Transcriptive

Transforming The Art of Transcription

Automated Transcriptions, Video Search, and Captions

Digital Anarchy
Smart tools for creative minds.

www.digitalanarchy.com

Transcriptive Contents

Transcriptive
Transforming The Art of Transcription

Keyboard Shortcuts

These shortcuts are available in Tracking Mode for easy editing. Learning to use these can make basic editing much faster. Especially when combined with increasing the Playback Speed to 1.5x or 2x.

Navigation

Play/Stop: Spacebar
Next Word: Right Arrow
Previous Word: Left Arrow
Next Sentence: Alt/Option+Right Arrow
Previous Sent.:   Alt/Opt+Left Arrow
Next Speaker: Alt/Opt+Shift+Right Arrow
Previous Sp.:   Alt/Opt+Shift+Left Arrow

Capitalization

Make Upper case: Up Arrow

Make Lower case: Down Arrow

Line Merging/Splitting

Merge with line above: Alt/Option + Up Arrow

Split sentence and make new line: Alt/Opt + Down Arrow

Punctuation Shortcuts:

Period:	.	Left Single Quote:	'
Comma:	,	Right Single Quote:	Command/Control + '
Question Mark:	?	Left Double Quote:	Shift + "
Exclamation Point:	!	Right Double Quote:	Cmd/Cntrl + Shift + "
Colon:	:	Semi-Colon:	;
Left Parenthesis:	(	Right Parenthesis:	)

Delete Word: Backspace/Delete   Delete Punctuation: Cmd/Cntrl+Backspace

About This Manual

While the manual is a good resource, we recommend checking out the videos on Transcriptive to really get an understanding of how it works. Often it's easier to show something than to try to explain it with words, especially with something that's visual to begin with.

However, it can be hard to go into lots of detail in a video, so this manual usually provides more info than you'll find in the tutorial videos. The videos and manual are designed to complement each other.

Here are the tutorial videos, they're all 8 minutes or less:

Transcriptive Tutorial - Editing Transcripts and Keyboard Shortcuts
https://youtu.be/foR464DHW0M

Transcriptive Tutorial - Speech Engine Setup
https://youtu.be/pdWiKhQWqsU

Transcriptive Tutorial - Requesting New Transcripts
https://youtu.be/nRIA09UQWwA

Transcriptive Tutorial - Caption Formats
https://youtu.be/3yPpoupTitk

Transcriptive Tutorial - Installation and Activation
https://youtu.be/C88lpddR1rw

Transcriptive Tutorial - Installation and Activation
https://youtu.be/C88lpddR1rw

Installation

You can always download the most up-to-date version of Transcriptive from our Demo page. Your serial number will activate the demo.

Mac Demo:

https://digitalanarchy.com/demos/psd_mac.html

Download the .dmg file and run the Installer that's on the dmg.

This will install the Transcriptive panel and associated files into:
Macintosh HD/Library/Application Support/Adobe/CEP/

It will also install the FLAC exporter into:
Macintosh HD/Library/Application Support/Adobe/Common/Plugins/7.0/MediaCore/Digital Anarchy

Once the installer finishes, it'll bring up the Digital Anarchy Registration page. Please fill out the registration and close the installer.

Restart Premiere Pro and Transcriptive should show up in Premiere under Windows>Extensions>Transcriptive

If at some point you wish to uninstall Transcriptive completely (this is sometimes helpful in troubleshooting problems with a new version), run the Un-Installer.

Windows Demo:

https://digitalanarchy.com/demos/psd_win.html

Open the .zip file and launch the Transcriptve_Installer.exe. If you are updating or upgrading, you'll be asked if you want to uninstall Transcriptive. Click OK to run the un-installer. (highly recommended)

Click through the installer, agreeing to the license terms. The installer will automatically select the correct directory, don't change this unless you have an unusual system and you know where the panel should go.

This will install the Transcriptive panel, Premiere Power Search panel and associated files into:
C:\Program Files\Adobe\Adobe Premiere Pro CC 2017\CEP\extensions

It will also install the FLAC exporter into:
C:\Program Files\Adobe\Common\Plug-ins\7.0\MediaCore\Digital Anarchy

Once the installer finishes, it'll bring up the Digital Anarchy Registration page. Please fill out the registration and close the installer.

Restart Premiere Pro and Transcriptive should show up in Premiere under Windows>Extensions>Transcriptive

If at some point you wish to uninstall Transcriptive completely (this is sometimes helpful in troubleshooting

Your Serial Number

After purchasing your plugin, you should see an email with the Subject: Your Serial Numbers for Order #XXXXXX

Look for the pink highlighted text in the email and that will have your serial number.

Transcriptive will ask you for the serial number when you first run it within Premiere Pro.

If you have any problems, you can always contact us at cs@digitalanarchy.com or call 415-287-6069

What is Transcriptive?

Transcriptive is a new way of automatically transcribing video and audio. Making it simple to search the video, create subtitles, closed captions, and more. We allow you to use two different Artificial Intelligence/Machine Learning services to achieve unprecedented automatic accuracy. Why two? It makes it more cost effective as you can choose the appropriate algorithm for the quality of your audio. The two differ in cost and accuracy... see the How It Works section for complete details.

Speechmatics is the most accurate of the two by far. Please use Speechmatics if accuracy is critical to you.

Once the transcript is processed it then becomes available in Premiere to be edited, searched, saved out to markers, or exported as a caption file. You can search the text either in the Transcriptive window or Premiere's metadata panel, making it easier than ever to find where something was said in your video.

You can also import an existing script or transcript and sync that to the audio! Transcriptive will add time code to every word, allowing you to easily search the text and jump to the corresponding place on the Premiere timeline.

Exported files can be taken into any other application or used in Premiere Pro. Common uses are for closed captions, sub-titles, and SEO for YouTube.

How it works?

Transcriptive will transcribe the active sequence in Premiere Pro or you can batch process multiple clips. You select your sequence, set in and out points if necessary, and open the Transcriptive panel and click TRANSCRIBE.

If you're batch processing, select the appropriate Batch command and select your files.

You'll then be asked what Speech Engine you want to use, Speechmatics or Watson.

You'll need to sign up for those services on their websites and enter your credentials into Transcriptive. You only need to sign up for one of them.

Why do we make you do this?

It's for your benefit actually. By signing up directly with the services, you get the wholesale cost of the minutes. You only pay whatever Speechmatics or Watson charges. Digital Anarchy doesn't make any money by marking up the cost of the minutes.

What are the costs?*

As of this writing (September 2017), the costs were as follows:

Watson: 16 hours/mo for free then $.02/minute afterwards.

Speechmatics (Based in the UK, so this is in Pounds): 6p for the first minute, 1p for every 10 seconds after that. As of 9/2017, that makes it about $.07/min.

* There's been no indication from either service that they plan to change these prices, but they could change. Please double check their websites. The Digital Anarchy website will also show the most up to date pricing. (we try to keep the manuals up-to-date but it doesn't always happen)

Which Speech Algorithm Should I Use?

Speechmatics will give you superior results. It has about 95% accuracy on clean audio and slightly less on somewhat noisy audio.

If you have a lot of background noise, the choice is easy: Speechmatics. Watson does not handle this type of audio very well and you'll probably be unhappy with the results.

Speechmatics is also easier to setup.

If you have good, clean audio the choice is harder. Speechmatics is better but Watson is also good. Watson has about 90% accuracy rate vs the 95% of Speechmatics. If cost is no concern, use Speechmatics. You'll have less clean up. The time you save will be worth the pennies you spend.

However, Watson gives you 1000 FREE MINUTES a month. If you're transcribing a lot of video per month, this adds up to a substantial savings. Yes, there's more clean up required but, hey, the minutes are free.

If you're just using the speech-to-text to allow you to search your video and jump to where the text ia spoken, Watson also is compelling choice. Yes, it's less accurate but if you're just trying to search a lot of video it's a good option.

Of course, we suggest you run a few tests of your own on video footage with varying audio quality to see the results yourself. We've found the above to hold true for the vast majority of tests we've run.

Overall, we really recommend Speechmatics. However if cost is an issue, and especially if you're only using it to search the video, Watson is fine.

System Requirements

Transcriptive requires an internet connection. The only way around this is to use a different computer to upload the audio file and get the transcript. You can then move the resulting transcript file to the original computer and import it into Transcriptive. It's easy but is a bit more work. (see Importing JSON Files for more info)

You MUST sign up for at least one of the Speech Services. We recommend signing up for both, but you must enter credentials for at least one of them in order for Transcriptive to work.

Premiere Pro 2015.3 or later. Adobe added more robust panel support in 2015.3 and we need that functionality. We always support the most recent version of Creative Cloud, so Transcriptive should work whenever a new version is released.

Transcription Basics and User Interface

Launching Transcriptive

You can launch Transcriptive by going to:

Window>Extension>Transcriptive

This will open the Transcriptive panel. Feel free to dock the panel anywhere in the Premiere Pro workspace. You can also leave it as a floating window.

First Thing...

The first thing you'll need to do is to enter in your login credentials for the Speech Services in Speech Engine Setup area of the Main menu. You need to set up at least one speech services to use Transcriptive.

Signing Up with Watson and Speechmatics

Speechmatics is the easier of the two. Just go to their website and create an account. Once the account is created go to your Dashboard and you'll see the info needed to login from Transcriptive. Very easy. You get 60 minutes free to start. After that you need to purchase a block of credits. 6 credits = 1 minute. Speechmatics charges by every 10 seconds, so if you transcribe 2m20s, you'll use 14 credits. (2min (12 credits) + 20sec (2 credits))

Speechmatics is the most accurate of the two by far. Please use Speechmatics if accuracy is critical to you.

Watson is a bit of a pain in the ass. They assume you're a developer, which means it's more confusing. It's not hard, but more people have problems with it. Here are the steps for Watson:

- Go to https://console.bluemix.net and create a free account

- After 30 minutes or so you should receive a confirmation email from IBM. Confirm your account

- Login and you'll be asked to setup your account

- Set up your Organization: This can be anything, however, we recommend you select US South as the server. If you're in Europe, Asia or elsewhere this won't make sense but it's the option that will potentially cause the least problems.

- Give your Space (workspace) a name. Anything will be fine.

- Once you're at your dashboard, go up to where it says '29 trial days remaining'

- Select Add Credit Card and enter in a credit card. You have to do this. If you don't go over 1000 minutes you won't be charged anything. But IBM needs a credit card to get started.

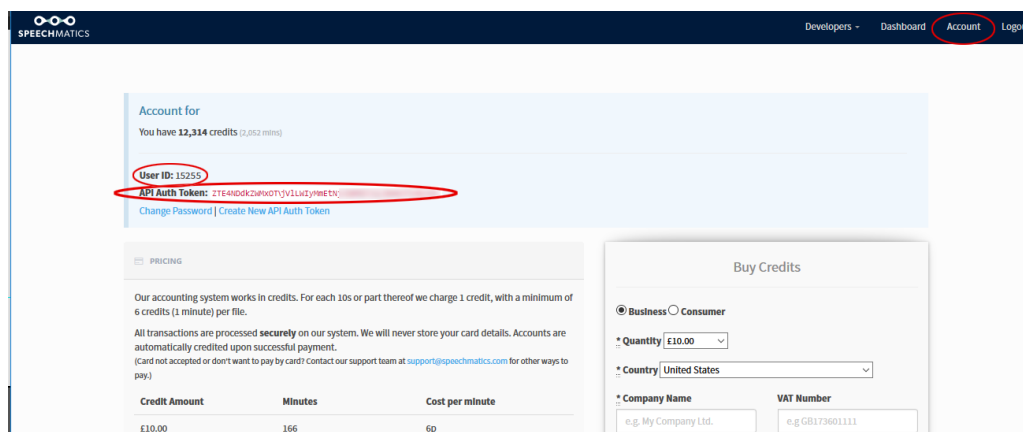You should be all set! Go forth and transcribe!

Speech Engine Setup

Transcriptive uses two different Speech Services and these need to be set up before you can use the plugin. You can enter in your credential information by going to the Transcriptive Main Menu and selecting Speech Engine Setup.

Once you open Speech Engine Setup you'll see areas to enter in your login credentials for both Services (You'll need to sign up for the services first. See the section prior to this for more info).

Watson: you only need to enter in your login info. Transcriptive will handle all the other setup.

Speechmatics: In addition to your User ID, you need to enter in the API Key. This can be found in your Speechmatics Account Dashboard. Log into Speechmatics and go to your Dashboard. You should see it as follows:

Copy the API key and paste it into Transcriptive.

Once all that is done, this is the only time you'll need to open Speech Engine Setup unless your password changes (which we recommend doing periodically).

Critical Things to Know

Clicking on The Transcriptive Panel

In order for Transcriptive to work, you have to click on the Transcriptive panel. If you just select your sequence and hit play, Transcriptive won't do anything. This is a limitation of Premiere. You have to select the Transcriptive panel and hit play from within Transcriptive otherwise Transcriptive won't have any idea what Premiere is doing. Transcriptive is technically a separate app running within Premiere. It will communicate with Premiere but not vice versa. So control Premiere from within Transcriptive, it won't work otherwise.

Load Transcripts:

If you've already gotten a transcript for that sequence, you need to select Load Transcripts from the Main Menu. This will load the transcript for the sequence you have selected. Premiere is not that great about telling Transcriptive what's going on, it's a bit like pulling Adobe's teeth. Not fun. So if you switch to a different sequence you need to tell Transcriptive that you've done so by selecting Load Transcripts. Of course, if you've never generated a transcript for that sequence, you'll need to click the Transcribe button and have it transcribed.

Accuracy:

Make sure you're using Speechmatics. It has much higher accuracy than Watson. Also, background noise and unclear speaking can make it harder for the speech services to figure out what's being said. The less noise you have and the better your talent enunciates words, the higher the accuracy will be. Make sure your talent is mic'd, speaks clearly and any background noise is minimized. Talent that has a heavy accent or is off mic will get transcribed with less accuracy (sometimes significantly less) than talent that is mic'd and well spoken. Regardless, Speechmatics is the better Speech Service, especially with recordings that have a lot of background noise. We highly recommend you use Speechmatics if you're doing captions or subtitles.

Undo/Redo:

You can also use Control+Z for Undo and Control+Shift+Z for Redo. NOTE: On the Mac it is also Control+Z as a Premiere panel can't use Command+Z as a normal application would.

Maximum Length of transcript:

On Speechmatics there is no maximum length. It'll transcribe every length of video you throw at it. Watson does have a limit. We can't upload anything larger than a 100mb file. To optimize this, for shorter sequences (<70 minutes) we use uncompressed audio. For longer sequences, we upload an mp3. This is lower quality and can be much lower quality resulting in less accurate transcripts. The lowest quality we do is 64kps. At that rate the maximum length you can transcribe is 3h30m. So that's the limit with Watson. It won't do more than that.

Closed Captions:

To create closed captions in Premiere, you need to export the transcript as SMPTE-TT and import that into Premiere. Premiere's Caption panel will see that as 608 or 708 captions.

Edit Mode vs. Tracking Mode

There are two editing modes in Transcriptive.

By default you're in Tracking Mode. If you press the spacebar to play your video, a yellow highlight will move from word to word as it's spoken. In Tracking Mode you can't edit the words, however you can delete them or add punctuation with the Punctuation Bar or keyboard shortcuts. It's a VERY fast way of doing certain edits. If you just need to add some punctuation and capitalize some words, you can do it all very quickly with keyboard shortcuts.

For example, the keyboard shortcuts work on the highlighted word. So if you highlight the word 'attention' and click the period key (.), a period will be added to 'attention' and 'so' will be capitalized. Likewise, you can select the word 'wednesday' and click the Up Arrow and it'll be capitalized. Just select the word, hit one key and keep playing back. Very fast and efficient.

However, if you need to actually edit the text, then you need to go into Edit Mode.

You enter Edit Mode by clicking <Enter> and you exit Edit Mode by clicking <ESC>.

Once you're in Edit Mode, you can edit text as if you were in a word processor. You can position the cursor anywhere you want and just start typing.

When you first enter Edit Mode, the word that was highlighted in Tracking Mode is selected. This makes it easy to replace. Just type the new word. This is very common, as the Speech Services will frequently get a word slightly wrong. For example, you might get 'lean' instead of 'keen'. So it's often fastest to enter Edit Mode, type out 'keen', exit Edit Mode and keep playing back.

The buttons on the Punctuation Bar don't work in Edit Mode.

Once you've mastered switching between Tracking Mode and Editing Mode and know the keyboard shortcuts in Tracking Mode, you can edit your transcript incredibly fast. You can go even faster if you set the Playback Speed to 1.5x or 2x (That's the button on the Bottom Bar in the lower, right corner).

Selecting Video/Audio to Transcribe

Transcribing A Sequence

If you're transcribing a sequence, Transcriptive will transcribe whatever audio is in the active sequence you have selected (we'll cover Batch processing in a moment). So if you're looking at the timeline of MySequence, that will be the audio that gets transcribed. ALL the audio. So if you have multiple tracks of audio it'll send all of it to the Speech Services.

This is great if you have dialog on a couple tracks. Both tracks will get transcribed. However if you have dialog on one track, music on another track, goats braying on another track, etc., etc.... that's not so great. Mute the music and the goats before transcribing your dialog. You'll get much better results.

In/Out Points

Transcriptive will honor IN and OUT points when transcribing. So if you want your entire transcript transcribed or just a portion of it, make sure the 'Sequence Length' matches what you're expecting. If it says '5 minutes' and your sequence is 30 minutes, there's a problem. It might be the IN/OUT points.

Batch Processing

There are multiple ways of batch processing. Let's discuss.

Ultimately they all do the same thing: process multiple clips and put the entire transcript into the Speech Analysis section of Premiere's Metadata panel for each clip. This means you can search the metadata of each clip in the Source window to find any text in the clip you are looking for. If you drop the clip into a sequence, Transcriptive will use the metadata to build the transcript, so you don't incur additional charges from the speech services. It's very cool if you have a bunch of clips you want transcribed.

When you batch process, Adobe Media Encoder gets launched so Premiere doesn't get tied up while all the clips get transcribed. If you have a lot of clips that could take a while and we didn't want Premiere to become unusable while waiting for the transcripts to get back.

Batch Files

This will bring up a dialog letting you select multiple files to be transcribed. This is more or less the same as any Open dialog you'll see.

Batch Folder

This will bring up a dialog letting you select a folder. Every video/audio file in that folder will

be transcribed.

Batch Project

If you have all the files you want transcribed imported into Premiere, create a bin named 'Transcriptive'. Place all the files you want transcribed into that bin. When you select Batch Project, Transcriptive will take all the files in that bin and transcribe them. Once that is done, you can move the clips out of the Transcriptive bin.

Once you've selected your files, you'll see the Transcribe dialog. See below for more details on that.

Transcriptive User Interface

The user interface is divided into five sections:

1- Top bar with the Menu, Transcribe button, and Status field

2- The main text area for editing the transcription

3- Speaker Identification column showing speakers (this only shows if there speakers were identified)

4- Timecode column

5- Bottom bar with the Export options

The Top Bar
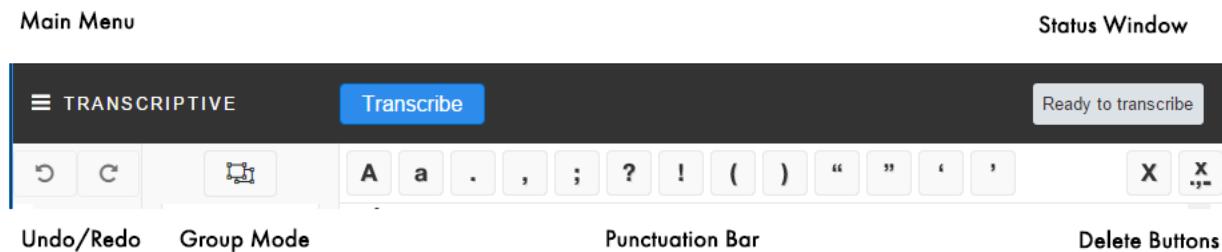
The top bar in the UI has many critical functions. The Main Menu has all the setup dialogs, Import and Load dialogs, and Batch Processing dialogs. The lower portion of the bar has Undo/Redo buttons, Group button and punctuation buttons.

Main Menu

This is your control panel of sorts. Speech Engine Setup and Preferences are located here. You can Batch Process files using the three options. You can also manage transcripts here.

New Transcripts

This simply clears out whatever text is in the Editing Window. You don't _need_ to do this. If you click the Transcribe button it'll overwrite whatever is in the Edit Window. However, it can be useful to visually see that you're creating a new transcript. So if you want a blank page before getting a transcript, just select New Transcripts.

Load Transcripts

If you switch sequences in Premiere, Premiere doesn't tell Transcriptive that you've done so. Premiere is not that great about telling Transcriptive what's going on, it's a bit like pulling Adobe's teeth. Not fun. What this means for you is that when you switch sequences you need to select Load Transcripts to load the correct transcript. Any time you switch sequences in Premiere, select Load Transcripts.

Of course, if you've never generated a transcript for that sequence, you'll need to click the Transcribe button and have it transcribed.

Import Transcript

This is a pretty important dialog and lets you get transcripts into Premiere in a number of ways.

Plain Text

This imports any plain text file into the Transcriptive window. The text will come in without timecode, so it's going to be of limited use unless you use Speechmatics to sync up the text and audio...

Existing Text File as A Transcript

If you already have a transcript for a video file and want to sync that up to the audio of the video, you can use the Plain Text importer to do that. Import the plain text file and then click on Transcribe. You'll need to select Speechmatics and select 'Align Text'. This will sync up the timecode of the spoken words with the plain text file. See Using an Existing Text File as A Transcript for more info.

SRT

If you have an existing transcript as a SRT file, you can use the SRT Importer to get that into Transcriptive. Since SRTs only have timecode for whole lines, the individual words do not have timecode.You can still search and jump to where a sentence was spoken, but you can't do so for individual words.

JSON

JSON files are created by the Speech Services when they generate the transcript. This is what Transcriptive gets when you request a transcript. Sometimes you'll submit a file to be transcribed and it'll fail for one reason or another (usually an internet hiccup), when this happens you can get the JSON file from Speechmatics or Watson and import it here.

This also lets you get the transcript on one machine and move it to another. So if your facility only has a couple machines connected to the internet, you can use them to get the

transcription, grab the JSON file and move it to the machine without an internet connection.

Speechmatics

If there's an internet failure of some sort, this makes it even easier to grab the JSON file. This shows a list of all the jobs you've submitted to Speechmatics. You can select the one you want to use for the sequence you've selected.
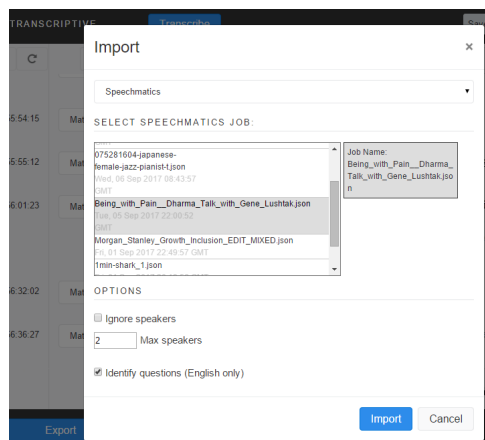
Speaker Management

By default the Speech Services return speakers as Speaker 1, Speaker 2, etc. This dialog lets you assign real names to the speakers. The Speaker dropdown that's next to each entry in the transcript will show whatever names you enter in instead of the default names.

Getting Rid of Speakers

One interesting use of this is that if you decided you don't want to worry about speakers, you can delete all the names in this and it'll remove the Speaker column in the Transcriptive panel.

TRANSCRIBE Button

The heart of transcriptive is the Transcribe button. This is where the magic begins.

Once you've selected the sequence you want to transcribe and muted any non-dialog audio tracks, you'll press the TRANSCRIBE button.

This will bring up the Transcribe dialog box. See the next page for more info about the dialog box.

Transcribe Dialog Box

Before Transcriptive sends the audio file off to the Speech Services we need to tell you a few things and we need to know a few things.

Most importantly, which Speech Service do you want to use, Speechmatics or Watson? (We may add more in the future)

Please use Speechmatics if accuracy is critical to you. It's much more accurate than Watson.

So you'll need to select one or the other. If you have not setup the Speech Services, there's a button that will take you to that dialog box. You can't use transcriptive without setting up at least one Speech Service.

What we tell you

First off, we tell you how long the audio is that you're transcribing. You should check this to make sure that you're not transcribing 3 hours of video when all you really want is 15 minutes.

It'll then give you some info about your account for whichever speech service you have

selected. For Speechmatics, we also tell you the system status. If this is Good, then you should expect to get your transcription back quickly. If it's something other than Good, it will probably be delayed.

Options

Languages: Speechmatics supports 18 languages and Watson supports 7 (currently). Obviously Speechmatics supports languages that Watson doesn't, but Watson also supports a couple unique languages like Mandarin. Select the language of your video and away you go. Neither service supports multiple languages in a single video.

Speaker Identification

Next, select if you want Speaker Identification and tell us how many speakers there are. See the Speaker Identification section for more details on this. Both speech services do a lousy job of identifying speakers. It's better than nothing, but it definitely requires a lot of clean up.

Identify Questions

You can select if Transcriptive will try to identify questions. This also doesn't work great. However, if you know you have an interview or something where there were lots of questions, you should turn this on. Again, it's better than nothing for interviews. For clips with few questions, you should leave this off as it'll give you more false positives than it's worth.

Align Current Text

If you're getting a new transcript from one of the speech services, make sure this is turned off. See the section on Importing An Existing Transcript. Basically this allows you to take an existing transcript and sync it to the audio, adding timecode to the existing text. Incredibly useful if you already have a transcript from a human transcription service but useless otherwise. You MUST import an existing text file for this to work.

When you click OK and Transcriptive exports the audio and sends it to the Speech Services.

Transcription Status Field

The field in the upper, right corner will tell you the status of your transcription and if it succeeds or fails.
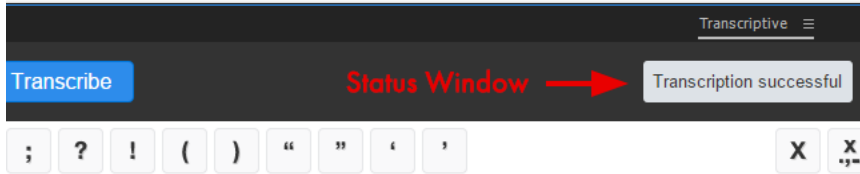

The messages you most commonly will see are:

Encoding FLAC file: This means we're exporting the audio as a lossless FLAC file, which is what the Speech Services require.

Waiting for Speechmatics/Watson: Once the audio file is sent, we wait for the Speech Services to respond either with the transcription text file or an error message. Usually this will take about 1/4th to 1/2 the time of the audio file. So if you send a 10 minute audio clip, expect it to be done in 2-5 minutes.

Transcription Successful: Transcriptive has received the text file and you should see the transcription in the Transcriptive panel.

Transcription Failed: This can happen for a number of reasons, including the service being down. We try to provide helpful error messages but the Speech Services aren't always helpful in this regard, so sometimes we don't know why it failed.

Undo/Redo

↺ ↻

00:54:19:15
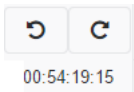
These buttons allow you to undo or redo the last action. You can also use Control+Z for Undo and Control+Shift+Z for Redo. NOTE: On the Mac it is also Control+Z as a panel can't use Command+Z.

Undo/Redo doesn't work for everything. You can't undo merging and splitting sentences, for example. Eventually we'd like to have everything undoable, but for now it's just text changes, which should over most peoples needs.

Group/Sentence View
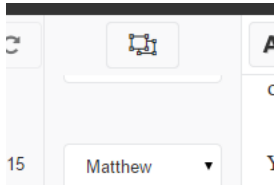
By default when you get a transcript back each sentence has a separate line in the Edit Window. If Speaker ID was turned on, the speaker is listed for each sentence. This is great if you want to see the timecode for each line or want to edit the speakers and make sure they're correct.
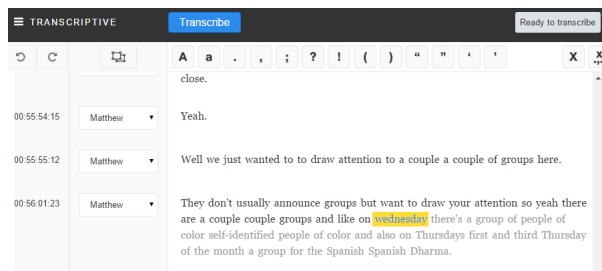
Sentence View

However, having each sentence as a separate line does take up a lot of real estate. If you just want to go through and edit/check the text it's usually easier if the text is in one big block as it would be in a regular word processor. Luckily the Group button lets you do this!

When you click on the group button it will merge all the text together temporarily so you can edit it more easily. If there are speakers, it'll group the sentences together by speaker. If there aren't speaker, it'll create blocks of text with about 1000 words each.
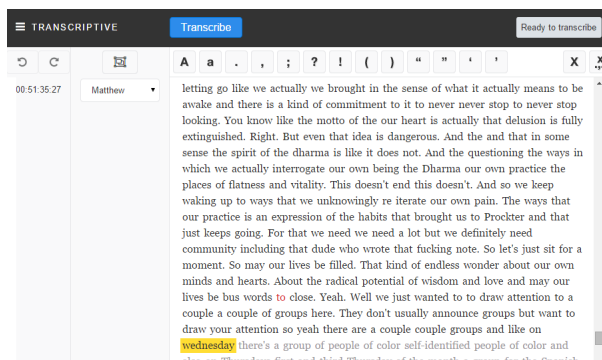
Group View

You can click the Group button again to go back to Sentence view and Transcriptive will retain all the sentence information.

Punctuation Bar

These buttons let you easily add punctuation in Tracking mode. Just select the word and hit the button or corresponding keyboard shortcut (the punctuation key on the keyboard, for example to add quotes, select the word and press the " key). This makes it very easy to add basic punctuation.

The great thing about this is you don't have to go into Edit mode. Normally you'd have to enter Edit mode, position the cursor at the beginning of the word, press the " key, move the cursor to the end of the word and press the " key again. With Transcriptive, as you're playing back the audio, you can stop the highlight on the word you want to add quotes to and press the " key, then press Control+" and you're done! Two keystrokes vs. six keystrokes.

These buttons don't work in Edit mode.

Text Edit Window

As you might guess, this is where you can edit the transcription. It also highlights the text as the video plays, allowing you to easily follow the text as you listen to the audio playback.

You can edit with keyboard shortcuts as the video plays or you can go into 'word processor' mode and edit as you usually would with a text editor.

We cover all this in more detail in the Editing Text section. Also check out the Keyboard Shortcuts section.

Speaker Identification Column

If you selected 'Speaker identification' when you transcribed, the Speech Services will attempt to identify how many speakers you have and which text is attributable to each one. Neither Speechmatics or Watson are awesome at this, and this typically needs a fair amount of cleanup.

The Speaker Column shows which speaker the Speech Service thinks said that particular sentence.

You can change the speaker by clicking on the dropdown menu and selecting the correct speaker.

By default the column will show Speaker 1, Speaker 2, etc. You can change that to real names by going to the Speaker Management dialog in the main menu.

Timecode Column

This shows the start timecode of any given sentence or block of text. In most cases, each word will have timecode associated with it. As you click on each word, the Time Marker in Premiere will jump to the portion of the timeline where it's spoken. The Timecode Column only displays the timecode for the first word.

If you split (Opt/Alt+Down Arrow) a sentence or block of text, the timecode of the word you split on will appear in the Timecode column.

This column is primarily for informational purposes only.

Bottom Nav Bar
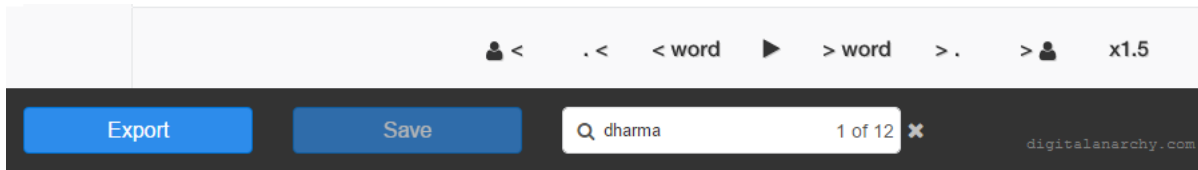
The lower navigation bar has controls to move around the text, Search toolbar and the Export dialog. All critical parts of Transcriptive.

Playback And Cursor Controls

The controls right below the text window allow you to move where the highlight is. Especially when you're using the keyboard shortcuts, it's fast and easy to highlight any word. The controls are:

- to the next/previous word (keyboard shortcut: right/left arrow)

- to the next/prev sentences (keyboard shortcut: Alt/Opt + right/left arrow)

- to the next/prev speaker (keyboard shortcut: Alt/Opt + Shift + right/left arrow)

- Play: Starts playback. The video/audio will play while the highlighted word moves along with whatever is spoken. (keyboard shortcut: Spacebar)

- Playback Speed: Lets you set the playback speed to realtime or 1.5x faster than realtime. This is very useful, especially if the transcript is accurate. You can scan for incorrect words much faster. It takes a little getting used to, but once you do it'll speed up your cleanup.

Export

There's a lot going on in this little dialog. You can export the entire transcript to a closed caption file, export it to Premiere's metadata or markers, or export a plain text file for paper edits.

Closed Captions and Subtitles (SRT, VTT, STL, SMPTE-TT)

All these file types work similarly and have similar options. The only exception is SMPTE-TT which is what you should use for closed captions and has a couple additional options because of this.

When you select one of these file types you get these options:

SRT, VTT, STL

These are all text files that meet different standards. They are usually used for subtitles or Open Captions. They usually can't be used for Closed Captions without being converted to something else.

All three formats have exactly the same options:

Include Speakers: This will include the speaker names in the exported file. So you'll get something like this, which will show the speaker name on screen:

00:00:08.710 --> 00:00:09.669
<v Jim>You're talking about her.

Starting Timecode: If you want to have different timecode than what's shown in the Sequence window, you can use this to set the start time to whatever you want.

Maximum character length: This sets the maximum number of characters a line of text can have. It defaults to 32 which is the max for closed captions. You can set this to any value.

Number of lines on the screen: This sets how many lines of subtitles you want on the screen at any one time. The maximum is 3 and if it's set to that you'll see three lines on the screen when the subtitles are viewed.

Keep subtitles on screen: This is pretty important. It sets the amount of time between when one subtitle is turned off and another is displayed. This prevents a short sentence (i.e. "Oh, yeah.") from flashing on and off the screen. It'll be displayed until it's 1/4th of a second away from another subtitle. 1/4th of a second is the default.

SMPTE-TT (XML)

SMPTE-TT is an XML file and is what you want to use if you're doing closed captions. **It's the only format that Premiere Pro's Caption Tool supports for closed captions.** To use it with the Premiere's too, you'll need to Export the SMPTE file and then re-import it. When you drop onto the Premiere timeline you can then use the Caption panel to manipulate the captions. It has most of the same options as the other three file formats, but there are a couple differences:

Captions Standard: You can specify whether the captions meet the 608 or 708 standard.

Caption Type: Let's you set how many lines you want and how those lines behave... either pop-up (one caption is replaced by another) or roll-on (the captions scroll up).

Plain Text

This saves the transcript out as a tab-delimited text file. You can then open this up in Word or Excel to print it out. It's great for giving someone a printout for paper cuts.

Speech Analysis

Exporting to Speech Analysis puts the entire transcript into the Speech Analysis field that's visible in Premiere's Metadata panel.
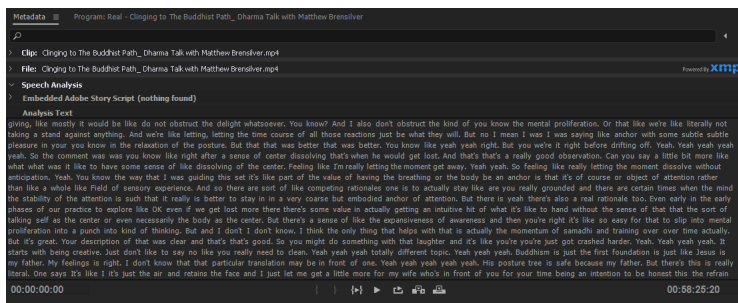
This is just another way of searching the text to find a word or phrase.

To be honest, if you've transcribed a sequence, you're better off searching in the Transcriptive panel. For sequences, the metadata panel is not that useful.

However, FOR VIDEO CLIPS this is incredibly useful. When you batch process clips, the transcript gets put into this field. If you're trying to find something in the Source window for a clip, you can search in the Metadata panel, find what you want and then drag that into your sequence. It makes searching long video clips very easy.

Sequence Markers

This exports the transcript to sequence markers. You can then search the text via the Markers panel. This can be very useful if you want to close the Transcriptive panel or if you need to send the project to someone that doesn't own Transcriptive. The entire transcript will be embedded into sequence markers which will show up in both the timeline and Program window.
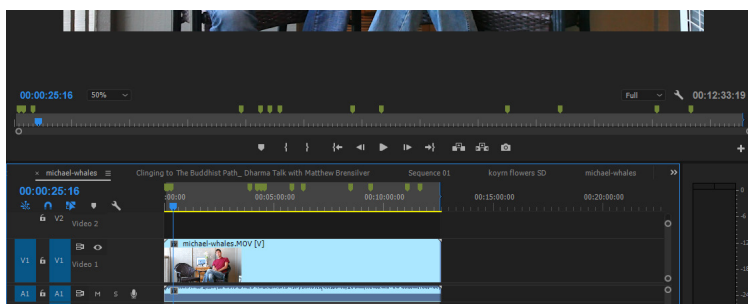
When you export you get two options:

Include Speakers: This will include the speaker names in the text of the markers.

Seconds Between Markers: This sets the frequency of the markers. If you set this to 10 seconds, a marker will be laid down about every 10 seconds. Each marker will include the text that STARTED being spoken in the 10 seconds after the marker (to use the above example). So if you have a sentence that starts at 10:05, another one at 17:15 and a third at 19:20, those three sentences will be included in the marker at 10:00.

But, you know... it's complicated.

If the sentence that starts at 19:20 is long and the next sentence is at 31:10, no marker will be laid down at 20:00 as no sentence starts between 20:00 and 30:00. The next marker will be at 30:00.

... because size matters.

When you get a transcript back, it should be broken down by sentences. Each entry has a single sentence. However, it is possible to merge sentences into a block of text. So it's possible for one entry to have, say, 2 minutes worth of text. In this case, you'll have a marker at 10:00 with two minutes worth of text and the next marker won't get laid down until the 2:00:20 mark.

One sentence per entry			Merged sentences with multiple sentences per entry		
00:00:26:17	Michael	Each frame will be a new focus.	00:00:08:21	Jim	You're talking about continuous focus.
00:00:28:17	Michael	And generally if you have a whale jumping as the whale jumps that first shot no matter how good you are it's	00:00:12:12	Michael	That's very important.
00:00:36:03	Michael	Maybe the second shot will be almost perfectly in focus and the third depending on how how many feet the	00:00:14:29	Michael	AI Servo for cameras when you have. The camera can continuously focus in between each frame. So even if your camera shoots at frames a second, each frame will be a new focus. And generally if you have a whale jumping as the whale jumps that first shot no are it's not going to be perfectly in focus. Maybe the second shot will be almost perfectly in focus and the third depending on how second your camera shoots. The more the better but as it shoots each each shot gets more and more in focus. So shooting.
00:00:42:16	Michael	The more the better.			
00:00:43:23	Michael	But.	00:00:49:00	Michael	You actually have to do both you have to have. Continuous focus and continuous shooting both. AI Servo and continuous shooting.

So... generally markers will be laid down according to the time you set in Seconds

Clip Markers

Clip Markers can be an awesome way of moving the transcript around. Clip Markers are embedded into the video file, so if you import the clip into a different project or move it to a different machine, the clip markers (and the transcript) will go with it.

You'll see these show up ON the clip itself in the timeline. Sequence markers show above the timeline.

The options are the same as Sequence Markers (see above) and work the same way.

JSON

These files are mostly for moving the transcript to a different machine with transcriptive. You can just move the entire Premiere project to the other machine. However if you just want to get the transcription over to the other machine, you can export out as JSON file and import it on the other end.

Saving The File

Once you've figured out what file type you want and all the settings, you can click Export. This will bring up a file dialog asking you where you want to save. Select the location on your hard drive and click Save.

Text Edit Window

(For a visual explanation, check out this 8 minute tutorial: https://youtu.be/_-KajNnQkPU)

We've already talked a bit about Edit Mode vs. Tracking Mode, but to recap...

The overall goal with this is that you don't have to take your hands off the keyboard. Usually you can do everything with keyboard shortcuts. There are exceptions but most of the time, if you know the keyboard shortcuts, your hands never have to leave the keyboard. This makes for fast, efficient editing.

There are two editing modes in Transcriptive.

By default you're in Tracking Mode. If you press the spacebar to play your video, a yellow highlight will move from word to word as it's spoken. In Tracking Mode you can't edit the words, however you can delete them or add punctuation with the Punctuation Bar or keyboard shortcuts. Keyboard shortcuts are critical in Tracking Mode. Learn them.

To actually edit the text you need to switch to Edit Mode.

In Edit Mode you can edit text as if you were in a word processor. You can position the cursor anywhere you want and just start typing or select text and replace it. Although, again, hopefully you're just changing a word here and there and don't need to move the cursor.

You enter Edit Mode by clicking <Enter> and you exit Edit Mode by clicking <ESC>.

Tracking and Edit Mode are described in more detail on the next page.

Tracking Mode

It's a VERY fast way of doing certain edits. If you just need to add some punctuation and capitalize some words, you can do it all very quickly with keyboard shortcuts in Tracking Mode

This is designed to let you easily playback the video and see any transcription errors quickly. If the transcript is mostly correct you'll be able to clean it up very quickly. This is why we recommend Speechmatics if you're going to be using the transcript for captions or subtitles where you need a perfect transcript. Speechmatics is more accurate than Watson and clean up will go much faster. The time you save on clean up will usually offset the extra cost. The more you have to go in and out of Edit Mode, the longer it'll take.

If you can do most of your editing in Tracking Mode, you'll be able to move incredibly fast.

For example, the keyboard shortcuts work on the highlighted word. So if you highlight the word 'attention' and click the period key (.), a period will be added to 'attention' and 'so' will be capitalized. Likewise, you can select the word 'wednesday' and click the Up Arrow and it'll be capitalized. Just select the word, hit one key and keep playing back. Very fast and efficient.

Here are the keyboard shortcuts:

Navigation

Play/Stop: Spacebar
Next Word: Right Arrow
Previous Word: Left Arrow
Next Sentence: Alt/Option+Right Arrow
Previous Sent.:   Alt/Opt+Left Arrow
Next Speaker: Alt/Opt+Shift+Right Arrow
Previous Sp.:   Alt/Opt+Shift+Left Arrow

Capitalization

Make Upper case: Up Arrow

Make Lower case: Down Arrow

Line Merging/Splitting

Merge with line above: Alt/Option + Up Arrow

Split sentence and make new line: Alt/Opt + Down Arrow

Punctuation Shortcuts:

Period: .   Left Single Quote: '
Comma: ,   Right Single Quote: Command/Control + '
Question Mark: ?   Left Double Quote: Shift + "
Exclamation Point: !   Right Double Quote: Cmd/Cntrl + Shift + "
Colon: :   Semi-Colon: ;
Left Parenthesis: (Right Parenthesis:)

Delete Word: Backspace/Delete   Delete Punctuation: Cmd/Cntrl+Backspace

Edit Mode

However, if you need to actually edit the text, then you need to go into Edit Mode.

You enter Edit Mode by clicking <Enter> and you exit Edit Mode by clicking <ESC>.

Once you're in Edit Mode, you can edit text as if you were in a word processor. You can position the cursor anywhere you want and just start typing. Although, again, hopefully you're just changing a word here and there and don't need to move the cursor.

When you first enter Edit Mode, the word that was highlighted in Tracking Mode is selected. This makes it easy to replace. Just type the new word. This is very common, as the Speech Services will frequently get a word slightly wrong. For example, you might get 'loan' instead of 'alone'. So it's often fastest to enter Edit Mode, type out 'alone', exit Edit Mode and keep playing back.

Also, since you enter Edit Mode with the word selected, you can use the arrow keys to go to the beginning or end of the word. For example, if the speech service returned 'engage', you can click the left arrow to go to the front of it and add 'dis' (for 'disengage') and then hit ESC. You could click the right arrow to go to the end of it and type 'd' (for 'engaged') and then hit ESC.

So it can be very fast for making quick edits. But it is a full word processor so if you need to change multiple words or you see something further down that needs to be corrected, you can move the cursor wherever you need to make those changes.

The buttons on the Punctuation Bar don't work in Edit Mode.

Playback Speed Button

Once you've mastered switching between Tracking Mode and Editing Mode and know the keyboard shortcuts in Tracking Mode, you can edit your transcript incredibly fast. You can go even faster if you set the Playback Speed to 1.5x or 2x (That's the button on the Bottom Bar in the lower, right corner). It's surprisingly easy to listen to a video that's been sped up by 150% or 200%. It takes a little getting used to, but once you do, you can fly through the editing process especially if the transcript is very accurate to begin with.

Search

Search is one of the revolutionary things about Transcriptive. Finding where a word or phrase is spoken in your video is as simple as Googling something on the web. Just type it in and you jump to wherever on the timeline it's spoken. it makes it simple and fast to find the right moment for your documentary, news program or any time you have hours of footage.
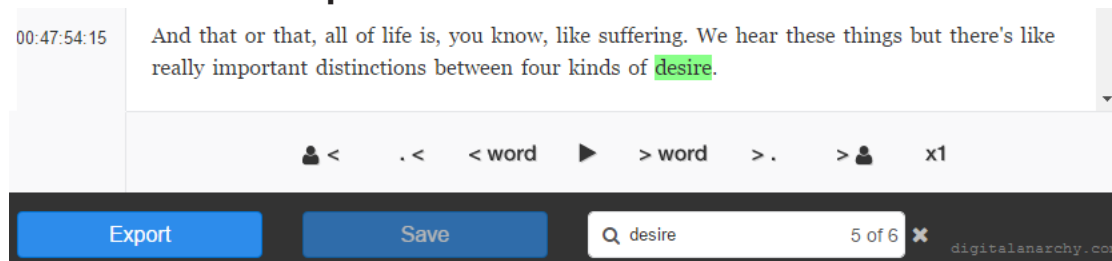
There are three ways to search the transcript to find where a word or phrase is spoken.

The Transcriptive Panel
The Metadata Panel (Speech Analysis section)
The Markers Panel (if you export to markers)

Search in The Transcriptive Panel

If you have the entire transcript in the Transcriptive Panel, this is the easiest way to search. Just type in the word or phrase you're looking for and it'll highlight it. You'll see how many results Transcriptive found in the Search field.

In 1.0 you need to scroll down the panel to find the search results. 1.0.1, which will be released by the end of Sep 2017, will let you tab or move through the search results with arrow keys (similar to a browser).

When you find the word you're looking for, just click on it and the time marker will jump to the point on the timeline where it's spoken.

Every word in Transcriptive has time code associated with it, so we know where in the video it's spoken. The only exception to this is if you import a text file like an SRT file. In that case we may just have the timecode for each sentence or possibly not at all. In that case you'll want to use the Align Text function to get timecode for the word. (See Syncing an Existing Transcript with Your Audio)

Search The Metadata Panel (Speech Analysis)
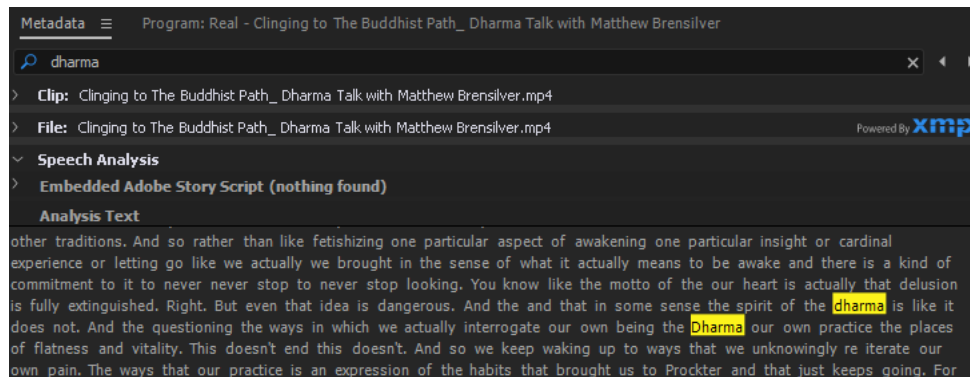
Speech Analysis for Video Clips

This is most useful if you've batch processed a bunch of video clips and want to find something before they're in a sequence. For example, finding a specific part of an interview in the Source window so you can grab that segment and drag it into a sequence.

To search the metadata panel the transcript needs to be in the Speech Analysis field.

The most effective way of doing this is to batch process a number of clips. When video clips (or audio clips!) are batch processed the transcript is placed in the metadata which shows up in the Speech Analysis field of the Metadata Panel in Premiere. You can then use the search field in the Metadata Panel to find what you're looking.

Searching The Source Panel

When searching in the Metadata panel, you can select text there... a sentence or block of text... and in and out points will be set in the Source Panel where that block of text starts and ends. This makes it easy to rough cut together a bunch of clips or segments. Just drag the selected portion of the clip into your sequence like you normally would.

It's MUCH easier than scrubbing through an hour long clip looking for that point where someone said something. You can now essentially google your video.

The other cool thing about this is that you the transcript will now go along with the clip. If you import it into a project other than the one you batch transcribed it in, the transcript will be there.

Speech Analysis for Sequences

While you can send the transcript to the Speech Analysis section for a sequence (instead of a video clip), this isn't very useful. We recommend searching in the Transcriptive Panel. But if you're moving the

sequence to a machine that does not have Transcriptive installed, it's one way of getting the transcript embedded in the Premiere project.

Search The Markers Panel
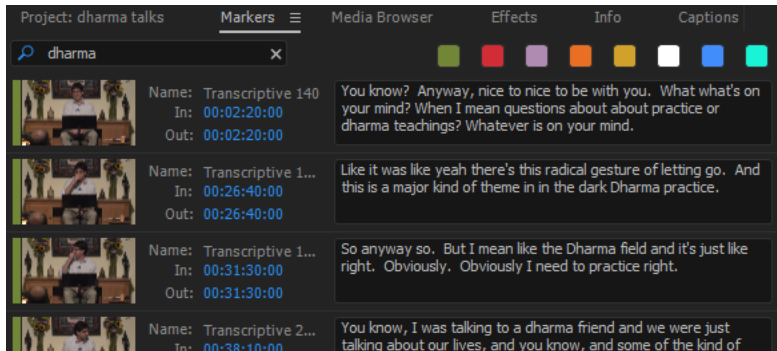
You can also export the entire transcript to Premiere markers. (See the Export section on how this works)

This is great if you want to search markers instead of the Transcriptive Panel. The Transcriptive Panel does take up a lot of real estate on screen, so sometimes searching markers fits better into a given workflow.

The markers also get embedded into the Premiere project. So if you're moving the project to a machine without Transcriptive, this is one way of getting all the transcript data over to the other machine.

The downside to this method is that the search result is approximate. Most markers will have 5 or 10 seconds worth of the transcript. So you'll still need to scrub through a few seconds to find the exact word or phrase you were looking for.

Working with Closed Captions

With Transcriptive 1.0, the best way to export Closed Captions out of the panel is to export as a SMPTE-TT file. You can specify 608 or 708 closed captions as well as a number of other features.

This is the only format that Premiere's Caption Panel will recognize as closed captions. Most other captioning applications should work with this as well.

We do plan on supporting .scc files in the near future and hopefully Adobe will let us save it out directly as a Premiere caption file. However, for now, SMPTE-TT is the way to go if you want closed captions.

Using SMPTE-TT with Premiere's Caption Panel

Once you've exported a SMPTE-TT file out of Transcriptive, you need to import it into Premiere. It's just a XML file and Premiere should recognize it as a caption file.

Once you've imported it Premiere, you'll need to drag it into the sequence you want captions for. Just put it on any video track.

If the talent doesn't start speaking at 00:00 you'll probably need to align the beginning of the SMPTE file with where the talent does start speaking. Otherwise the captions will be out of sync.

You'll also need to go to the Caption settings in the Program panel and tell Premiere what kind of captions to expect, 608 or 708. (and you'll need to enable captions if they aren't already)
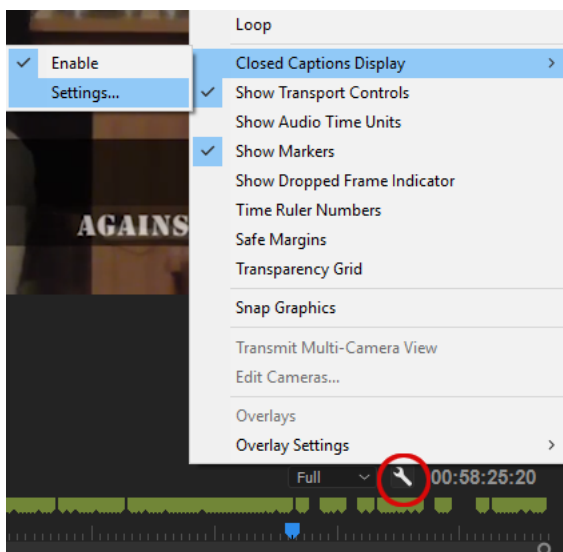
At this point the captions should be showing up on screen.

You can then open the Caption panel in Premiere and start formatting them and changing the position. Premiere's caption tool is a little limited but you can set the position and some basic formating for each caption or all of them.
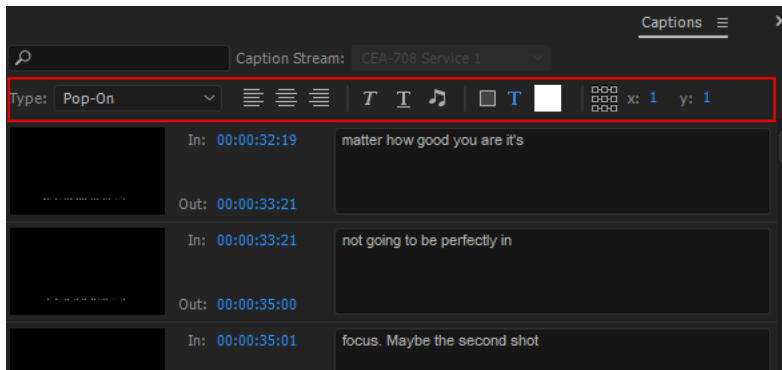
One thing to remember when formatting the captions... You'll want to make sure their all selected when you make a change. Control + click on one of the captions to bring up a 'Select All' pop-up. This will let you select all the captions and make changes at the same time.

Syncing An Existing Transcript to Audio (adding timecode to a text file)

One of the most important things Transcriptive can do is take your existing script or transcript and sync it up with the speech in your video. This will add timecode to every word in the text file.

So if you already have a transcript, either the written script, a transcript from a transcription service or a caption file created elsewhere, you can easily add the correct timecode to it. Given the existing transcript is likely to be more accurate than the one generated from scratch, you should definitely use this option if at all possible.

Importing A Text File

Under the Main Menu is the Import Transcript command. Select that and choose the file type of your transcript. In most cases it'll be plain text but it could be SRT, VTT or one of the other formats. See the Import section for more details on each of those.

Aligning The Text with Timecode (Syncing it with the speech in your video)

Once the text is in the Transcriptive panel, you need to align it with your video's audio. To do this:

- Click the Transcribe button
- Select Speechmatics (only Speechmatics can do this, it's not an option for Watson)
- Check the Align Text button at the bottom of the Transcribe dialog
- Click OK

This will behave similarly to a normal transcription session: Transcriptive will encode the audio, upload it to Speechmatics, and get back a file from Speechmatics that we can use to add timecode to your text.

However, there are a few differences:

1- You're charged less. It costs 4p instead of 6p/min... or if you're not in the UK... 5 cents/ minute vs. 7 cents/minute. So you'll see the number of minutes you have left increase. This number is only correct if everything you do from that point out is aligning text. If you try and transcribe something normally, the number of minutes you have left will decrease dramatically.

2- It takes longer. It seems to take significantly longer for Speechmatics to analyze it and send back timecode than if it's just transcribing the audio from scratch. We're not sure why this is.

Troubleshooting section

How to Get The Best Results from Transcriptive

Speechmatics will give you the most accurate transcription. So if you're doing captions and subtitles and will be cleaning up the text, Speechmatics will save you a lot of time.

If you're just using the transcript to allow you to search your video, Watson is still quite accurate and a lot less expensive. For search, it's a good, economical solution especially if you have a lot of video footage.

Recording Your Audio

The higher the quality of your audio, the better the accuracy of the transcript. If someone is well mic'd in a sound proof room, you're going to get better results than if you're using the microphone on your iPhone on a busy, New York street corner.

The speech services do an amazing job but they can't overcome garbage audio. Garbage in, garbage out.

This doesn't mean it has to be perfect.

In fact, it's rarely perfect and the transcriptions are still usually excellent (especially from Speechmatics). However, the closer you follow the below guidelines, the better chance you have of getting a highly accurate transcript.

The Microphone

A good microphone goes a long way. So if you have your talent speaking into a good mic, even if you're on a street corner (maybe not the busiest New York street corner ever, let's not push it), the transcript should still come out excellent.

Our favorite microphone is the Behringer Ultravoice Xm8500 Dynamic Cardioid Vocal Microphone. You can get it for $20 and it's great. Omni-directional mics pick up more noise. They can be great, the Sennheiser ME2 lav mic we have is awesome, but good ones are more expensive and, regardless, they pick up more noise.

Background Noise

The less background noise the better. The quieter you can get the space you're recording in to be, the better off you'll be. However, all background noise is not created equal. White noise (waves crashing, cars passing) is less problematic than being in a room with lots of people talking.

Noise Reduction

Do NOT use noise reduction software on your audio. It will make the transcription worse. The AIs are trained with normal recordings, so the noise reduction software distorts what they're expecting and the resulting speech-to-text is worse. (Machine learning works by giving an AI test data that it can learn with. It will do a better job with files that are similar to the files it learned on. So if it's been trained with recordings with normal noise, it'll do less well on recordings that have been manipulated to not have noise.)

Talent That Can Enunciate

The speech services can handle accents. It's really not a problem if you have a British, Indian or American accent.

Enunciation on the other hand… that matters. If you're interviewing a drunk southern boy, the transcription will probably fare worse than your interview of that English barrister.

Try to get your talent to speak as clearly as possible. Again, it doesn't have to be perfect. If they speak a bit fast or use a little slang, it should still result in a good transcription. But the further away you get from clear, well enunciated language, the worse the transcription is going to be.

Watson Fails

Transcriptive
Transforming The Art of Transcription