

# VistatecSpeech



An Exciting New Multimedia AI Offering



## VISTATECSPEECH: AN EXCITING NEW MULTIMEDIA OFFERING

---

### Why VistatecSpeech is Better

- **Automates the transcription process** – produces an AI-generated script with speaker labeling and detailed timing information already inserted; output is highly accurate with minimal review & touch-up needed
- **Only a single localization pass required** - localizing the script and using it for both audio recording as well as subtitling improves the output consistency and the TM leverage
- **Less translation and rework** - translators see the entire sentence instead of fragments – improved context leads to better translations at the start, and less need to review or correct
- **TMS independent** - it works with any TMS that can import / export .JSON, .SRT, or .TXT files
- **MT independent** - it works with any neural MT engine that supports the desired language pairs
- **Makes content updates easy and painless** – adding a new chapter, updating a set of user instructions, etc. formerly would involve studio time and sound engineering time – higher costs and longer TAT. With VistatecSpeech, making updates is similar to editing a document.
- **Potentially huge cost savings** - overall, up to 80% savings on multimedia costs

### The Best of AI and Human Talent, Working Together for Your Business

The power of AI, guided by the wisdom and experience of multimedia professionals

- **Human oversight** - reviewed by humans at multiple steps in the production workflow
- **Highly customizable** - users can customize pronunciations, three-letter acronyms (TLAs), placenames, unique brand names, etc.
- **Variety of voice types for all quality needs and budgets** - clients can select any type of voice - human, synthetic, or synthetic voice-matching of human source voice
- **Automatically synchronized audio** - audio timings are automatically synched to the video for each target language, adjusting for the length of each translated sentence.

## Lower Costs and Greater Flexibility

- **Higher quality and faster TAT** – because VistatecSpeech has fewer touchpoints and less file-handling, this results in less opportunity for human error, which in turn reduces cost and removes schedule risk
- **Variety of voice types for all quality needs and budgets** - clients can select any type of voice - human, synthetic, or synthetic voice-matching of human source voice
- **Audio on demand for reviewers** - output audio can be generated at any time during the translation review process – no need to wait until the end of production to hear and evaluate the output – so as corrections are being made, the reviewer can check that the new output is correct
- **TMS independent** - works with any TMS that can import / export .JSON, .SRT, or .TXT files
- **MT independent** - works with any neural MT engine that supports the desired language pairs
- **Benefits scale up or down** - clients can use some, most, or all the features - with increasing benefits at each step
  - *For example, you can still use studio-recorded human voice with VistatecSpeech, if you prefer*
  - *You will still get the benefit of AI-powered transcription, translation, timing, and speaker tagging as well as improved consistency between translations of audio script & subtitles*



## HOW TO DECIDE

---

### If You Need:

- Video material with topics of an informational, general purpose, or lightly specialized nature
- Scenarios where the general idea needs to be clear, and precise details are "nice to have" but not critical
- Frequently changing content that needs constant curation and updates
- Fast-changing content with a short shelf-life
- Material that is instructional, "how-to", informational, or internally facing to the client
- Step-by-step process breakdowns, storytelling, discussions
- Suitable for Marketing, HR training, user training, product demos, interface walk-throughs
- Processing large volumes of legacy multimedia content

*Then  
VistatecSpeech  
is for you!*

### If You Need:

- Video material on a highly complex or technical subject
- Scenarios where every detail must reflect the local experience
- Highly visible content - brand-sensitive, financial, or regulatory in nature
- Discussion-based content with multiple speakers or a lot of talking heads
- Content with emotional scenes
- Audio with specific character traits, i.e., accents, distinctive intonations
- Lip syncing (precise matching of mouth movements to audio dialog) is required
- Videos with lots of on-screen text (excluding subtitles)

Then traditional  
voiceover services  
may be your better  
option.

## AVAILABLE LANGUAGES AND PRICING MODEL

---

### Languages

**Now:**

English	Italian	Spanish – Spain	Chinese - Simplified
German	Dutch	Spanish – Latin America	Arabic - MSA
French - France	Portuguese – Iberian	Japanese	
French - Canada	Portuguese - Brazilian	Korean	

**Soon:**

*Polish, Russian, Swedish, Turkish, Hindi – and more*

*New languages being added constantly – ask your Vistatec Account Manager for details*

### Pricing Model

Included	Not Included
<ul style="list-style-type: none"> <li>• AI-enabled automated transcription</li> <li>• Human QA of automated transcription</li> <li>• Synthetic and/or cloned voiceover</li> <li>• All video production steps, start to finish</li> <li>• Final human QA performed over entire deliverable before performing the hand back</li> </ul>	<ul style="list-style-type: none"> <li>• Actual translation – clients still need to select the type of translation required for their project and the quality level: raw MT, MTPE, full human translation, etc.</li> <li>• On-screen text – this still requires special handling and needs to be priced separately</li> </ul>