

Guidelines for English transcripts

These guidelines describe how to perform the revision of (English) automatic transcripts from a video presentation in English. Each transcript consists of a scientific presentation of a research paper, often delivered by non-native English speakers. Each video is approximately 6 minutes long and features one speaker discussing technical and scientific topics. For this reason, and to support you in the activity, we will provide you with the associated scientific paper, too.

Please find below more detailed information on how to perform the revisions by ensuring the accuracy and consistency of the final transcripts.

If you have any questions about the task, encounter unclear audio, or need guidance, please don't hesitate to reach out and tag us for assistance.

Thank you for your hard work and commitment to this project!

Accuracy and Clarity

Given the scientific nature of the videos' content, it is extremely important to ensure that the content of the **transcript adheres to the original audio as much as possible**, capturing the speaker's exact words. Thus, you should avoid adding interpretations or assumptions, or further clarifications: the transcript should only contain what is actually said.

See some special cases that might require particular attention during the activity.

- **Technical Terms and Proper Nouns:** Verify the spelling of technical terms, names, and scientific jargon where possible. Use online references or project glossaries as needed. Also, you can check in the original associated paper that is being presented.
- **Contractions:** Use contractions (e.g., "don't," "can't") if the speaker uses them, and vice versa if the speaker explicitly uses the full form (e.g., "do not").
- **Mispronunciations:** Ensure the transcript contains correct English words, even if the speaker might have an accent or might mispronounce them—this is possible given that you are dealing with non-native speakers of English.

! We admit some exceptions to the above-mentioned rules of perfect alignment to the original audio for the sake of simplification, as discussed below.

Exclusions and Simplifications

- **Disfluencies:**

- Do not transcribe **fillers** like "um," "uh," unless they are contextually significant (e.g., "right?" for emphasis).
- Do not transcribe phrase and word **repetitions** (including false starts). For example:
 - Audio: "*I-I mean, the data shows...*"
 - Transcript: "*I mean, the data shows...*"
- Do not transcribe **self-corrections** or **revisions**. For example:
 - Audio: But that's it's not, it's not, it's, uh, it's a word play on what you just said.
 - Transcript: But it's a word play on what you just said.
- **Background Noise:** Do not transcribe background sounds (e.g., "[dog barking]").

! If a portion of the audio is unclear or unintelligible, cancel the automatic transcript and write '[INAUDIBLE]' including a timestamp indicating the time where the unintelligible portion occurs—e.g. [INAUDIBLE 1:12]--and reach out to us.

Spelling, Style and Formatting

- **Spelling:** Use US English variant
- **Punctuation:** Use punctuation to reflect natural pauses or intonation. For example:
 - *Periods* for full stops.
 - *Commas* for short pauses within sentences.
- **Capitalization:** Follow standard capitalization rules for proper nouns, acronyms, and sentence beginnings.
- **Non-English Words:** It's possible you may encounter non-English audio portions (e.g. "As you can see in the example shown, in Italian we find a morphological distinction between *ragazzo* and *ragazza*").
 - If you are able to transcribe the language, please do an in-text annotation as follows: [FOR:LANG text]
 - "As you can see in the example shown, in Italian we find a morphological distinction between [FOR:IT ragazzo] and [FOR:IT ragazza]"
 - If you are unable to transcribe the language, cancel the automatic transcript and write [FOR:UNK timestamp] (where "timestamp" is the time when the foreign word occur)
 - "As you can see in the example shown, in Italian we find a morphological distinction between [FOR:UNK 3:04] and [FOR:UNK 3:11]"