



Guidelines for a quality automatic audio/video transcript

Automatic transcription of audio and video files presents the advantage of high working speed and low costs compared to manual processing. If operator-assisted transcription requires three to five hours of work for each hour of recording, an automatic transcription engine, such as Scriptoman, achieves this in half the recording time. In both scenarios, several elements are needed for a quality final transcript, such as: accurate identification of the recording's particulars (if any), enhancing the audio quality to the highest possible level, sound post-processing, etc.

Recording Analysis and Particulars Determination



The recording analysis aims to identify specialty terms or professional jargon. Many industries (IT, healthcare, legal) speak their own specific language. Scriptoman allows you to enter a custom dictionary (industry-specific terms, acronyms, company names, etc.), and these will then become much easier to identify in the transcription process. . In the absence of such dictionary, transcription accuracy is impaired, but there is still the possibility that these terms are corrected in the post-processing editing.

In this stage, the audio quality of the recording can be analyzed also to help users set realistic expectations as to the transcription results. Automatic transcription services (voice to text or speech to text) are highly dependent on the quality of the audio recordings they process. It is possible to reach an accuracy of above 90%, but this only when an audio or video file features a clear, well-articulated, accent-free, loud-enough voice, and is free of background noise, overlaps or echoes. When a recording is impaired by any of this factors, the transcription accuracy drops, and the time required to process the file increases.

Recording Preparation and Organization



Recording preparation and organization are critical steps, particularly when these are not beyond your control. As much as possible, inform participants that the discussion is recorded and ask them to speak slowly and plainly. For conferences, use individual microphones for each speaker, turn them on only when these are speaking, and try to prevent any voices overlapping. If you are recording a phone call, speak as slowly and clearly as possible, and ask your collocutor to do the same. The effects on the final results can be surprising.

Quality Microphones

The recording is key to the entire transcription process, and because of this we recommend that you pay utmost attention to this aspect. To get a good audio recording, you need to use quality microphones. There are huge differences between the standard microphones built in phone sets, tape recorders, cameras and multimedia headphones, and the



microphones made precisely for conferences, podcasts, radio, interviews, stage, etc. The first category can only be used in small, closed and quiet spaces, and even so, the end result can be often rather disappointing. For any other conditions: large or outdoor spaces, ambient noise, amplification systems, etc. you will definitively need state-of-the-art external wired or wireless microphones. As a general comment, microphones fall into three categories: unidirectional, bidirectional, or omnidirectional.

- Unidirectional microphones (cardioid, for those who speak Technical) have a maximum sensitivity to the front of capsule, which decreases slightly on the sides, and becomes very poor to the back. They are most often used and are recommended to capture stand-alone sources, or when a good separation of a person's voice from background noise is required, including reflection of the original sound from walls or other objects. They are best fit to capture voice when recording is made in large rooms with good acoustics.
- Bidirectional microphones are somewhat rarer and used mainly by professionals, whereas omnidirectional microphones are mostly those built in different pieces of equipment. This latter category is suitable for a contextual operation, but are able to provide for a good recording only in ideal conditions: small rooms with free of any background noise, and at a very short distance from the speaker.
- In order to obtain a quality audio recording, use dedicated microphones: static office microphones for podcasts, video conferencing, video-blogs, etc.; clip on microphones for interviews and MCs on stage (particularly if they move around a lot); unidirectional microphones for large spaces with background noise; protections (sponges, windshield) for the mic capsule when you record outdoors.

Removal of Background Noise



Quality microphones help mitigate background noise, but ideally this should be removed at source. Have your recordings done in the quietest rooms, and see that no door are opened/closed in the process, colleagues are chatting down the hall, or phones are ringing.

Also, turn off other electrical/electronic equipment in the room: air conditioning systems, fans, computer or servers, refrigerators, etc. All this background noise will affect the final sound quality, and removing them with editing systems could prove rather complicated.

Audio Editing

CONFIDENCE



LEVEL	STATISTICS
 Very confident	89.63% 3007 words
 Fairly confident	8.70% 292 words
 Slightly confident	1.67% 56 words

Editing just got easier—see each word's confidence level by clicking on the thermometer icon 🌡️.

To increase the transcript quality, use the Scriptoman editor!

It shows you the the degree of words that have been transcribed with a high level of confidence. It also marks any inaccurate words.

Thus, you will end up spending much less time editing, and have transcript that can be even 100% accurate.

Practical Tips from Professionals:

- ✓ Record indoors as much as possible, having first attempted to identify the quietest room. If this is not possible and you have to record outdoor, use sponges or windshields on your mic to mitigate the background noise. If you can, remove any background noise with an audio editing software.
- ✓ Set up and fit out a recording-only space. It doesn't have to be a studio *per se*. A small room with walls covered (in curtains, drapes, paintings) shall do the job when it comes to recording because it has a very low echo.
- ✓ Invest in a quality microphone. The difference between a built-in microphone and a dedicated one is huge, even when the latter is just an entry-level model. Most built-in microphones are omnidirectional, which means they will record almost every unwanted noise.
- ✓ Make use, as much as possible, of unidirectional microphones. In most cases, unidirectional microphones provide the best recordings. These are highly sensitive, and when directed in the right direction, the captured background noise is significantly reduced. Also, use microphones clipped onto the speaker's lapel as often as possible because they will record a clear and strong voice, even when a person is moving around.
- ✓ Speak close to the microphone. The shorter the distance between the speaker and the microphone, the better the recording quality. Sound fidelity is enhanced, whereas the background noise is reduced. When the speakers are advised to articulate well their sounds, high quality is guaranteed.
- ✓ Permanently check the sound. Do not wait for the recording to end to find out of quality it is. Always check the microphone level and settings, and whether any background noise is present or not. The ventilation system in a room may go unnoticed during recording, it will be definitely heard as background noise on the final audio file. Always check with a pair of headphones.

Conclusions

A quality audio recording is critical for an equally good transcription. Background noise, overlapping voices, echo, speaking with an accent, etc., affect the audio recording and, consequently, the quality of the transcript produced. Removing these elements will help achieve better results, in a shorter time, and reduce the load of final editing work required.