaker detection labels in the transcripts?"	Podcast listener	S43
"Captions... with names misspelled repeatedly and nothing noting change of speaker."	Caption user	S44
"I'm thinking to dub my videos in English. But don't know where to find the perfect person for this work."	Creator trigger	S32

Source Index

Public IVOC sources

- S01 [Question] Dubbing videos into other languages to grow viewers - r/youtubers. [Open source](#)
- S02 How can I translate videos I make without an actual translator? - r/youtubers. [Open source](#)
- S03 Are there any good tools that can automatically dub a video from German to English? - r/VideoEditing. [Open source](#)
- S04 Video editor for dubbing based on subtitle timestamps - r/VideoEditing. [Open source](#)
- S05 Why do you guys hate Auto Dubbing? - r/youtube. [Open source](#)
- S06 The auto-dubbing/translating feature has reached an obnoxious point - r/youtube. [Open source](#)
- S07 Best YouTube MCNs for growth and localization? - r/PartneredYoutube. [Open source](#)
- S08 Been spending more time fixing AI-tool outputs for video content - r/ContentCreators. [Open source](#)
- S09 Anyone else noticing viewers are getting tired of AI content - r/ContentCreators. [Open source](#)
- S10 PSA: DO NOT trust the Dubbing Studio - r/ElevenLabs. [Open source](#)
- S11 Eleven Labs vs. Competition: Observations, Feedback - r/ElevenLabs. [Open source](#)
- S12 ElevenLabs - r/selfpublish. [Open source](#)
- S13 Producing audio books isn't worth it - r/selfpublish. [Open source](#)
- S14 Authors: question about audiobooks - r/selfpublish. [Open source](#)
- S15 Is this the ruin of audiobooks? - r/audiobooks. [Open source](#)
- S16 AI voices are so good that I don't think audiobook narrators will exist - r/audiobooks. [Open source](#)
- S17 Would you hate it if an audiobook's voice sounded a little robotic? - r/audiobooks. [Open source](#)
- S18 What's the first sign that tells you the audiobook is going to be bad? - r/audiobooks. [Open source](#)
- S19 Mispronounced words are driving me crazy - r/audiobooks. [Open source](#)
- S20 What is the most frustrating part of closed captioning? - r/deaf. [Open source](#)
- S21 What makes good closed captioning stand out from bad captioning? - r/deaf. [Open source](#)
- S22 Do deaf people prefer subtitles to include utterances and repeated words? - r/deaf. [Open source](#)
- S23 Opinions on AI sign-language tools - r/deaf. [Open source](#)
- S24 How are you handling video localization at scale these days? - r/localization. [Open source](#)
- S25 What determines using translated videos vs translated VTT/SRT? - r/elearning. [Open source](#)
- S26 Translate Training Videos - r/elearning. [Open source](#)

- S27** AI dubbing tools for converting one Indian language to another - r/Filmmakers. [Open source](#)
- S28** Kingdom Come Deliverance 2 dev says he was replaced with AI - r/Games. [Open source](#)
- S29** Yes, localization can be done badly, but that doesn't mean it shouldn't be done - r/CharacterRant. [Open source](#)
- S30** AI narration - r/audiobooks. [Open source](#)
- S31** ElevenLabs Dubbing v2 - r/singularity. [Open source](#)
- S32** I'm earning nothing at 200K subscribers - r/PartneredYoutube. [Open source](#)
- S33** Easy way to do AI voice over in video? - r/VideoEditing. [Open source](#)
- S34** What's the best service for transcription with timecodes? - r/podcasting. [Open source](#)
- S35** What does your editing process typically involve? - r/podcasting. [Open source](#)
- S36** Automatic dubbing - r/PartneredYoutube. [Open source](#)
- S37** How MrBeast scales his content internationally - r/PartneredYoutube. [Open source](#)
- S38** If AI helps you create, what still makes it yours? - r/ContentCreators. [Open source](#)
- S39** What if there is an AI tool that can help me grow? - r/ContentCreators. [Open source](#)
- S40** My publisher wants to turn my novel into an AI-narrated audiobook - r/audiobooks. [Open source](#)
- S41** I published my debut and became a 6-figure author - r/selfpublish. [Open source](#)
- S42** Can't afford Otter.AI anymore - r/Journalism. [Open source](#)
- S43** Version 2026.4 released: Transcripts - r/overcast. [Open source](#)
- S44** FCC complaints for ASR captions - r/deaf. [Open source](#)
- S45** I know that LTT wants to expand its reach, but this is ridiculous - r/LinusTechTips. [Open source](#)

Official product capability sources

- P01 ElevenLabs Productions overview.** Human-edited localization, audiobooks, custom voices, and audio engineering on the ElevenLabs platform. [Open source](#)
- P02 Localization details.** Dubbing and subtitling with human-edited translation and audio; 11 listed languages; turnaround starting from 24 hours. [Open source](#)
- P03 Captions and transcripts details.** Professionally formatted and synced captions/transcripts for complex multi-speaker content; same-day starting turnaround; CC and SDH options. [Open source](#)
- P04 Audiobook details.** Single and multicast; pronunciation, prosody, pauses, and post-processing handled by experts; sound design add-on. [Open source](#)
- P05 Enterprise service.** End-to-end managed production with content ready for distribution. [Open source](#)
- P06 Ordering workflow.** Orders placed in the account; an ElevenLabs Producer works on the content; email notification and final assets appear in the account. [Open source](#)
- P07 Pricing and delivery transparency.** USD per source-audio minute; exact price and turnaround shown before checkout; custom style-guide options affect price. [Open source](#)
- P08 Producer network.** Expert linguists and localization professionals vetted and trained by ElevenLabs. [Open source](#)

Official market and accessibility context

- A01 YouTube Help: Multi-language features.** YouTube says multi-language audio lets creators keep language tracks on one video and reports that creators using it saw more than 25% of watch time come from non-primary-language views. [Open source](#)
- A02 W3C WAI: Captions/Subtitles.** Captions include speech and non-speech audio information, speaker identification, and meaningful sound effects. [Open source](#)
- A03 FCC: Closed Captioning on Television.** Caption quality is organized around accuracy, synchronicity, completeness, and placement. [Open source](#)