



ALPHA



White Paper

The ultimate guide to voice-over localization (including dubbing, lip sync and more)

If you want global audiences to connect with your video, film or multimedia content (such as video games), using localized voice-over to replace original spoken-word audio is proven to increase audience engagement and boost viewer retention.

Certainly Netflix thinks so – you'll find out why at the end of this article.

ALPHA LTD
St Andrew's House
St Andrew's Road
Cambridge CB4 1DL
United Kingdom

 @thisisalphalive

alphacrc.com

In fact, the huge growth in streaming platforms delivering content across multiple global regions means the market for lip-sync dubbing (more on exactly what this is later) is set to expand tenfold in the next five years.

But what are the different types of [voice-over localization](#) on offer? And which will best suit your audience (and your budget)?

To provide some answers, we've created the Ultimate Guide to Voice-over Localization, which details the key differences between specialized services used instead of subtitling, including:

- Voice-over
- Dubbing
- Lip-sync dubbing
- Phrase-sync dubbing
- Non-synced dubbing
- Bilingual delay dubbing
- Voice-over narration
- Wild track recording

What is voice-over?

More generally, in film, television and video, "voice-over" refers to narration by a person off-screen who is providing some kind of commentary for on-screen events (a film narrator, for example, or a journalist's script read over news footage).

But in the translation and localization industry, we refer to voice-over as any type of translated or localized spoken-word audio, which either replaces or overlays an original spoken word audio track for film or video content.

This being said, voice-over is the generic term for all types of specialist spoken-word audio translation/localization.

The different sub-categories of localization voice-over – such as dubbing, lip-sync dubbing and wild track recording – are described below.

What is dubbing?

[Wikipedia](#) defines dubbing as "the substitution of the voices of actors shown on the screen by those of different performers". This is also known as "revoicing" in the film industry.

Beyond film, dubbing can also be used to describe voice substitution of any people on screen who are not actors, such as presenters, interviewees, animated characters, or computer-generated characters in video games.

There are, however, different kinds of dubbing which require varied levels of expertise, resources and budget. These include:

- Lip-sync dubbing
- Phrase-sync dubbing
- Non-synced dubbing
- Bilingual delay dubbing

Let's find out more about each of these different types.

What is lip-sync dubbing?

At the top of the voice-over food chain is lip-sync dubbing (also referred to just as "lip sync"), the most complex and demanding of all the voice-over formats.

Lip sync is when the voice in the new language replaces and closely matches the lip movements in the spoken vocals of the original language. The aim is that the synchronization of new spoken word with the on-screen speaker's lip movements is so closely matched that the audience barely notices that the original audio has been replaced.

Of course, this is no easy task. Languages use different amounts of words to convey the same meaning; for example, Spanish dialogue is often around 25% longer than its English equivalent. It's a complicated process to create new target-language dialogue that has the same meaning, tone of voice and can sync with the original speaker's lip movements.

Lip-sync dubbing involves specialist translators who are trained to work within these tight constrictions. For lip-sync dubbing of big-budget movies, native-speaking script writers and editors often also get involved in the process to ensure the dialogue is pitch-perfect for its target audience.

Then of course you have the recording process itself. Specialist lip-sync voice-over professionals are experts at matching dialogue to lip movements; then, an editing team will aim to create the perfect synchronization of new spoken word to original visual content.

All these factors mean that lip sync is certainly not a budget option. That's why it's mostly used for high-end film and television dramas, where the exact synchronization of new dialogue with on-screen action increases audience immersion and emotional engagement.

For an example of lip-sync dubbing in action, [here's a video](#) of Robert de Niro selling cars in *Analyze That* dubbed into German. (Find out about the foreign-language voice-over actors who revoice the Hollywood stars [here](#)).

However, there are also other voice-over choices which are quicker, cheaper and equally effective for many types of multimedia content.

What is phrase-sync dubbing?

In contrast to lip-sync dubbing, phrase-sync dubbing (also referred to as “phrase sync”) does not attempt to match the new spoken-word audio to the exact lip movements of the on-screen speakers. Instead, spoken-word dialogue is matched phrase by phrase rather than word by word. This means that when a person stops and starts speaking in the original, the localized phrase-synced version has to match the same timings.

Because the level of synchronization is much less precise, phrase-sync dubbing is on average six times faster than lip-sync dubbing to complete.

Inevitably, that makes it a much cheaper option which is often used when a lip-sync budget is not available and/or when it’s not necessary to “believe” that the speakers are actually speaking the new language, but when it’s still valuable for audiences to engage with the new dialogue in a close approximation of the original spoken word.

What is non-synced dubbing?

In contrast to lip-sync dubbing and phrase-sync dubbing, with non-synced dubbing there is no requirement to match new localized spoken word audio with any of the on-screen speakers’ movements.

Translated dialogue should always be accurate and faithful to the original, but the need for close or approximate synchronization is removed, making this a cheaper and quicker voice-over localization option.

This type of dubbing is therefore used more commonly in factual content (such as documentaries or training/educational programmes, for example) where it is less important that the viewer believes the speaker is actually delivering lines in the new target language.

What is bilingual delay dubbing?

Bilingual delay dubbing is a specialized form of dubbing where the spoken word content starts with the on-screen speaker delivering the words in the original language, which is then overlaid with a translated version within a few seconds.

The original-language audio volume is lowered enough to allow the new-language version to be heard distinctly, but it is still faintly audible in the background.

This approach is often used in translating speech in news reports. As the beginning of the original-language being spoken is clearly audible (and the rest could potentially be carefully retrieved if required), it confers a certain amount of authenticity on the dubbed track follows. This is because speakers of the original language are able to confirm that the translation is accurate.

What is localized voice-over narration?

In contrast to dubbing, localized voice-over narration is used to replace spoken-word content which is not delivered by an on-screen speaker.

All types of content use this approach, including films, television, video games, corporate videos, marketing campaigns and more.

The key elements are delivering an accurate translation which engages with its target audience, using a voice-over artist who can deliver the script in the appropriate tone of voice, and ensure synchronization with on-screen content.

Of course, because there are no on-screen speakers to follow, synchronization timings are more flexible. This usually makes localized voice-over narration a quicker and less expensive option than dubbing.

What is wild track recording?

At the other end of the precision spectrum to lip-sync recording is wild track recording.

In wild recording, the localized spoken word content does not have to be a word-for-word translation of the original, but rather provide an accurate gist or summary of what is being said.

Likewise, there is no attempt at precise synchronization, as long as the new wild track basically matches the on-screen content.

There are many types of content when precise translation and synchronization of localized spoken word are superfluous to audience needs. Getting the main ideas in the right order is what counts, and wild track audio offers an affordable way to achieve this.

Why use voice-over? (as explained by Netflix)

Without the need to record audio, hire voice-over artists and (sometimes) sync audio precisely to the on-screen action, subtitles are nearly always a cheaper and quicker option than using voice-over localization.

However, for many types of content, voice-over localization has clear advantages over use of subtitles for translated spoken word.

Certainly Netflix thinks so. In an [article in Variety](#), the streaming giant revealed that it automatically sets its original foreign-language programmes, such as German drama *Dark* or Brazilian sci-fi show *3%*, to versions dubbed in English for US viewers, rather than offering subtitles. (These are, however, available as an option.)

Why? The answer is an interesting one.

When Netflix asked US viewers if they would prefer to watch foreign-language shows with dubbing or subtitles, there was an overwhelming preference for the latter option. But when it tested this theory and put the French show *Marseille* by default in a dubbed version for one US audience group and with subtitles to another, the results didn't match what viewers had previously said.

In fact, the group watching the dubbed version were more likely to complete watching the series than those viewing with subtitles. So Netflix made a simple business choice: its original foreign-language shows are set by default with dubbing for US viewers, who can choose a subtitled version if they prefer.

Ease of understanding, total viewer immersion and an enhanced "native brand experience" are some of the reasons why film studios, television companies, games studios and content providers prefer to use some form of rerecorded audio to subtitles when taking their products to new markets.

In conclusion

We hope this article has clarified what is a potentially confusing array of services and terminology used in the growing industry of voice-over localization.

Knowing your lip-sync from your wild track recording will help you determine what the best options are for your content, your audiences and your budget.

Of course, if you have any further questions or particular projects you'd like to discuss, we'd love to talk with you. Please send us an email at marketing@alphacrc.com.

Found that interesting?

If you're not already signed up to our newsletter, register on our [News](#) page to receive lots more great articles that will help you on your journey to going global.