



Rev Captioning Style Guide 3.4

Updated on: December 6, 2018

Recent updates to the style guide:

- [Slide 22](#) - Rating or CC logos added to pre-existing on-screen text exceptions.



Tip: To quickly search this guide use ⌘ + f, or Ctrl + f.

7 Rules of Captioning

We want to delight our customers with high quality captions.

To do that, always follow the 7 cardinal rules described in this guide. Use your best judgment to apply these rules to any circumstances you encounter.

Category	What You Should Do	
Caption Formatting	1	Break caption groups at logical places so the text is easily readable by the audience.
	2	Break caption groups so they are always under 60 characters long.
	3	Use a dash and a space “- ” to indicate a speaker change. Add speaker IDs and atmospherics as appropriate, to give additional description.
	4	Use the up-arrow caret “^” to indicate every time a caption group occurs at the same time as on-screen text in the lower 1/3 of the screen.
Word Accuracy	5	Listen carefully to the dialogue and accurately type out the words with minimal errors and guesses. Never add words that aren’t spoken.
	6	Pay extra attention to the spelling and capitalization of special words and proper nouns. Research spellings if you’re uncertain.
Syncing Alignment	7	Sync the start of a caption group within 1/2 second of when the sound begins. This applies to both speakers and atmospherics.

Table of Contents

Common Terms	Slide 4
Keeping Track of How You Are Doing	Slide 5
Project-Specific Instructions & Getting Help	Slide 7
Break Caption Groups at Logical Places	Slide 8
Caption Groups Must Not Exceed 60 Characters	Slide 10
The Sweet Spot for Caption Group Length	Slide 11
Advanced Caption Breaking	Slide 12
Speaker Labeling	Slide 13
Atmospherics	Slide 17
Handling Pre-Existing On-Screen Text ^	Slide 22

Accurately Type Out the Words	Slide 25
Lightly Edit for Readability	Slide 26
Common Homophones	Slide 27
Punctuation and Hyphens	Slide 28
Ellipses and Quotation Marks	Slide 29
Numbers and Equations	Slide 30
Acronyms and Non-Letter Symbols	Slide 32
Technical Terms and Common Conventions	Slide 33
Spell Words Correctly	Slide 34
Sync Caption Groups to the Start of the Sound	Slide 35

[Captioning with Dash Guide](#)

Common Terms

Here is a list of some terms you will encounter throughout this style guide.

Captions Captions are the audio content of a video in written form. Whenever you make a subjective choice in creating the captions, consider the viewer: if you could only read the captions and not hear the audio track, would you understand what's going on in the video?

Dash This is Rev's online proprietary software to caption videos.

Caption Groups A unit of text that is shown on-screen. These are the text boxes within Dash that you create to enter text such as dialogue or atmospherics. A caption group includes the timing of when to display its text during the video.

Caption Group Breaking This refers to when you should create a new caption group.

Caption Group Length The number of characters, including spaces, in a caption group. Maximum is 60 characters, including spaces, per group.

Atmospherics These are the non-dialogue sounds you hear during a video such as music or sound effects.

Keeping Track Of How You Are Doing

The Grading Process

The goals of the grading process are to ensure customers receive high quality work and Revvers receive feedback about how to improve as a captioner.

After submitting a project, your work may be reviewed by another experienced Revver. As a rookie, all of your projects will be sent to grading to provide you feedback on your work. When you advance to the Revver level, a smaller percentage of your work will be sent to grading to provide ongoing feedback.

When a project is graded you will receive written comments about what you did well and what you can improve. You'll also receive a score of 1-5 across three dimensions: Accuracy, Formatting, and Alignment.

- 5 Excellent = Near perfect caption.
- 4 Good = Customer-ready as-is, but noticeable number of minor errors.
- 3 Fair = Almost customer-ready, **one or more major errors**.
- 2 Poor = Some usable sections but major improvement needed.
- 1 Very Poor = Unusable, or the transcript is incomplete or content is omitted.*

The scores you receive from Grading determine your metrics and are very important. You will need to maintain certain [metrics](#) to continue working with Rev.

However, if your metrics fall below our Revver requirements you will receive an email warning notifying you of the areas we need to see improvement. If your quality and/or professionalism does not improve, your Rev account may be closed.

* If you submit incomplete work, your pay for the job will be removed and your account may be closed after Support review.

Keeping Track Of How You Are Doing

What are Major Errors?

Accuracy	<u>Egregious Mishears</u> - Typing the incorrect words from what is spoken.
	<u>Omission or Addition</u> - Leaving out dialogue or adding in unspoken dialogue.
	<u>Incorrect spelling</u> of researchable terminology.
Alignment	<u>Excessive incorrect alignment</u> - caption groups appearing more than half a second too early or late.
Formatting	<u>Incorrect use of dashes</u> to indicate a new speaker or change of speaker.
	<u>Omission of up-arrow carets ^</u> when needed or adding up-arrow carets when unneeded.

Why are Major Errors important?

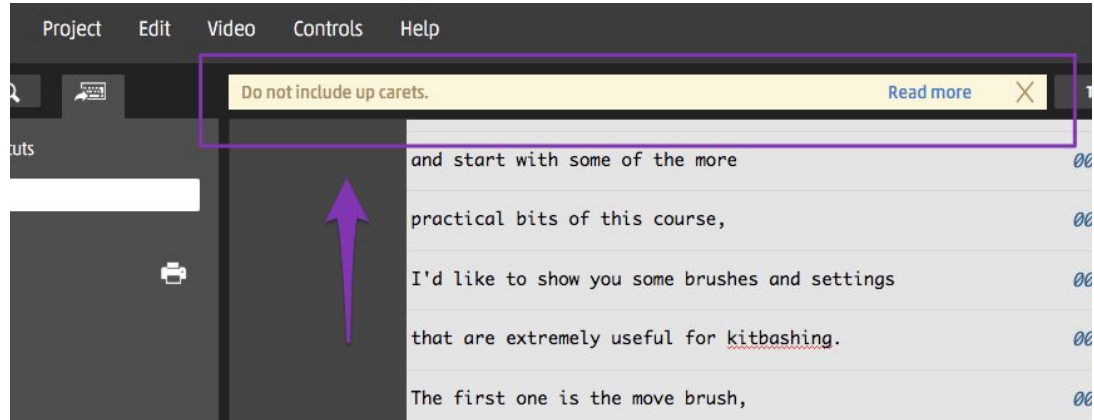
Major errors are the most common reasons that customers return files to Rev to be re-done. The presence of major errors in a captions file can cause the file to be rejected by video platforms and cost our customers large amounts of money.

Submitting work with major errors will have a negative impact on your metrics, and your metrics determine whether you're able to continue working with Rev. As noted on the prior slide, the presence of a major error will result in a grade of 3 or less.

Project-Specific Instructions & Getting Help

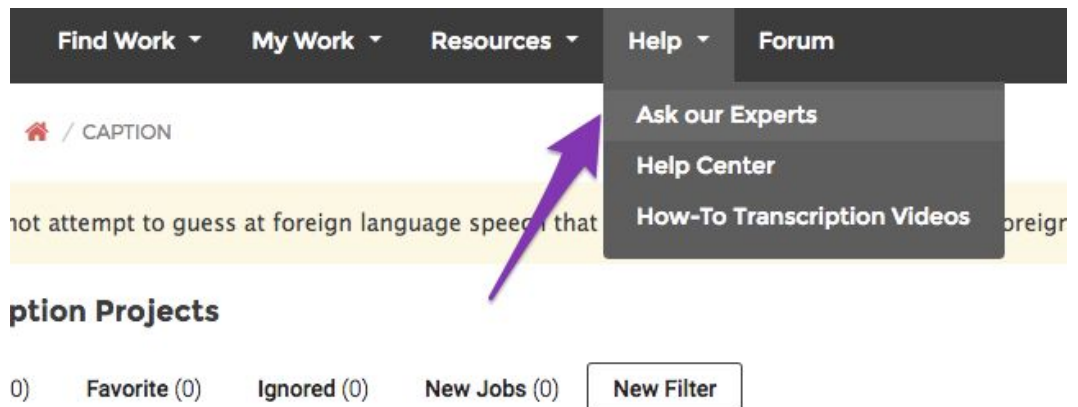
Look for special instructions from customers or support

1. Occasionally, a project may have special instructions that deviate from our normal guidelines. These instructions will appear in a yellow box in Dash.
2. **Only instructions in the yellow box are to be followed.** Customers may sometimes include instructions in an uploaded script or with provided glossary terms. These are not approved and should be ignored.



Get help easily from Rev Experts or support

1. Use the "Ask our Experts" feature to get help or answers on project-specific questions.
 - Include a link to your project in Dash.
Select *Video > Share Video* > then CTRL + V to paste the URL.
2. For general questions, email our support team at support@rev.com.



Break Caption Groups at Logical Places

1

Rule of thumb: A caption group is the unit of text that is shown on-screen at any one time. Split up speech and atmospherics at logical places using the Enter key, so the text is easily readable. This splitting is called caption breaking.

How to decide when to end one caption group and create a new one?

- Look for natural breaks using the video context and audio keys.

1. Pauses in speech, usually accompanied by:

- a. Punctuation , ! ?
- b. Pronouns, adverbs, and prepositional phrases such as: that, who, in order to, not only, as we, in which, where, with, what, how, for, through, until, to, as, of, yet, so, by
- c. Conjunctions such as: and, nor, but, or, because

When breaking for punctuation (bullet a), hit Enter after the punctuation. For prepositional phrases and conjunctions (b and c), hit Enter before the word.

2. Ends of sentences

- a. Create a new caption group by pressing enter after the end of every sentence.
- b. You **cannot** exceed 60 characters in a caption group. The caption group will turn **red** in Dash to let you know you have exceeded the character count limit and that you need to break the caption group.
 - Exception: If two short sentences are said by a speaker in under one second you can put them in the same caption group. For example: - Like, I know. You don't.
You can never combine a partial sentence with another partial or whole sentence in one caption group.

3. Changes in Speaker

- a. Each time a new speaker begins, create a new caption group.

If a speaker is talking very slowly and a sentence takes longer than 5 seconds to say...

Separate the sentence into more than one caption group.

Example: what it means for us (3 second pause) to be a high performance team

Use the speaker pausing to properly separate the captions.

what it means for us

Align caption groups to when dialogue is spoken in the video.

to be a high performance team

Break Caption Groups at Logical Places - Examples

1

Sentence Breaking Example

This is an example of *incorrect* caption breaking. These breaks make for awkward reading:

hit the earth. Or this kind of
meteorite, this is the current theory that the meteorite
hit the earth, and mass extinction occurred. So we think
that happened 65.5 million years ago.

This example *reads much better* because it follows the rhythm of speech, breaking at slight pauses:

hit the earth.
Or this kind of meteorite,
this is the current theory that the meteorite hit the earth,
and mass extinction occurred.
So we think that happened 65.5 million years ago.

Caption Breaking Example Sentence

When splitting up a long sentence, balance it between the neighboring caption groups. You cannot have a couple of words from the beginning or end of a sentence hanging off in their own caption group. For example, it looks bad to have “team” on its own:

what it means for us to be a high performance
team

- Instead, if the sentence is said in less than 5 seconds, combine them into one caption group since it is less than 60 characters:

what it means for us to be a high performance team

- Or if it takes longer than 5 seconds for the speaker to say the sentence, separate into balanced caption groups.

what it means for us
to be a high performance team

Caption Groups Cannot Exceed 60 Characters

2

Break caption groups so they are always under 60 characters long.

Dash includes a simple character count to help guide you:

- When you type 50 characters, the typing box turns yellow because you are nearing the limit. It's fine to be in the yellow, but break the caption group (hit Return / Enter) soon, preferably at a logical place as mentioned on the [previous slide](#).

If I type a caption group that exceeds 50 characters	00:00.00
the caption box turns yellow.	00:00.00

- When you go over 60 characters, the typing box turns red because you are over the limit. Remove some text from that caption group until the red color disappears.

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
---	----------

Dash will not allow you to submit your project if you have captions exceeding 60 characters.

The Sweet Spot for Caption Group Length

2

Breaking caption groups well will help you meet the syncing requirement of having readable captions

- It is better to break long sentences into the fewest possible number of caption groups to maintain readability.

This is an example of *poor* caption breaking because it is rapid continuous speech so each short caption will flash across the screen when you sync them:

hit the earth.
Or this kind of
meteorite,
this is the current
theory that
the meteorite

This example *reads and syncs much better* with longer caption groups because the captions will be on-screen long enough to be easily readable.

hit the earth.
Or this kind of meteorite,
this is the current theory that the meteorite

Dash shows you the sweet spot in caption length

It's fine to have short caption groups, such as when a speaker says a single word or when you have a short atmospheric. However, when you have continuous speech, aim for caption groups that show up as green or yellow in the typing window:

This caption groups is like Goldilocks.

00:00.00

Advanced Caption Breaking

When **multiple** sounds happen simultaneously, or very close together, AND the sounds are all vital to the story being told, use Advanced Caption Breaking by pressing **Shift+Enter** to include a second line within the caption group.

- If there is enough time between sounds or dialogue to include them each in their own caption group, do so.
- Two lines is the maximum per caption group.
- **NEVER** use a speaker ID when using advanced caption breaking. Include a dash on the second line to signify the change of speaker.
- Each line cannot exceed 30 characters. The typing box will turn red if a line exceeds 30 characters.

How to caption crosstalk (two characters talking over each other):

- Another character starts talking over the first speaker, so his speech is put on its own line within the same caption group.
- The next speaker after the caption group with advanced caption breaking must follow our regular [speaker ID guidelines](#).

Main speaker is speaking →

Second speaker talks at the same time

Next speaker follows normal ID guidelines →

- Sir, our scouts have returned and report	01:50.38
that the orc army is nearby.	
- I hear something.	01:50.38
- [Head Scout] We're awaiting your orders.	01:50.38

How to handle important simultaneous atmospherics:

- Use your judgment for whether each atmospheric is necessary for telling the story.
- Each atmospheric must be on its own line.

(spooky haunting music)
(door slams)

00:00.00

Rule of thumb: Speaker labeling is informing the viewer who is speaking.

Three Rules

1. **Use a dash and a space EVERY time a NEW speaker starts speaking or when the speaker CHANGES.** I.e., “- ”
 - Atmospherics do not count as a change of speaker. If an atmospheric is used that breaks up dialogue from the same speaker, do not include a dash after the atmospheric.
Example: - I don’t want anything, I’m not hungry.
 - Well, I am starving.
 (bell rings)
 I think I’ll get a hamburger.
2. **If the speaker CANNOT be visually identified, identify the speaker with a speaker ID.** E.g., “- [Mark]”, “- [Narrator]”
 - Details on how to identify a speaker are on the next slide.
 - If a speaker is unknown, use an appropriate identifier. Some possible examples could be “- [Interviewer]”, “- [Guest]”, “- [Ghostly Voice]”
 - It’s ok to reuse the same speaker ID in the video for multiple speakers throughout the video as long as those speakers are never on-screen at the same time.
 - Exceptions:
 - Do not identify the speaker if they begin talking on-screen OR off-screen but they ARE seen speaking during their current dialogue AND they’ve not been interrupted by another speaker since they started talking.
 - If a visible speaker is expressing unspoken thoughts, use an ID to help clarify to the viewer. E.g., “- [Mark Voiceover]”.
3. **If there are MULTIPLE speakers talking at the same time, you can:**
 - Use [Advanced Caption Breaking](#) if the crosstalk captions cannot be on-screen long enough to be read clearly.
 - If a group of people are speaking in unison you need to use a descriptive ID such as “- [Congregation]”, “- [Students]”, “- [Crowd]”, etc. to provide clarification to the viewer that multiple people are speaking in unison.
 - Focus on the main story thread. As stated in our [light editing for readability](#) guideline, it’s okay to not type meaningless interjections such as an interviewer saying “mm-hmm”.

When you need to identify a speaker, follow these guidelines:

- Use the speaker's *first name* if it is known: e.g., “- [Mark]”. You must be consistent with speaker ID's. Use the same speaker ID for a specific speaker throughout the video.
 - Exception: Use the name as it is known to the audience at the time. For example, if a character is introduced as Mark but then later is revealed to actually be Theodore, start the captions using Mark and then switch to Theodore when the reveal happens.
 - If you are captioning a TV series you may use the names of the main characters as the audience is already familiar with them.
- If a character name is not known, use a visible descriptive identifier: e.g., “- [Blonde Woman]”
 - It is ok to use “- [Man]” or “- [Woman]” *if they are the only man or woman at the time*; if more than one man or woman, then provide more description.
 - If gender is not known, use “Person” and a descriptor. E.g., “- [Tall Person]”
 - Never use numbers such as Narrator 1 or Narrator 2 in a speaker ID.
 - Never use race or other discriminatory identifiers. Instead, use a descriptor such as occupation, clothing, height, etc.

Some Tips and Tricks

- When doing a movie, first check for credits at the end of the video.
 - If they exist, they will tell you the proper spelling of character names.
 - It may also suggest helpful descriptors for characters that are unnamed.
 - E.g., “- [Policeman]” or “- [Screaming Girl]”
 - You can also check [IMDB.com](https://www.imdb.com), Google, or LinkedIn for names, spellings, and pictures of characters.
 - Review this [Help Center Article](#) for additional speaker ID examples.

How to Label Speakers on Audio-Only Projects

Occasionally there may be projects that are audio-only. These will display a solid black screen when playing the video.

- **Since there is no video, assume the speaker is ALWAYS on screen.**
 - Use a dash and a space EVERY time a NEW speaker starts speaking or when the speaker CHANGES. I.e., “- ”
 - Never use a speaker ID.

Example 1: Basic use of the dash.

First speaker spoke these two lines.

- What are we doing here? 00:20.40

(car honks loudly) 00:20.40

I need to go get food. 00:20.40

Second speaker begins talking on-screen, so use "-" to indicate the change of speaker.

- But you just ate. 00:20.40

Example 2: Speakers are never seen speaking on-screen.

Off-screen narrator spoke these two lines.

- [Narrator] The community came together that day. 01:36.14

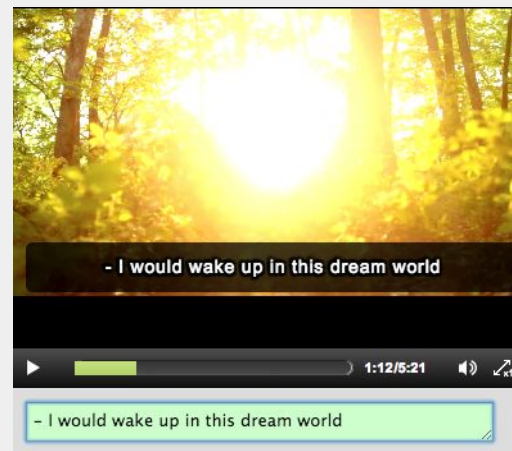
It was the most fun many of the children had ever had. 01:37.14

Mark, who is off-screen but was seen on-screen previously, speaks this line.

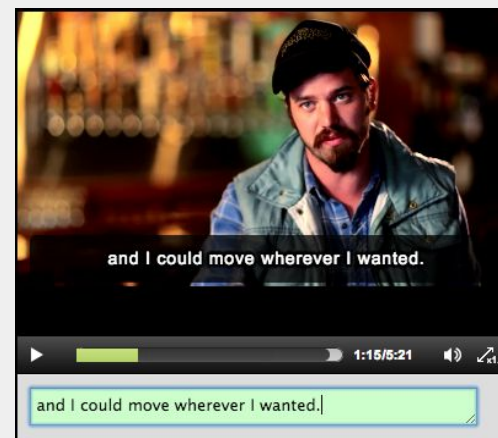
- [Mark] Seeing so many children laughing and playing, 01:38.82

Example 3: A speaker starts off-screen but then moves on-screen during the scene.

Speaker starts talking off-screen,

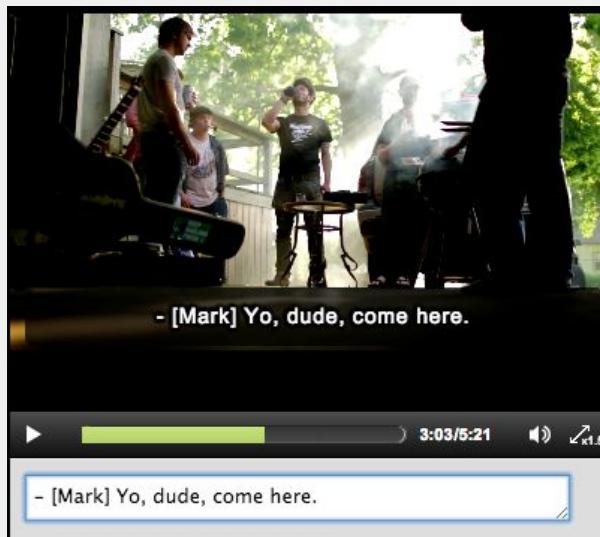


then is shown on-screen.



Example 4: Multiple people are shown on-screen and the viewer can't tell who is speaking.

Mark is part of the group but the viewer can't identify who is speaking. Here you need to include a speaker ID.



Example 5: Two speakers talking simultaneously, or crosstalk.

The original speaker is talking while another character speaks simultaneously, use [Advanced Caption Breaking](#).

I don't like this anymore.
- This is the best thing ever.

Rule of thumb: Atmospherics are typed descriptions of non-dialogue sounds. They allow viewers to understand important sound effects or what may be happening in a video when there is no dialogue.

You need to include atmospherics for 4 main conditions:

1. **A sound effect which is integral to the story or message of a video.**
 - Only include significant sounds effects that help tell the story. Use your best judgment.
 - i. If in doubt, include the atmospheric.
 - ii. If a character reacts to a sound, e.g., “(gun bangs)”, “(plane engine roars)”, or “(car honking)”.
 - iii. If a sound is the main focus point, e.g., a group of children playing, use “(children laughing)”.
 - iv. Include sounds made by the speaker, e.g., “(laughs loudly)”.
2. **Background music sets a specific mood as part of the story telling.**
 - Only include a background music atmospheric if there’s a significant gap in speech and the music seems important. E.g., “(dramatic orchestral music)”
 - Introductory music is a common atmospheric. E.g., “(gently chiming bells)”
3. **Music with clearly audible lyrics that are part of the story telling.**
 - If the music lyrics add content to the story, e.g., TV theme songs or when a character puts on a song and dances to it. You need to include the music note symbol. E.g., “♪ We will, we will, rock you”
 - i. Check out more details on lyrics on the [next slide](#).
4. **A speaker speaks in a foreign language.**
 - Follow the speaker labeling conventions and describe with the atmospheric “in foreign language.” E.g., “- What was that? (yells in foreign language)”

Atmospheric Formatting

- Atmospherics are put in parentheses and are always in lowercase, e.g., “(loud snoring)”, unless a proper noun is used, e.g., “(Mark laughs)”
 - Parentheses can only be used for atmospherics, never dialogue.
 - Atmospheric-only caption groups do not need a dash or speaker ID.
- Atmospherics must always be present tense, e.g., “(laughs loudly)”, never past tense, e.g., “(laughed loudly)”.
- Always describe with an action verb. E.g., “(frogs croaking)”, never with an onomatopoeia of a sound, e.g., “(ribbit, ribbit)”.

Caption song lyrics when one of these conditions is met:

- It appears the creator of the film took the time to make sure the lyrics were heard.
- It's a theme song for a TV show or songs performed by a speaker in the video.

How to caption lyrics:

- Add a musical eighth note “♪” at the start of every caption group containing the lyrics. You can add these notes in Dash by typing “##” followed by a space.
- Do not include a dash or speaker ID for lyrics.
- Capitalize the beginning of each caption group.
- Do not use ending punctuation such as . ! ? when typing lyrics, but do create a new caption group for each lyric phrase.
- When the lyrics pause or end, the next caption group with non-singing dialogue must follow normal conventions for a change of speaker. For example, if the person who was singing is now speaking, no dash is needed. However if the speaker is different from who was singing, add a dash and a space “- ” and possibly a speaker ID, if needed.

♪ Whenever you need me

00:00.00

♪ I'll be there

00:00.00

♪ Don't you know baby

00:00.00

Music is often used in videos to help set a mood or underscore actions.

If there's at least a 2 second gap in speech AND it does not seem the lyrics are intended to be clearly heard AND the background music is setting a specific mood, then caption the atmospheric as mood music:

- If you know the title and artist, put the song title in quotes, along with the artist, inside parentheses:

("Jailhouse Rock" by Elvis Presley) or ("Jailhouse Rock")

- Use the smartphone app "Sound Hound" to identify specific music when possible.
- You can also check the credits of a video to see if songs are listed to help you identify the correct song title or research lyrics.
- Most of the time you won't know the artist or title, so you need to include descriptive adjectives:
 - E.g., "(gentle music)", "(bright pop music)", "(heavy metal music)", "(electronic dance music)".
 - A list of music adjectives can be found here: <http://pianoaddict.com/map-words-characters/>
- If someone talks over the music, focus on the speaker and don't add the atmospheric. Dialogue is always more important.
- Sometimes music is used to signal a change in the scene. If this occurs, include an atmospheric at the start of the music and then when the tempo or style of the music changes dramatically, add a new atmospheric describing the change. E.g:
 (lively banjo music)
 (moves into slow dance music)
- When in doubt, ask Support for help via the Help Button.

Atmospheric Positions

When incorporating atmospherics, there are 3 possible positions you can put them.

1. If the atmospheric lasts for at least 2 seconds OR has some silence before and after the atmospheric, put the atmospheric in its own caption group. E.g.:

(bar patrons talking loudly)

02:27.84

2. Otherwise, if the sound is made by the speaker, put the atmospheric in-line with the dialogue at the same point you hear it occur. E.g.:

I (laughs) couldn't believe it!








01:38.00

3. Otherwise, if an important atmospheric overlaps with dialogue OR is less than 2 seconds long, use [Advanced Caption Breaking](#) to put it on its own line within the same caption group. E.g.:

(child wailing)
If the rules had been enforced

02:33.86

Here are some common issues when using atmospherics:

			Example
	DO	Include adjectives describing the type of music.	(slow violin music)
	DO NOT	Use “sound of”. The viewer knows the words in parentheses are sounds or music currently being heard.	(sound of a bell) (sound of people laughing)
	DO NOT	Use an onomatopoeia	(ribbit ribbit ribbit)
	DO NOT	Use the words intro, introductory, ending, fading, playing, etc. when describing music.	(introductory music playing) (music fading out)
	DO NOT	Use “again”, “previously”, “repeated”. If a sound is being heard repeatedly, simply include the same atmospheric again and sync it to the sound being heard.	(bell rings again) (clock ticking heard previously)
	DO NOT	Describe the action that created the sound.	(opening box) (kicks desk)
	DO NOT	Indicate a speaker’s inflections through atmospherics.	- (emphasizing) When is it? - (whispering) Hello?

Rule of thumb: Videos sometimes have pre-existing text in the lower $\frac{1}{3}$ of the screen. If you need to type captions that will appear at the same time, even for only a split second, we need to flag the caption with an up-arrow caret ^ (Shift+6).


What is On-Screen Text?

On-screen text is text that has been added by the filmmakers in post-production that contains important information for the viewer. We add an up-arrow caret ^ to caption groups that appear at the same time as on-screen text in the lower $\frac{1}{3}$ of the screen. These caption groups will then be shown at the top of the screen for the customer, instead of covering the on-screen text shown in the lower $\frac{1}{3}$.

These are some examples of on-screen text in the lower $\frac{1}{3}$ when you need to use an up-arrow caret ^:

- Speaker names, social media usernames, storytelling information.

The following types of text do not count as on-screen text. **Do NOT** use an up-arrow caret ^ for:

- Text on the environment in a video. Example: A gas station sign, a number on a race car.
- TV or corporation logos, including rating/CC information such as: 
- A running timecode throughout the video.
- Pre-existing text shown as slides, graphics, whiteboards, or a software interface. This is common for educational presentations.

How to indicate captions with on-screen storytelling text in lower $\frac{1}{3}$ of the screen:

- Up-arrow carets ^ must be used whenever there are captions at the same time as pre-existing burned-in text on the lower $\frac{1}{3}$ of the screen. Dash provides pre-existing text guide markers to help you identify when to use the up-arrow caret.
- When this happens, add the up-arrow caret ^ any place within the caption group and it will display as a blue marker to the left of the caption group.
- You must add an up-arrow caret ^ to EVERY caption group that occurs at the same time as the burned-in text, even if it is only for a split second.
- IF there is on-screen text in both the bottom $\frac{1}{3}$ and the top $\frac{1}{3}$ of the screen at the same time, **do NOT** include the up-arrow caret since important text will be covered in either location.

Example 1: pre-existing captions or subtitles.

Usually captions and subtitles will be in full sentences, displayed on the bottom center of the screen. Do not type captions here.

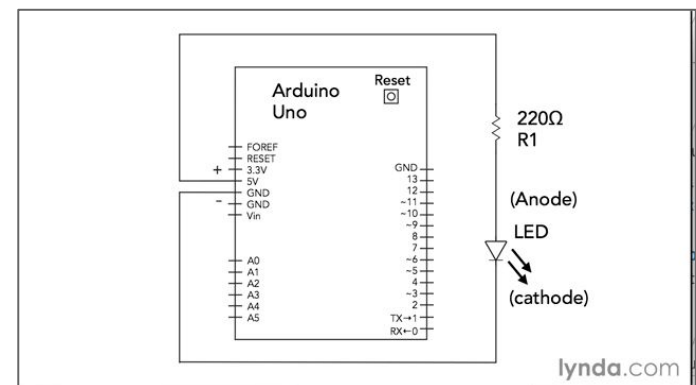
DON'T retype this burned-in subtitle



Example 2: pre-existing text in powerpoint slides, graphics, logos, charts, etc.

Type out the dialogue as usual.

DON'T use the up-arrow caret ^



Example 3: pre-existing text that's not in the lower 1/3 of the screen.

Type out the dialogue as usual.

DON'T use the up-arrow caret ^



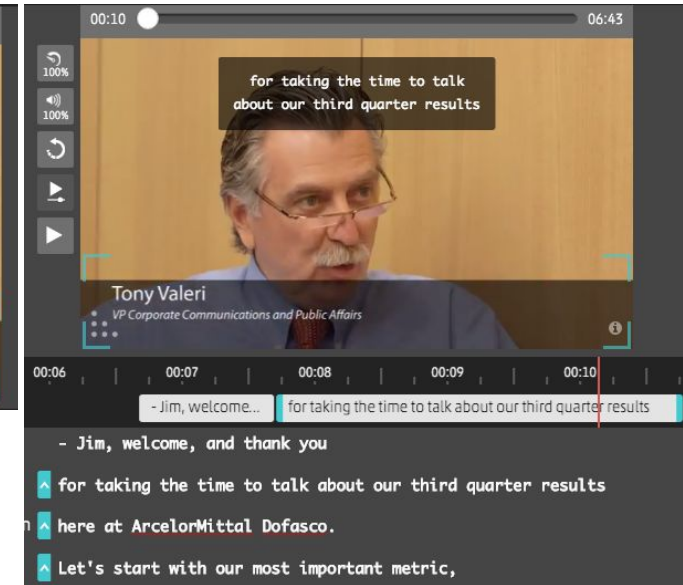
Handling Pre-Existing On-Screen Text - Examples

MAJOR ERROR

4

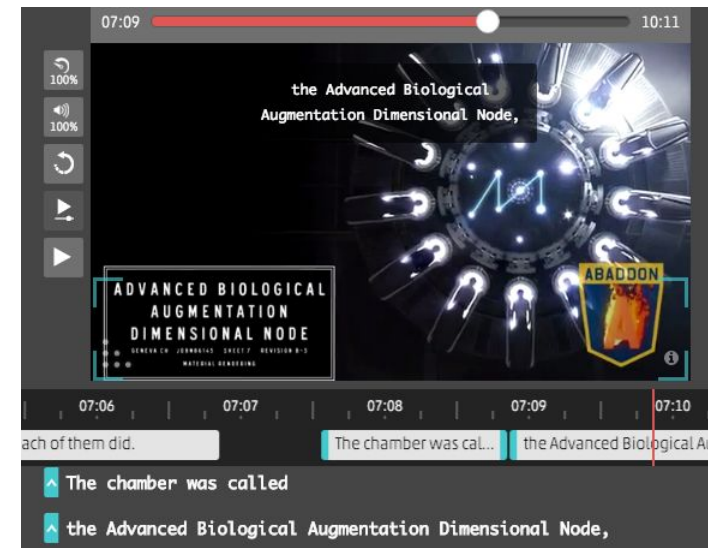
Example 4: Information shown about the video in the lower $\frac{1}{3}$ of the screen.

DO use an up-arrow caret ^ in every caption group that occurs at the same time as the burned-in text. Use the Dash guide marks to help identify the area.



Example 5: Storytelling information shown to the viewer in the lower $\frac{1}{3}$ of the screen.

DO use an up-arrow caret ^ in every caption group that occurs at the same time as the burned-in text.



Rule of thumb: Listen carefully to the dialogue and accurately type out the words with minimal errors and guesses. Never correct the speaker's grammar or add words that aren't spoken. Be consistent with punctuation and symbols.

Type what the speaker says. You must ALWAYS caption what is heard and *always use American English spelling*.

- Never correct (edit) the speaker's grammar (morphology, syntax, and semantics).
- Never paraphrase.
- Never substitute words.
- Never add words that are not spoken.
- Never rearrange the order of speech.
- Don't correct phonetics unless it distracts from readability. See the next slide, [editing for readability](#).
- Do not remove speech disfluency that distracts from readability. See the next slide, [editing for readability](#).



Correct:
captioned the
speaker's words
as-spoken

- I live right by a beach, so I go to the beach
with my friends during summer.
I like dancing, playing volleyball, and going driving.
That's pretty much what I do.



Incorrect:
modified tense of
verbs and added
words.

- I live right by a beach, so I go to the beach
with my friends during *the* summer.
I like to drive, dance, and play volleyball.
That's pretty much what I do.



Incorrect:
extreme
paraphrasing and
re-ordering.

- *I like dancing, playing volleyball, going driving,
and going to the beach with my friends.*

Our goal is readability, so it's preferable for you to remove extraneous text that will likely distract a viewer from the core message:

- Omit speech disfluencies*: unnecessary filler words, false starts, stutters, repetitions, etc.
- Omit quick interjections, such as an interviewer saying “mm-hmm”, unless a direct response to a question.
- Correct *egregious* phonetic and pronunciation errors that *inhibit readability*.

However, never change the story being told:

- Don't correct a speaker's grammar or pronunciation that is easily understood. E.g., “gonna” must stay as “gonna”.
- NEVER censor or edit expletives. If the word is censored with a beep sound, use (beep) or (bleep) in-line where the sound occurs.
- Never omit special words, entire sentences, or expletives.
- Don't be excessive. If in doubt, don't omit the word(s).



- ~~And so, um, I guess...~~ I think we should go to the, ~~the~~ movies tonight.



- ~~Um ...~~ You, ~~you know,~~ have to go underground.



- We need to fix the ~~washer~~ washer right away.



- Those guys are driving ~~fly, tricked-out~~ rice rockets.



- ~~Damn Eddy,~~ did you really have to tell them about that?

* For more information on speech disfluency, read http://en.wikipedia.org/wiki/Speech_disfluency

The following is a list of common homophones in which two words with different meanings sound very similar. Please make sure you use the correct word so that it makes sense in the context of the sentence.

Word	Alternative Word	Rule
Its	It's	"It's" means "it is" or "it has." "Its" is the possessive form of it. "It's keeping its rules straight."
Your	You're	"You're" is a contraction of "you are," while "your" is possessive. "You're working on your project."
Were	We're	"We're" is a contraction of "we are". "We're hoping they were chosen to win the competition."
Are	Our	"Our" is the possessive form of "we". "Are" is the simple present tense form of the verb "be". "Are you going to see our parents tonight?"
Let's	Lets	"Let's" is a contraction of "let us". "Let's go have dinner," as compared to, "She lets us use this."
Affect	Effect	Typically, "affect" is a verb and "effect" is a noun.
This	These	"This" is singular whereas "these" is plural.
Their	They're There	<p>"Their" means belonging to them. "There" states where someone or something is. "They're" is a contraction of "they are".</p> <p>For example: "They're keeping their bicycles over there."</p>

The primary role of punctuation is to aid readability by marking the structure and intonation of the dialogue as spoken. Spoken language can vary from written grammar rules but as a captioner it is our responsibility to capture what is spoken for the viewer as closely as possible.

- Use . ! ? as you normally would as sentence endings.
- Use commas to add clarity to the idea the speaker is trying to convey. Pay attention to grammatical structure and use commas:
 - To connect two independent clauses by a coordinating conjunction.
 - To connect dependent and independent clauses.
- Use colons “:”, semicolons “;” and single quotation marks ‘ sparingly as they are often misused/overused.
- Use your best judgment to break up run-on speech into shorter sentences for readability.

Hyphens -

You should hyphenate two or more words that precede and modify a noun as a unit, especially if the words include a past participle, a present participle, a single letter, or a number. For example:

- accident-prone
- custom-built
- one-by-one

If a speaker is abruptly interrupted by another speaker or sound effect, use a double-hyphen. NEVER use a single hyphen to indicate interruptions or trailing off. For example:

- The symphony should really get its--

- I don't care about the symphony.

Ellipses ...

Ellipses should be used sparingly and only at the end of a caption group if the speaker trails off their current thought AND pauses afterward for at least 1 second. For example:

- I'm not sure...

Never use ellipses at the beginning or in the middle of a caption group. Instead, you should hold off displaying the next caption group until the dialogue has resumed.



Incorrect

- Before I couldn't ... we always wanted to go there.



Correct

- Before I couldn't...

We always wanted to go there.

Double Quotation Marks “ ”

Quotation marks are only used for direct quotations when a speaker quotes another person directly or implicitly. Do not use single quotation marks ‘ ’ for direct quotations.

Format quotes by including “ at the beginning of every caption group that is still in the same quote. Only put a closing ” at the very end of the full quote. For example:

- She told me, "It's ridiculous.

"Maybe I can go ask for 12 planes,

"and they'll make the drawings in the sky."

Numbers

- Spell out single-digit numbers zero through nine, e.g.: “zero”, “one”, and “nine”. Use numerals for all other numbers. E.g.: “\$1,456,020.23”, “1.5”, “1/4”.
 - Some exceptions apply with common conventions for numbers with the goal of readability.
 - Caption common conventions as they are spoken, such as: “Genesis 1:1”, “\$2”, “February 1st”.
 - Percentages should be written: 4%, 8%, 15%
 - Proper nouns that contain numbers. For example: “Article III of the Constitution”, “Fifty Shades of Grey”.
 - If the speaker says “million,” “billion,” “trillion,” use those words instead of typing the numerical equivalent. E.g., “Eight billion.”

Time

- If an exact time is mentioned, format it as follows: “9:30 a.m.”
 - If spoken, always use lowercase a.m. (ante meridiem) and p.m. (post meridiem).
 - If the speaker says “o’clock”, caption as spoken. E.g., “nine o’clock.”
 - If the speaker doesn’t mention an exact time, follow the number conventions above. E.g., “I’ll see you tonight at nine.”
- Spell out words and phrases as spoken that do not include actual numbers. E.g., half past, quarter of, midnight, noon.

Math

- Follow basic number rules; use numerals for fractions (since a fraction uses more than one digit). For example:
 - - [Teacher] One plus one equals two.
 - - [Teacher] 2,000 plus three equals 2,003.
 - - [Teacher] 1/2 of this and 10 3/4 of that.
- Follow the conventions for ordinal place. For example: “Zeroth”, “first in line”, “10th place”, “21st century”, “100th time”.
- For anything that’s not a number / digit or percentage, write out the full word instead of the symbol. For example:
 - - X squared times 0.32%.
 - The derivative of y squared.
 - Five x times negative three x equals negative 15x squared.
- A teacher may write this on a whiteboard:

$$\text{Percentage of Bankroll} = \frac{(\text{Odds Received} \times \text{Probability of Winning}) - \text{Probability of Losing}}{\text{Odds Received}}$$

while speaking the equation aloud. You would caption his speech as:

- Percentage of the bankroll equals the odds received, multiplied by probability of winning, minus probability of losing, divided by odds received.

Graphing Terms

- Write it out as the speaker says it, following basic number conventions such that:
 - (-10,3) becomes “negative 10 comma three” or “negative 10 three” depending on their word choice.
 - Quadrants are labeled with Roman numerals, such as “quadrant IV”
 - Axes and coordinate references are hyphenated as follows: x-coordinate, y-axis.

Acronyms

- Only type an acronym if spoken that way by the speaker.
- For acronyms like FOIL (First, Outer, Inner, Last, used for binomial multiplication), be sure to keep it capitalized for all uses (e.g., “Now you try FOILing this next problem”).
 - To pluralize an acronym, add a lowercase “s” without an apostrophe (e.g., SATs,), unless the acronym ends in an S, in which case, add an “es”.
- Do *not* abbreviate unless the speaker specifically says it.
 - For Example: If the person says “gigabytes” then type “gigabytes”, not “gigs” or “GB”.
- If a speaker spells out a word, use the format “W-O-R-D”.

Non-letter Symbols

- We cannot use symbols or special characters such as é, £, €, or ². Only use what is available on a standard American keyboard.
- Non-letter symbols, such as pi, should have spaces in between them and both the preceding and next variable or term.
- Try to be as clear and consistent as possible using spaces as needed to avoid confusion, such as pi being mistaken for p times i. For example:
 - “Two pi r” NOT “Twopir”
 - Or if the speaker says “Two times pi times r”, then type it as spoken.

Never Use ALL CAPS

- We use proper capitalization which means only upper case the beginning of sentences and proper nouns. Never capitalize all the letters in a word or sentence except when you’re typing out an acronym such as “FCC” (Federal Communications Commission) or indicating a word that the speaker is spelling out such as “W-O-R-D”.

Technical Terms

- Videos that contain technical terminology, such as software walkthroughs or how-to's, may display technical terminology on-screen that the speaker also speaks. We must ALWAYS include dialogue that is spoken, even if it is shown on-screen, but you can type out these terms as shown on-screen to aid readability.
 - For example, the speaker says “console dot log” but is shown on-screen as “console.log” Use the term as shown on-screen to help readability for the viewer, i.e. “console.log”
- If a speaker in a technical video mentions keyboard commands, caption the commands as shown on-screen. If they are not shown on-screen, use the following format:
 - The speaker says “press shift command X”, this would be captioned as “press Shift + Command + X”
 - If multiple keyboard commands are mentioned in a row, separate with a comma: “Shift + Command + X, Control + Up Arrow.”

Common Conventions

- In addition to common conventions regarding numbers ([see slide 30](#)), you are allowed to follow these common conventions:
 - Website addresses may be captioned as “http://www.google.com” instead of “h-t-t-p, w-w-w dot google dot com”
 - Social media handles, such as for Twitter and Instagram, can be captioned as “@mytwitterhandle”

Rule of thumb: Pay extra attention to correctly spelling and the capitalization of special words and proper nouns. Research spellings every time you're uncertain. Use an atmospheric when you're still uncertain of what is said.

Take time to research the proper spelling and capitalization of important words and proper nouns

- Look at text on-screen. If the words are spoken, caption the words using the same spelling as what you see on-screen.
- Research proper spellings and capitalizations for important words. Google proper spelling of proper nouns (e.g., names, brands, and places) and topic-specific vocab (e.g., Adobe Premiere Pro). We provide helpful resources when they're available.
- Put in a reasonable effort to research terms but do not devote excessive time to this. If spelling is not easily findable, make your best guess using a common or phonetic spelling of the word.
- Spell words consistently throughout the captions.



- Ozzy Osbourne can't sing a *cappella* to save his life.



- Talk about a *yoo-foo-iz-uhm*!

Acceptable phonetic spelling of actual word used: "euphuism"



- Hi, I'm Rich *Thornet*, the founder of *Dribble*.

This is researchable.
Correct spellings: "Thornett" and "Dribbble"

When you cannot confidently hear or understand a word, use an atmospheric

- If the video intentionally creates an inaudible situation, use an appropriate atmospheric such as "(mumbles)" for a single speaker or "(background noise drowns out other sounds)". **NEVER** use "(inaudible)" or "(indistinct)" in captions.
- If I stub my toe, one more time, on this stupid (mumbles).
- View this [Help Center article](#) on how to handle unclear audio for additional examples.

Sync Caption Group to the Start of the Sound

MAJOR ERROR

7

Rule of thumb: Sync the start of each caption group so it appears on-screen at the same time as when the dialogue or atmospheric begins in the audio.

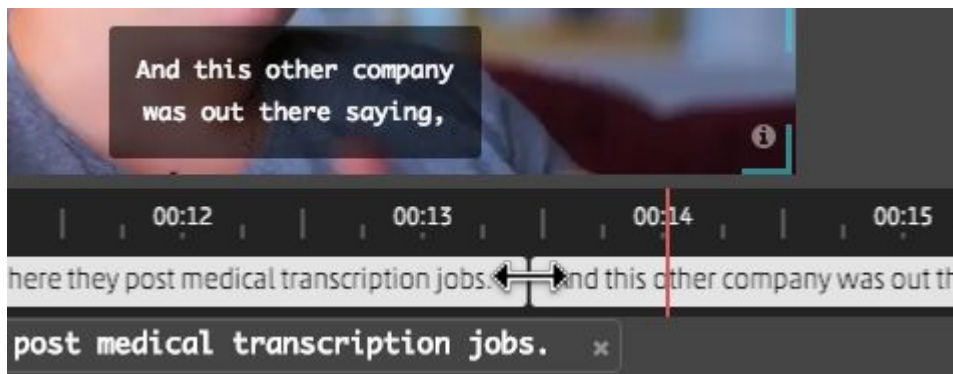
Start Time of a Caption Group

- The start time needs to align with the beginning of the sound. This applies to both atmospherics and speech.
- While the video plays, press the UP arrow key on your keyboard to align the start of a caption group precisely when the first word or sound of a caption group is heard.
 - Note: When you use the UP arrow key, Dash automatically bumps your timestamp forward by $\frac{1}{4}$ second to compensate for normal lag in human reaction time.
- Aim for precision, but it's ok for the start time to be up to a $\frac{1}{2}$ second early or late from the true beginning of the sound.

Adjusting Start and Stop Times

In the horizontal timeline of the “sync” step, you can use the left or right arrow keys on your keyboard to adjust the start time of the caption group.

You can also use your mouse to click and drag the caption group to adjust the start time.



Auto-Calculated End-Times and Minimum Caption Length

Rule of thumb: Dash automatically calculates the end time for a caption group. You are only responsible for syncing the start time of a caption group.

When to Edit Captions in Order to Adjust Caption Timing

1. If a speaker is talking very slowly or has a long, dramatic pause in the middle of a sentence, you can break up caption groups so that they do not end too EARLY.
2. If multiple speakers are talking very quickly you can combine captions using [advanced caption breaking](#) so that the caption does not end too LATE.