Captioning Style Guide v6
Welcome to the Captions Style Guide!

This guide explains Rev’s expectations for captions quality. In addition to this guide, which covers the fundamental elements of a quality captions project, we have a robust Help Center for additional guidance.

Captions are used as part of accessibility services. As such, we need to ensure that a deaf or hard-of-hearing viewer receives a similar experience to that of a hearing viewer.

Browser Compatibility

Rev recommends that you use the most up-to-date version of Google Chrome when working with the Dash editor.
Capturing Content

Spoken Content
Speaker Identification
Atmospherics
Lyrics and Music
Atmospheric-Only Projects
Foreign Language
Difficult or Challenging Content
Glossary Terms & Resource Files
Scripts

Formatting Captions

Caption Length
Caption Grouping

Syncing Captions

Caption Timing

Help Center Articles

These are additional resources. They are also linked throughout the guide.

Accuracy
Lightly Editing
Commas & Terminal Punctuation
Quotation Marks
Hyphens and Ellipses
Speaker Labeling
Audio-Only Projects
Atmospherics
Lyrics
Atmospheric-Only Projects
Foreign Language
Unclear Words
Scripts

Help Center Articles (cont...)

Glossary Terms and Resource Files
Caption Grouping
Music Atmospherics
Unworkable Projects

Grading

Grading Scale
Rubric

Additional Resources

Captions Article Directory (list of all Captions Help Center articles)
Dash User Guide
Help Center
Forum
**Rule** Caption **all** spoken English words, even if there are pre-existing captions or subtitles on the video. Lightly edit when necessary for readability. Use US spelling.

**Why** To provide a deaf or hard-of-hearing viewer the same experience of watching a video as anyone else.

**Rules of thumb for caption accuracy:**

- Maintain the integrity of the spoken words.
  - Do not paraphrase, rearrange, or change the speaker's words.
  - Caption contractions, formal and informal, as spoken.

- Most projects will have an auto-generated speech recognition draft. Do not overly rely on the draft.

- Lightly edit unscripted productions, but do not omit intentionally spoken words.

- You are expected to research proper nouns and terminology for representation and proper spelling.
  - Watching for terms on screen can be helpful.
  - Googling with a bit of context from your video/audio is also helpful.
  - URLs, hashtags, social media tags should be captioned using common convention: www.rev.com / #revcaptions / @rev

- Include proper punctuation and capitalization per common English grammar rules.

- Never type out a censored, silenced, obscured, or beeped/bleeped out word.
  - Use an appropriate atmospheric for the sound heard when the word is censored, e.g. (beep)
When the speaker **CAN** be visually identified:

Use a dash and space at the beginning of the speaker’s dialogue. This includes the first speaker.

**WHY?** So that a deaf or hard-of-hearing viewer will know someone different has started speaking.

---

When the speaker **CANNOT** be visually identified:

Use a dash, space, and bracketed identifier, or speaker ID, at the beginning of the speaker’s dialogue.

**WHY?** So a deaf or hard-of-hearing viewer will know who is speaking.

---

**EXCEPTION** For audio-only projects, only a dash and space is needed for speaker changes.
Captions need to indicate sounds heard on screen. We call these identifiers atmospherics.

Atmospherics provide visual indicators of non-verbal sounds to the viewer. This allows the deaf or hard-of-hearing audience to pick up on sounds that are important to the content of the video.

How to create atmospherics:

<table>
<thead>
<tr>
<th>Do</th>
<th>Don't</th>
</tr>
</thead>
<tbody>
<tr>
<td>● Use parentheses ( ) and lowercase unless a proper noun is used</td>
<td>● Don’t use a dash or speaker label in a caption group containing only atmospherics</td>
</tr>
<tr>
<td>● Describe the sound or sounds heard on screen by following this convention:</td>
<td>● Don’t use onomatopoeia e.g. (ribbit ribbit); instead, describe what’s creating the sound, e.g. (frog croaking)</td>
</tr>
<tr>
<td>○ noun + descriptor/verb in present tense form e.g. (water boiling), (door slams)</td>
<td></td>
</tr>
<tr>
<td>○ The noun lets viewers know who or what is making the sound, while the descriptor/verb lets them know what the sound is</td>
<td></td>
</tr>
<tr>
<td>● Always use present tense, e.g. (Erin coughs)</td>
<td></td>
</tr>
</tbody>
</table>

For extended periods of music, refer to the next slide: Music
Capturing Content

Lyrics and Music

**RULE**
Caption music and lyrics when there is no spoken dialogue occurring at the same time, even if there are pre-existing captions or subtitles on the video.

**WHY**
When there are no spoken words, the lyrics become the dialogue to be captioned. When lyrics are heard, it is important that they be captioned on-screen for the deaf or hard-of-hearing audience to experience.

**Lyrics**
- When there is no spoken dialogue occurring at the same time, lyrics should be captioned, even if there are pre-existing captions or subtitles on the video.
- Omitting clearly heard lyrics will result in a reduction in score and can be scored as low as 1/1/1.

**Music Atmospherics**
- When a project also contains spoken words, only include a background music atmospheric if there's a significant time gap and it would benefit the viewer to include.
- A common format is a descriptor followed by the word “music.” You can indicate the progression of music with words like *begins* and *continues*. E.g. (orchestral music begins)
- Introductory music is a common use case. E.g. (bells chiming)

**How to notate lyrics:**
Include a musical note at the beginning of the caption group. In Dash, use `##` followed by a space to create the musical note. Dash will add a second music note at the end of the caption group after the project is submitted.

*Tip:* Googling portions of the lyrics can be helpful.
Atmospherics-Only Projects

**RULE**

For **atmospherics-only projects**, accurately caption all sounds that are heard using atmospherics.

**WHY**

Even when there is no spoken audio in the project, atmospherics tell the story for the deaf or hard-of-hearing viewer.

---

**What is an atmospherics-only project?**

- No spoken dialogue
- Contains sounds or music
- Project with only cartoon gibberish

**What is NOT an atmospherics-only project?**

- All spoken dialogue and/or lyrics in a foreign language
- Audio is corrupt OR entirely silent
- Project has no spoken words OR meaningful sounds

*Projects should be unclaimed as “no audio present” or “no English audio/ no atmospherics”*

---

**How to caption an atmospherics-only project:**

- Include atmospherics more frequently than you would in a normal project.
- Caption all environmental sounds, action sounds, character noises or gibberish using an atmospheric.
- Use detailed atmospherics to capture music.
  - Does the instrument convey a tone?
  - Does the volume or tempo increase or drop off?

Atmospherics are inherently subjective, and we understand that this isn’t always a well-defined area. Please do your best to caption these projects in a way that creates valuable content for hard-of-hearing viewers.
For foreign language (non-English) in a project, follow the guidelines below.

**WHY**
Some customers place orders with foreign languages spoken within the content. We want to be sure we are delivering a product the customer expects, so we've put together guidelines on how to handle foreign language. Noting when non-English dialogue is being spoken lets a deaf or hard-of-hearing viewer know that a conversation is taking place.

**RULE**

**How much of the project is in a foreign language?**

- **100% foreign language (unworkable)**
  Unclaim the project and select “Contains only foreign language” as the reason for unclaiming.
  - Projects where no English is spoken/sung
  - Projects where sounds are present, but no English is spoken or sung

- **Partially English, partially foreign language**
  Caption the English as usual and use an atmospheric to indicate when English is not being spoken.
  **Do not caption non-English content.**

- **There are only a few words in a foreign language**
  Caption the word(s) if:
  - the word is part of a proper noun.
  - the word is common in the English language.
  - the word was supplied by the customer in the glossary or provided materials.

Learn more about: Foreign Language
**CAPTURING CONTENT**

**Difficult or Challenging Content**

**RULE**
You should do your best to caption all spoken words and/or lyrics. For extremely challenging content, follow the guidelines below.

**WHY**
When dealing with challenging content, consider the best interest of the customer and the final product that will be seen by the deaf or hard-of-hearing viewer.

If your project is entirely too challenging to accurately caption the spoken words:
- Unclaim the project and select “difficult audio” as the reason for unclaiming.

If your project contains **both** challenging content, and clear content:
- Accurately caption the spoken words that can be clearly heard.
- For certain challenging sections of audio where it’s not crucial to understand every word, use an atmospheric to provide context to the viewer.
  - e.g. (friends conversing quietly), (group chattering)
- If an occasional word cannot be understood, use (indistinct) in place of the word.
  - (indistinct) should only be used if you **absolutely** cannot determine the word; excess or inappropriate tags could result in a lower grade.

**Tip:** You can use **Lend an Ear** and see if the community of Revvers are able to help.
Customers may provide additional information in the form of glossary terms. Customers may choose to upload a resource file with glossary information.

This additional information is provided to help create accurate captions.

The glossary is a section where customers can provide clarification on the spelling of words and terms that may be presented in the video. These can include but are not limited to niche terminology, company names, and speaker names.

**Important things to note about glossary terms and resource files:**

- Always check projects for included glossary terms. Misspelling a term found in the glossary will result in a lower score.
- If a customer provides instructions in a glossary or resource file that go against the guidelines found in this guide, they should be ignored. Caption the project as usual.
  - These types of instructions are not approved by Rev.
  - Following instructions that go against our guidelines can result in scores as low as 1/1/1.

**Remember:** Glossaries and resource files should only clarify spelling. Any additional instructions are not approved and should not be followed.
Scripts should be used as resources to help create accurate captions. Regular Rev guidelines still apply to projects submitted with scripts.

When a customer provides a script, their expectation is that we will use their document to create the captions for their video.

**General guidelines when following scripts:**

- Follow scripts as closely as possible to ensure that all dialogue is captioned correctly. This will ensure that names and important terms are spelled correctly.
  - For some projects, it may be possible to use the scripts exactly as provided by copying and pasting the content into Dash.
  - If copy and pasting isn’t an option, the script should be used as a resource to ensure accuracy.
- Some scripts may not be complete and we may need to add in captions for spoken dialogue and/or atmospherics that have not been included.
- Scripts can provide valuable information about characters and scenes. This may be included as notes, separate from the dialogue.

**Remember:** A script should only be provided to assist with captioning dialogue, names, terms, and sounds. Any additional instructions that do not follow Rev guidelines are not approved and should not be followed.
Formatting Captions

**Rule**
Individual caption groups must always contain 60 or fewer characters.

**Why**
Captions that fall within certain character limits are easier to read quickly. This ensures that viewers do not miss spoken content.

The typing area turns **green** and then **yellow** as you add more characters and near the 60 character limit.

**It's perfectly acceptable to submit captions with a white, green or yellow caption group.**

The typing box turns **red** when you are over the 60 character limit. Split the text across multiple caption groups until the red color disappears.

Learn more about:
[Advanced Caption Formatting](#)

---

A caption will start white | 00:09.44

This is green due to the number of characters | 00:10.28

The group will turn yellow as you add more characters | 00:11.94

If you type a caption group that exceeds 60 characters, the caption box turns red to let you know to break the caption group | 00:00.00

See next page for more information on caption grouping guidelines
Captions should be split into individual caption groups such that whole phrases, nouns, sentences, and flow of dialogue is interrupted as **minimally** as possible, while abiding by the 60 character limit.

This allows the captions to be easily read by the viewer, with logical breaks and enough time on screen.

<table>
<thead>
<tr>
<th>Split captions</th>
<th>Do</th>
<th>Don’t</th>
</tr>
</thead>
<tbody>
<tr>
<td>After terminal punctuation</td>
<td><em>I need to call her.</em></td>
<td><em>I need to call her. I was in my office this morning</em></td>
</tr>
<tr>
<td></td>
<td><em>I was in my office this morning</em></td>
<td><em>and got the message that she called.</em></td>
</tr>
<tr>
<td></td>
<td><em>and got the message that she called.</em></td>
<td></td>
</tr>
<tr>
<td>Before conjunctions</td>
<td><em>They are just a handful,</em></td>
<td><em>They are just a handful, but they all have</em></td>
</tr>
<tr>
<td></td>
<td><em>but they all have their pros and cons.</em></td>
<td><em>their pros and cons.</em></td>
</tr>
<tr>
<td>Before prepositions</td>
<td><em>The flock of geese flew gracefully over the large</em></td>
<td><em>The flock of geese flew gracefully over the large</em></td>
</tr>
<tr>
<td></td>
<td><em>oak trees that were nearest to the</em></td>
<td><em>oak trees that were nearest to the</em></td>
</tr>
<tr>
<td></td>
<td><em>babbling brook.</em></td>
<td><em>babbling brook.</em></td>
</tr>
</tbody>
</table>
Sync each caption group (atmospherics and speech) so it appears on-screen when the audio begins.

The deaf or hard-of-hearing viewer should see the text on screen at the same time it would have been heard by anyone else.

1. **The start time needs to align with the beginning of the sound.**
   - Aim for precision, but it’s ok for the start time to be up to a ½ second early or late from the true beginning of the sound.

2. **Do not worry about the end time.**
   - Rev automatically calculates the end time for a caption group after you submit the project (post-processing).
   - **NEVER** add extra spaces to a caption group OR double up the captions in an attempt to adjust the amount of time the caption group is on-screen. This causes errors in the file format for customers.

3. **Keep in mind the readability of the caption groups.**
   - **Split the caption groups** if a speaker is talking very slowly (more than 5 seconds to say a sentence) or there is a long pause. This maintains proper timing with the speech and ensures the caption group does not end too early.
   - **Use advanced caption format** if multiple speakers are talking very quickly. This combines two quick phrases into the same caption group for readability.

**Tip:** For more details on how to sync captions, see the [Dash User Guide](#).
A grade consists of scores in three categories: **Accuracy**, **Formatting**, and **Syncing**. You can view the **scoring rubric** on the next page to see how graders will assess your work.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Status</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 - Excellent</td>
<td>Customer ready</td>
<td>Captions are high quality, having few if any errors.</td>
</tr>
<tr>
<td>4 - Good</td>
<td>Customer ready</td>
<td>Captions are acceptable quality, and while there are some noticeable errors present, the project is still customer ready overall.</td>
</tr>
<tr>
<td>3 - Needs Improvement</td>
<td>Not customer ready</td>
<td>A reasonable effort was made, but multiple edits would be needed before this is considered customer ready.</td>
</tr>
<tr>
<td>2 - Poor</td>
<td>Not customer ready</td>
<td>Some effort was made, but there are significant errors that would be time-consuming to correct before this is considered customer ready.</td>
</tr>
<tr>
<td>1 - Very Poor</td>
<td>Unusable</td>
<td>No reasonable or consistent effort was made.*</td>
</tr>
</tbody>
</table>

*Keep in mind that both the **severity** and **frequency** of errors is taken into consideration when assigning a score.*

* If you submit **incomplete or unedited projects**, your pay for the project will be removed, the project will be graded 1/1/1, and your account may be closed without warning.
<table>
<thead>
<tr>
<th>Accuracy</th>
<th>5 - Excellent</th>
<th>4 - Good</th>
<th>3 - Needs Improvement</th>
<th>2 - Poor</th>
<th>1 - Usable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Does the text accurately reflect the spoken audio?</td>
<td>Spoken audio is accurately represented throughout. While there may be minor mishears, they do not impact readability or change the meaning of what was said. <strong>AND</strong> Punctuation is correctly used. Captions are readable.</td>
<td>Spoken audio is accurately represented in the majority of the project. There may be some errors that impact readability or change the meaning of what was said, but they are infrequent. <strong>AND/OR</strong> Punctuation errors that impact readability may be present but are infrequent.</td>
<td>Spoken audio is sometimes misrepresented. Multiple errors are present that moderately impact readability or change the meaning of what was said, though the overall capture shows intentional effort. <strong>AND/OR</strong> Punctuation errors that impact readability may be present and are infrequent.</td>
<td>Spoken audio is frequently misrepresented. Multiple errors are present that significantly impact readability or change the meaning of what was said, though some intentional effort was made. <strong>AND/OR</strong> Punctuation errors make readability difficult.</td>
<td>Project appears to be incomplete, unedited, or capture is so poor that the final deliverable is unusable.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Formatting</th>
<th>5 - Excellent</th>
<th>4 - Good</th>
<th>3 - Needs Improvement</th>
<th>2 - Poor</th>
<th>1 - Usable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Are the captions styled according to Rev’s guidelines?</td>
<td>Captions are well-formatted, with no egregious style guide violations, though a few errors that don’t impact readability may be present. <strong>AND</strong> Speaker changes have been correctly indicated.</td>
<td>Captions are well-formatted, with no egregious style guide violations, though a few errors that somewhat impact readability may be present. <strong>AND/OR</strong> Speaker changes have been mostly correctly indicated, though there may be an occasional error.</td>
<td>Captions are sometimes poorly formatted. Egregious style guide violations are present that moderately impact readability, though the overall capture shows intentional effort. <strong>AND/OR</strong> Speaker changes are mostly correctly indicated, though there may be more frequent errors.</td>
<td>Captions are poorly formatted. Egregious style guide violations that significantly impact readability are present throughout, though some intentional effort was made. <strong>AND/OR</strong> Speaker changes are not correctly indicated, with several errors.</td>
<td>Project appears to be incomplete or unedited, or no attempt was made to follow formatting guidelines.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Syncing</th>
<th>5 - Excellent</th>
<th>4 - Good</th>
<th>3 - Needs Improvement</th>
<th>2 - Poor</th>
<th>1 - Usable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do the caption groups align with the audio?</td>
<td>Caption groups are synced within 0.5 seconds of speech throughout the project, with no significant timing errors.</td>
<td>The majority of caption groups are synced within 0.5 seconds of speech, but there are occasional timing errors. Overall readability is not impacted.</td>
<td>Several caption groups are timed incorrectly, somewhat impacting the readability of the captions.</td>
<td>The majority of caption groups are timed incorrectly, significantly impacting the readability of the captions.</td>
<td>Project appears to be incomplete or unedited, or no reasonable attempt was made to sync the project.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Overall</th>
<th>Maximum number of errors allowed to achieve this score (average per grading clip)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimal to no effort made</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>