

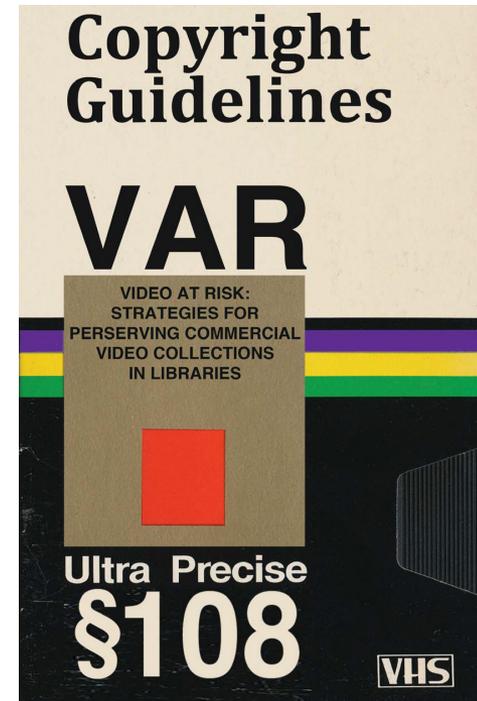
Handling Closed Captions in NYU's Circulating Video Collection

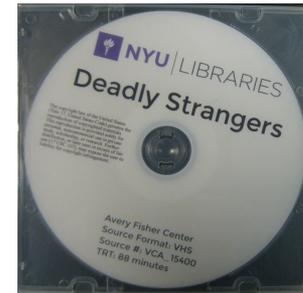
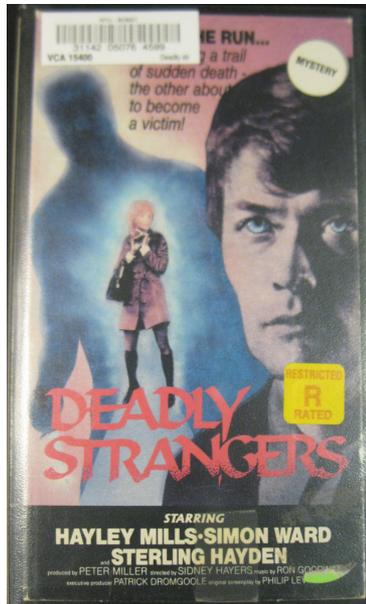
§108 and What We Do with Captions

Michael Grant
New York University Libraries
November 11, 2016

§108 Reformatting at NYU Libraries

- *Video at Risk* project presented the case for digital reformatting of commercially unavailable VHS tapes as eligible for preservation under §108 of the Copyright Act.
- Barbara Goldsmith Preservation & Conservation Department begins dedicated §108 reformatting in October 2015.

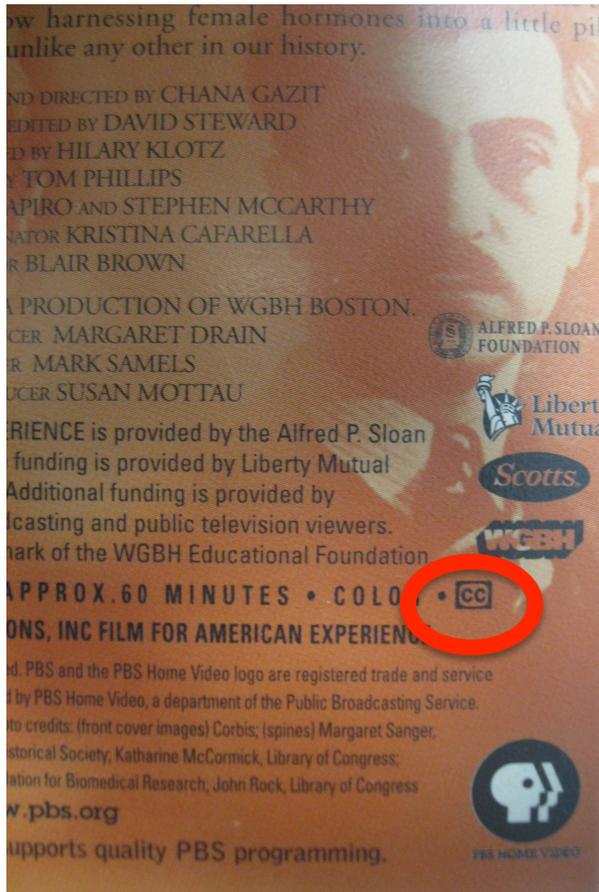




For each tape digitized we create a:

- **Preservation Master:** 10-bit uncompressed YUV QuickTime
- **Production Master / “Mezzanine”:** DV50 (50 mbps)
- **Access Copy:** DVD

But What About Tapes with Closed Captions?



The Pill (2003)

How Can We Save the Picture and the Captions and Maintain Their “Closedness”?



Miracle on 34th Street (1947, colorized 1985)



?



I BELIEVE, I BELIEVE.
IT'S SILLY, BUT I BELIEVE.

Well, How Do the Captions Get On Your Screen in the First Place?

Stretched 3x vertically for visibility!

942c 9420 942c 942f 9420 94d0 c845 a7d3 20d0 d5d3 c845 c420 4f46 4680 9470 4fd6 4552 2054 c845 20cd 4fd5 ce54 c149 ced3 ae80

- Encoded in hex pairs similar to ASCII, but with additional pairs designated for special characters, color, formatting, and spacing commands.
- Embedded in line 21 of the NTSC video signal.
- Signal dictates a caption's placement, format, and text.
- A caption remains onscreen until the decoder receives one of several signals indicating to terminate the caption.



Broken Glass (1996)

Well, How Do the Captions Get On Your Screen in the First Place?

Stretched 3x vertically for visibility!

942c 9420 942c 942f 9420 94d0 c845 a7d3 20d0 d5d3 c845 c420 4f46 4680 9470 4fd6 4552 2054 c845 20cd 4fd5 ce54 c149 ced3 ae80

- Encoded in hex pairs similar to ASCII, but with additional pairs designated for special characters, color, formatting, and spacing commands.
- Embedded in line 21 of the NTSC video signal.
- Signal dictates a caption's placement, format, and text.
- A caption remains onscreen until the decoder receives one of several signals indicating to terminate the caption.



Broken Glass (1996)

Well, How Do the Captions Get On Your Screen in the First Place?

Stretched 3x vertically for visibility!

942c 9420 942c 942f 9420 94d0 c845 a7d3 20d0 d5d3 c845 c420 4f46 4680 9470 4fd6 4552 2054 c845 20cd 4fd5 ce54 c149 ced3 ae80

- Encoded in hex pairs similar to ASCII, but with additional pairs designated for special characters, color, formatting, and spacing commands.
- Embedded in line 21 of the NTSC video signal.
- Signal dictates a caption's placement, format, and text.
- A caption remains onscreen until the decoder receives one of several signals indicating to terminate the caption.



Broken Glass (1996)

Well, How Do the Captions Get On Your Screen in the First Place?

Stretched 3x vertically for visibility!

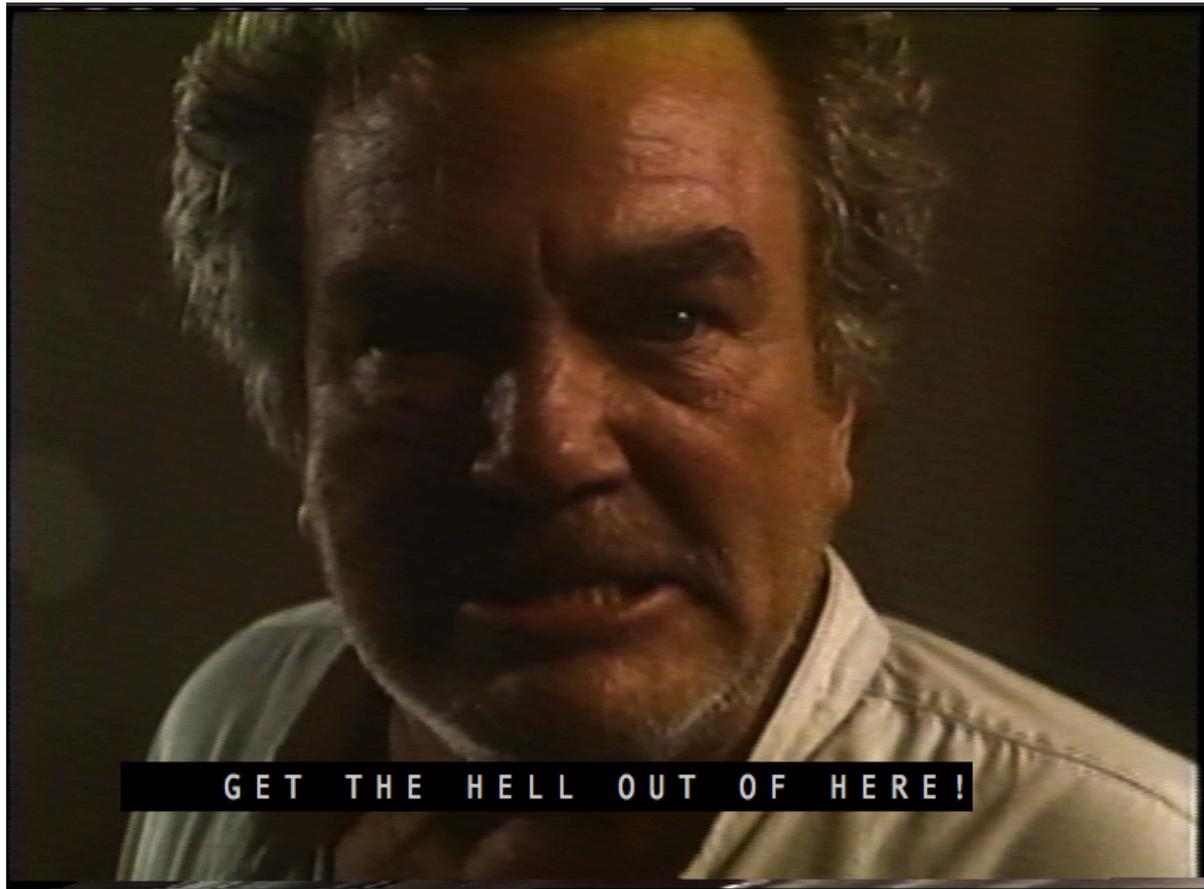
942c 9420 942c 942f 9420 94d0 c845 a7d3 20d0 d5d3 c845 c420 4f46 4680 9470 4fd6 4552 2054 c845 20cd 4fd5 ce54 c149 ced3 ae80

- Encoded in hex pairs similar to ASCII, but with additional pairs designated for special characters, color, formatting, and spacing commands.
- Embedded in line 21 of the NTSC video signal.
- Signal dictates a caption's placement, format, and text.
- A caption remains onscreen until the decoder receives one of several signals indicating to terminate the caption.



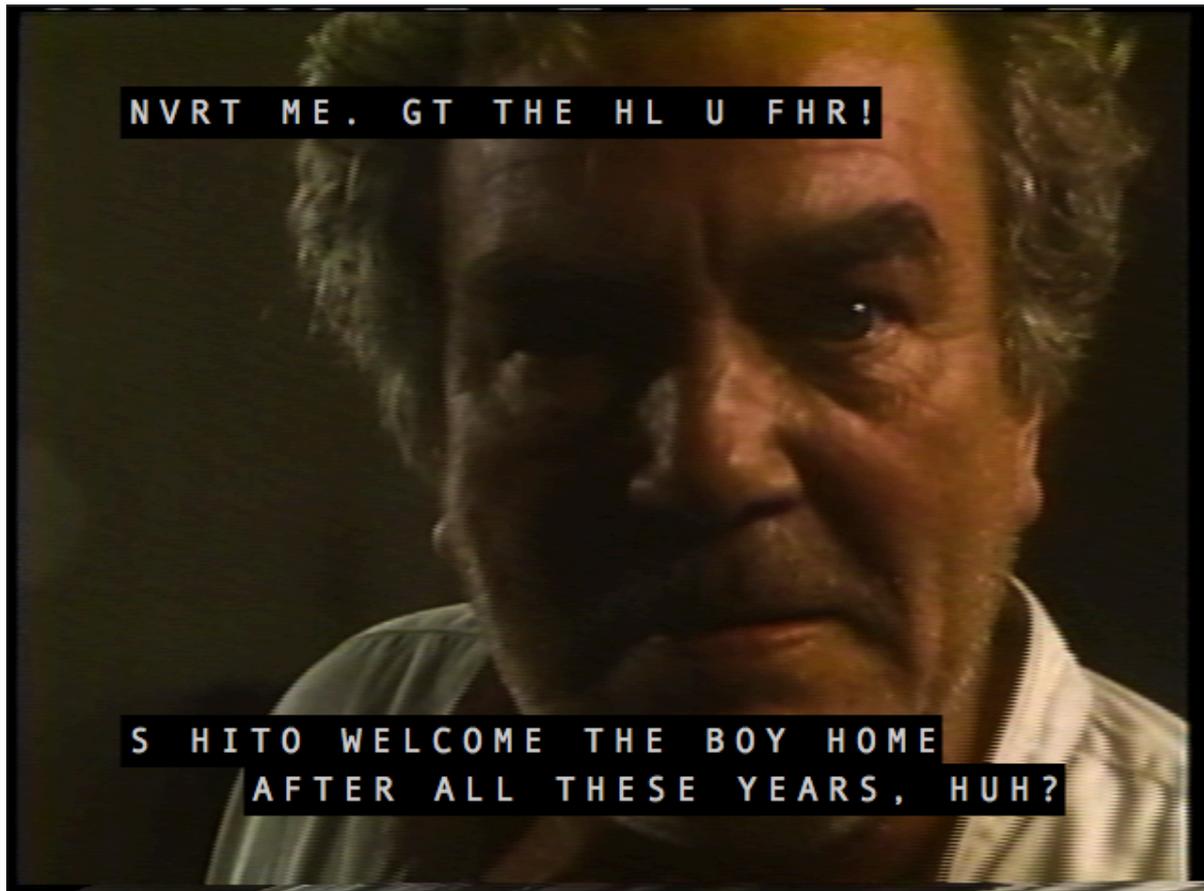
Broken Glass (1996)

Uncompressed digitization captures line 21, so the embedded caption data is preserved with no special effort.



Nostromo (1996)

But as soon as any compression is performed,
things start to get messy...



Nostromo (1996)

Converting from Line 21:

942c 9420 942c 942f 9420 94d0 c845 a7d3 20d0 d5d3 c845 c420 4f46 4680 9470 4fd6 4552 2054 c845 20cd 4fd5
ce54 c149 ced3 ae80
942c 9420 942c 942f 9420 94f8 57c8 d9bf
942c 9420 942c 942f 9420 9454 c245 43c1 d5d3 4520 c845 a7d3 20c1 2052 45c1 4c80 94f4 d9c1 cecb 4545 20c7
4fad c745 5454 4552 ae80
942c 9420 942c 942f 9420 9452 c845 20d3 c149 c420 544f 20cd 452c 20a2 43c1 d054 c149 ce2c 94f2 49a7 cd20
c845 c1c4 49ce c720 464f 5220 54c8 4520 43c1 d049 54c1 4c80
942c 9420 942c 942f 9420 9454 a254 4f20 c8c1 d645 20c1 2057 4f52 c480 94f4 5749 54c8 2054 c845 20d0 5245
d349 c445 ce54

to a Timed Script:

```
<p begin="00:15:28:10" end="00:15:29:17" region="pop1" style="basic" tts:origin="10% 84.67%"  
tts:extent="60% 5.33%">I COULDN'T DISSUADE HIM.</p>  
<p begin="00:15:29:19" end="00:15:32:01" region="pop1" style="basic" tts:origin="10% 79.33%"  
tts:extent="50% 5.33%">I TOLD HIM THERE ARE</p>  
<p begin="00:15:29:19" end="00:15:32:01" region="pop2" style="basic" tts:origin="10% 84.67%"  
tts:extent="65% 5.33%">ONLY MULE TRACKS UP THERE.</p>  
<p begin="00:15:32:03" end="00:15:34:27" region="pop1" style="basic" tts:origin="10% 79.33%"  
tts:extent="52.5% 5.33%">NOBODY'S EVER DONE IT</p>
```



Converting from Line 21:

942c 9420 942c 942f 9420 94d0 c845 a7d3 20d0 d5d3 c845 c420 4f46 4680 9470 4fd6 4552 2054 c845 20cd 4fd5
ce54 c149 ced3 ae80
942c 9420 942c 942f 9420 94f8 57c8 d9bf
942c 9420 942c 942f 9420 9454 c245 43c1 d5d3 4520 c845 a7d3 20c1 2052 45c1 4c80 94f4 d9c1 cecb 4545 20c7
4fad c745 5454 4552 ae80
942c 9420 942c 942f 9420 9452 c845 20d3 c149 c420 544f 20cd 452c 20a2 43c1 d054 c149 ce2c 94f2 49a7 cd20
c845 c1c4 49ce c720 464f 5220 54c8 4520 43c1 d049 54c1 4c80
942c 9420 942c 942f 9420 9454 a254 4f20 c8c1 d645 20c1 2057 4f52 c480 94f4 5749 54c8 2054 c845 20d0 5245
d349 c445 ce54

to Timed Line 21 Data:

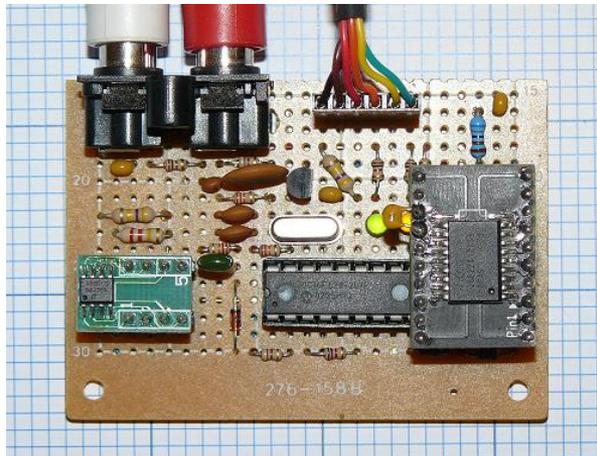
00:15:12:27 942c 9420 942c 942f 9420 94d0 c845 a7d3 20d0 d5d3 c845 c420 4f46 4680 9470 4fd6 4552
2054 c845 20cd 4fd5 ce54 c149 ced3 ae80
00:15:15:10 942c 9420 942c 942f 9420 94f8 57c8 d9bf
00:15:17:07 942c 9420 942c 942f 9420 9454 c245 43c1 d5d3 4520 c845 a7d3 20c1 2052 45c1 4c80 94f4
d9c1 cecb 4545 20c7 4fad c745 5454 4552 ae80
00:15:18:13 942c 9420 942c 942f 9420 9452 c845 20d3 c149 c420 544f 20cd 452c 20a2 43c1 d054 c149
ce2c 94f2 49a7 cd20 c845 c1c4 49ce c720 464f 5220 54c8 4520 43c1 d049 54c1 4c80
00:15:20:16 942c 9420 942c 942f 9420 9454 a254 4f20 c8c1 d645 20c1 2057 4f52 c480 94f4 5749 54c8
2054 c845 20d0 5245 d349 c445 ce54



Searching for Solutions

CC encoding is well-documented, with decoders built into every TV since 1993 and most modern video playing software (VLC, Apple DVD Player, Windows Media Player, etc.) – it *should* be simple! But no...

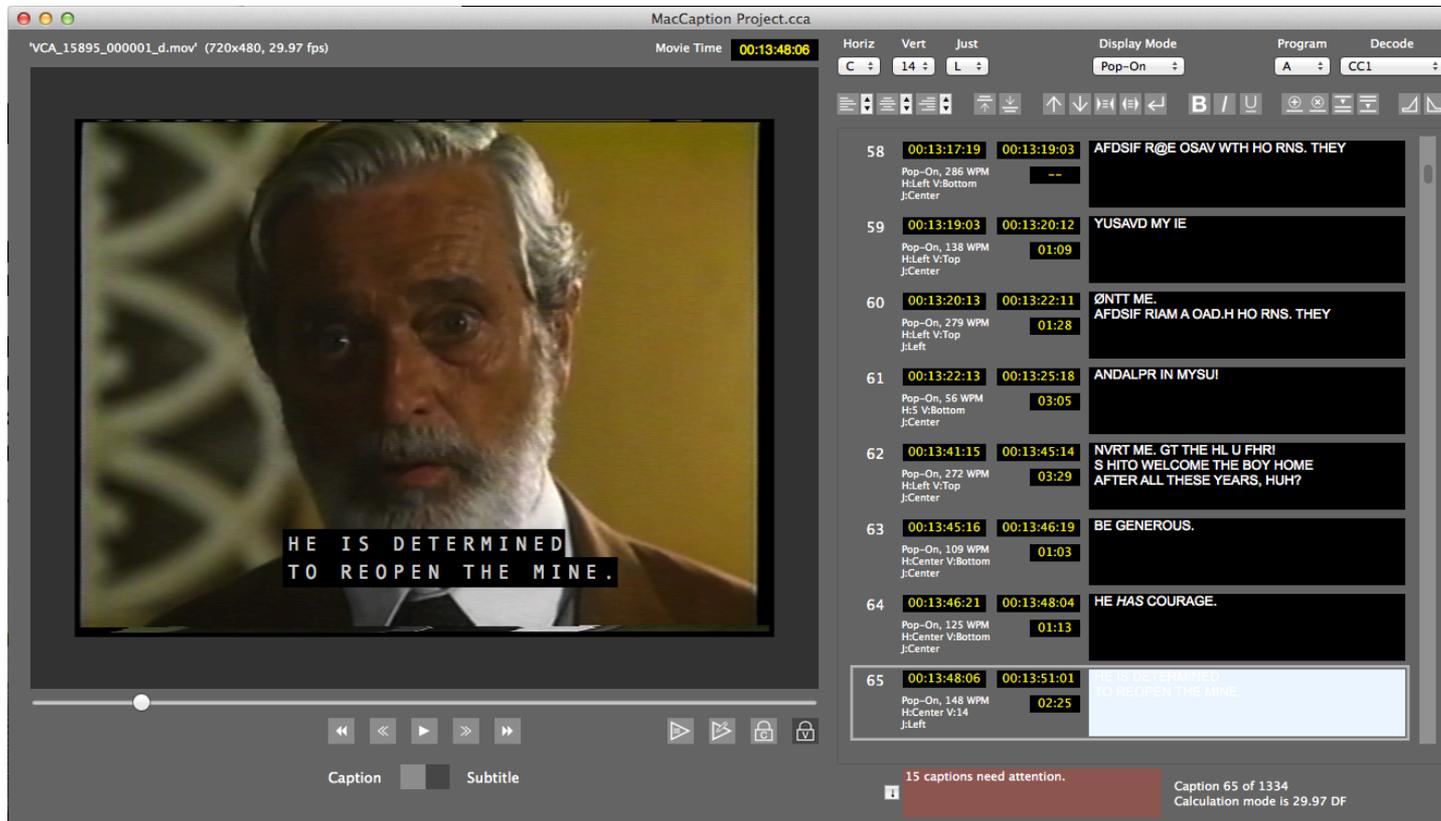
- No out-of-the-box hardware solutions – one site has instructions to *build* a decoder (Compendium Arcana).
- Software solutions exclusively oriented toward more recent digital formats (CCExtractor), or entirely unsupported on modern hardware and operating systems (IAMLine21Decoder).



Compendium Arcana (<http://www.compendiumarcana.com/vbi/>)

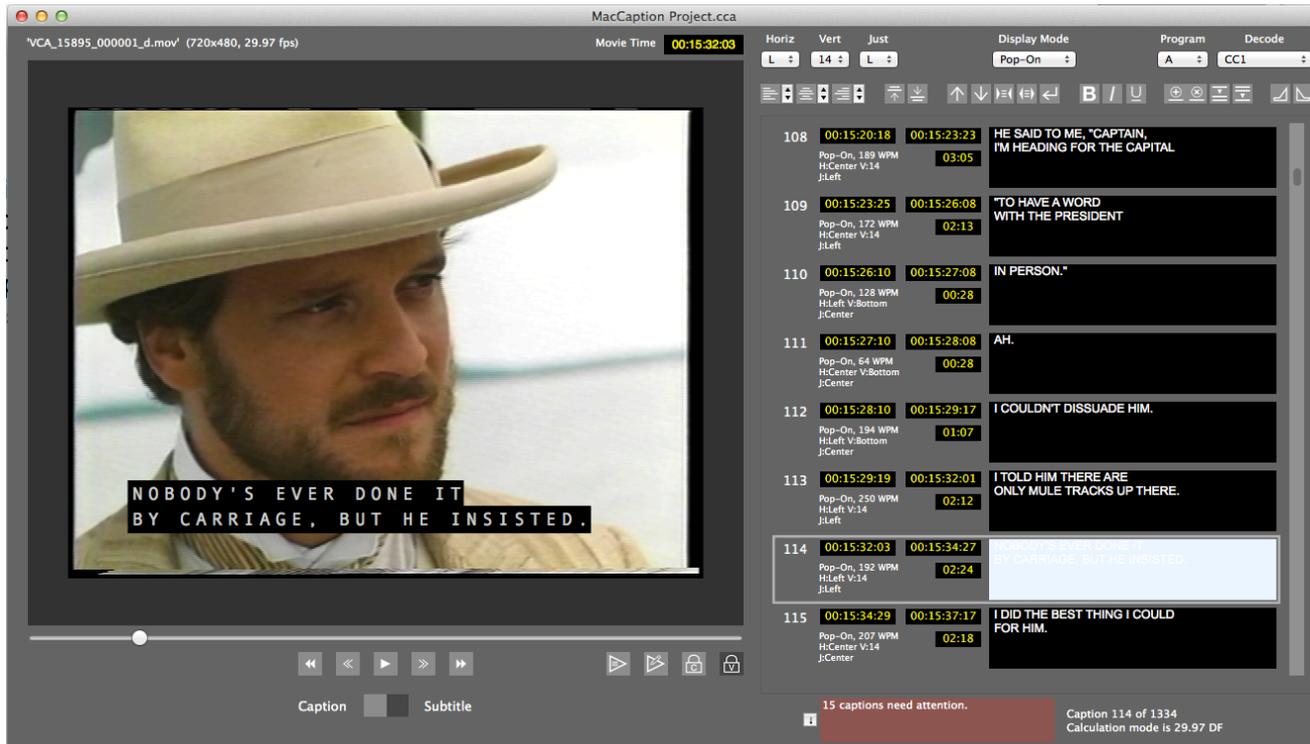
Except for MacCaption!

Available in three editions (*Desktop, Pro, Enterprise*) MacCaption offers this functionality at the Pro (\$6,325) level.



Movie pictured: *Nostromo* (1996)

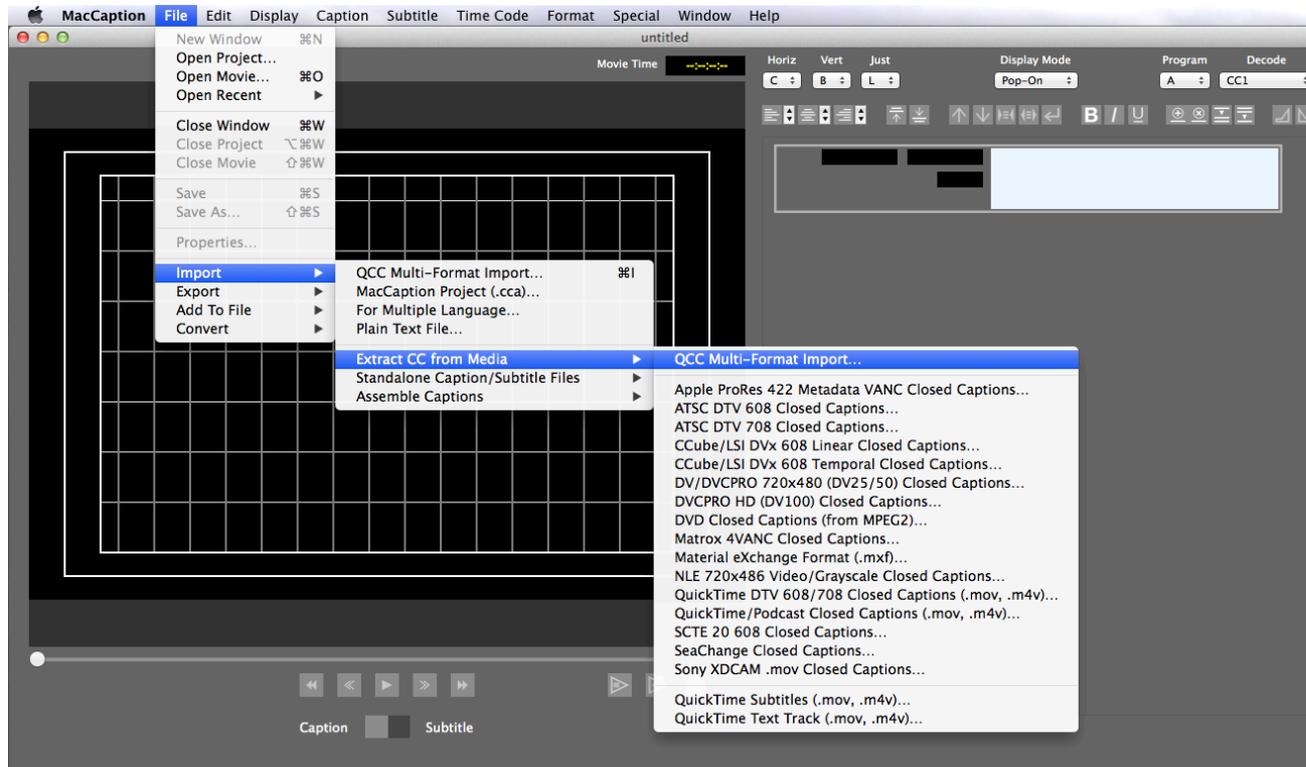
Working with MacCaption



Interface is designed primarily for creating and editing captions, but extraction function is easily accessed.

Movie pictured: *Nostromo* (1996)

Working with MacCaption



“QCC Multi-Format Import” easily handles many video and caption types...

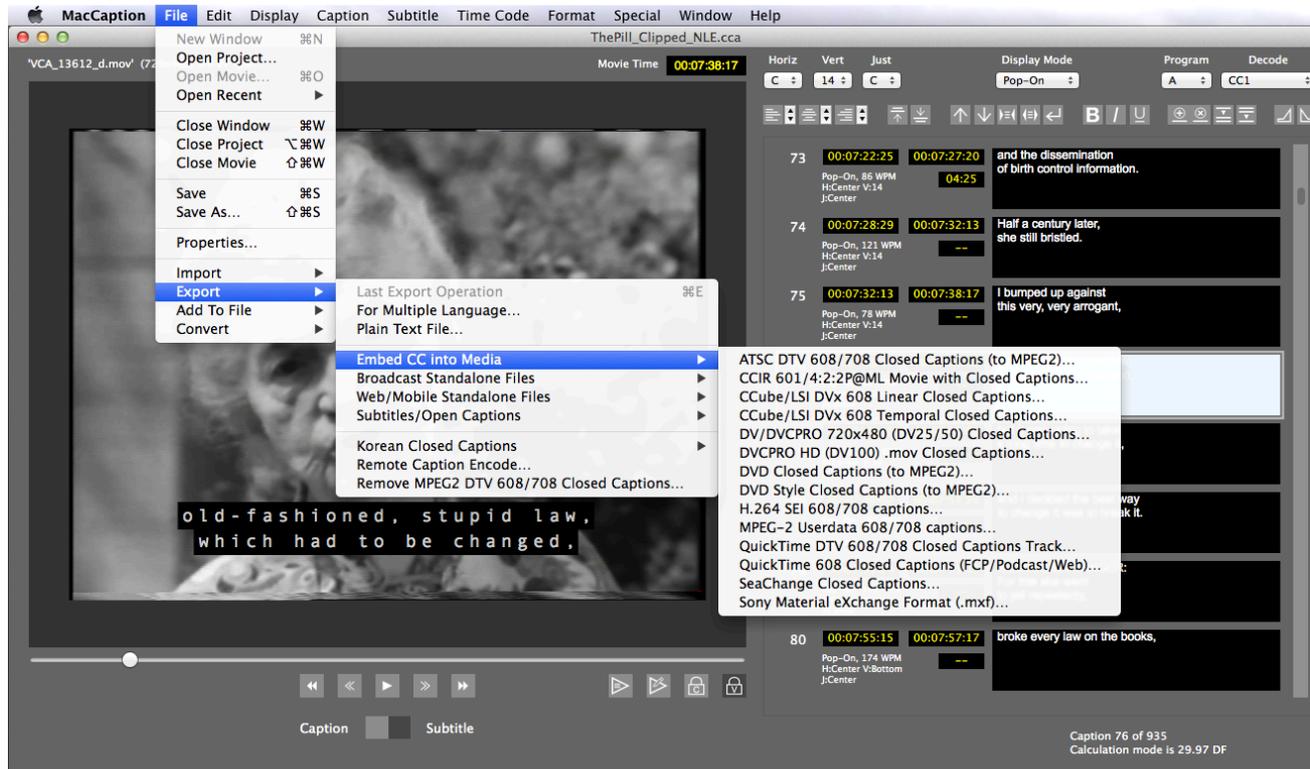
Working with MacCaption



Successful Import!

Movie pictured: *Nostromo* (1996)

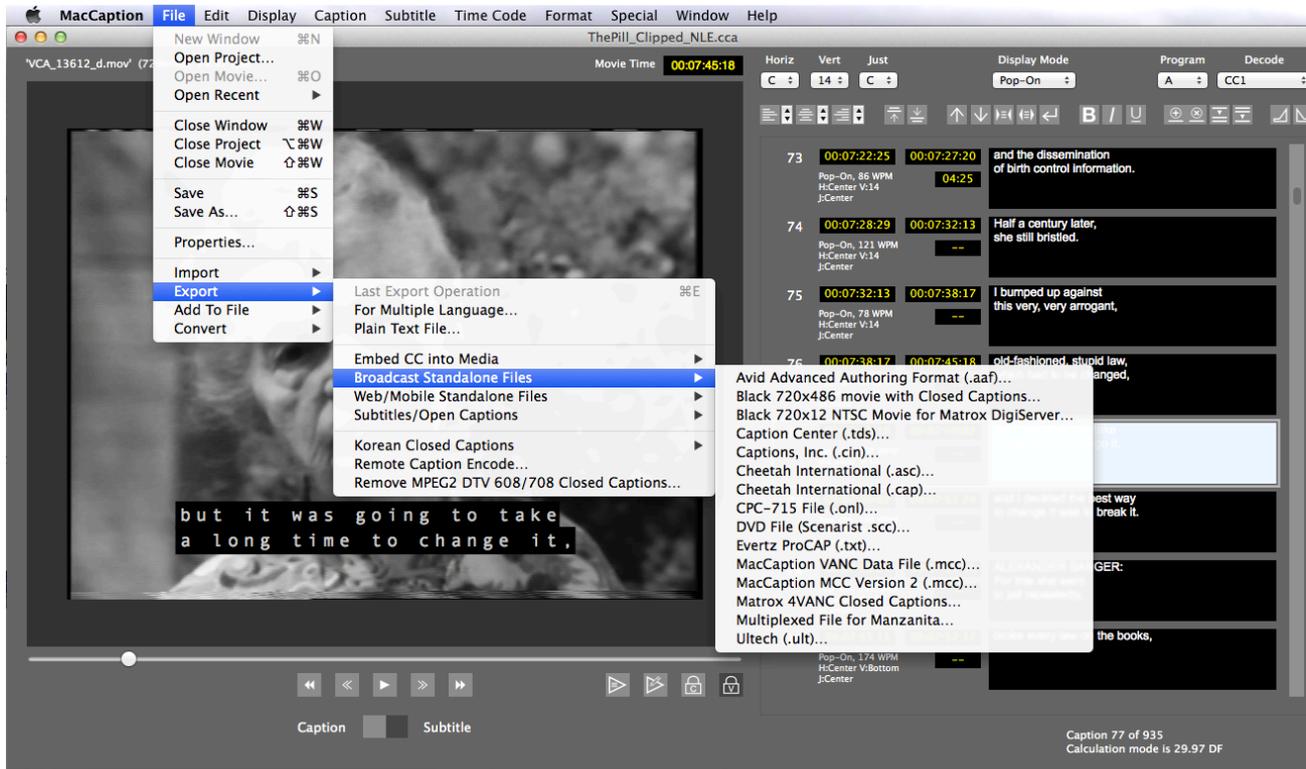
Working with MacCaption



And MacCaption offers many options both for embedding captions, and for exporting them as sidecar files.

Movie pictured: *The Pill* (2003)

Working with MacCaption



And MacCaption offers many options both for embedding captions, and for exporting them as sidecar files.

Movie pictured: *The Pill* (2003)

Failures with MacCaption

VCA_04722_Full_QCC.cca

'VCA_04722_d-MPEG-2 6.2Mbps 2-pass.m2v' (640x480, 29.97 fps) Movie Time 00:03:34:21

Horiz Vert Just Display Mode Program Decode
 L T C Pop-On A CC1

3 00:03:04:21 00:03:06:19 NO SIRVE PARA NADA
ESE HIJUEPUTA,
 Pop-On, 186 WPM 01:28
 H:5 V:14
 J:Left

4 00:03:06:21 00:03:08:04 GENTE BLANCA,
LES PIGMENTE MANFICIAL
 Pop-On, 209 WPM 01:13
 H:Center V:14
 J:Center

5 00:03:08:06 00:03:09:21 PAIN IN MY NECK.
 Pop-On, 160 WPM 01:15
 H:Center V:Bottom
 J:Center

6 00:03:11:06 00:03:12:19 HELLO, MR. PRODUCER?
 Pop-On, 125 WPM 01:13
 H:5 V:Bottom
 J:Center

7 00:03:12:21 00:03:14:19 I REFUSE TO GO ON.
 Pop-On, 155 WPM 01:28
 H:5 V:Bottom
 J:Center

8 00:03:14:21 00:03:22:04 I'M A METHOD TALK-SHOW HOST,
AND I NEED TO PREPARE.
 Pop-On, 80 WPM 07:13
 H:Left V:14
 J:Center

9 00:03:34:21 00:29:45:19 CN SAGNEHJUEPUA
CN[MTATNG MONKYZONS SUR,M UPA,
 Pop-On, 0 WPM 99:99
 H:Left V:Top
 J:Center

Caption 9 of 9
 Calculation mode is 29.97 DF

Movie pictured: *Mambo Mouth* (1991)

Consultation with Telestream



Top Rated Customer Support

Get service, support, and professional advice from our product experts

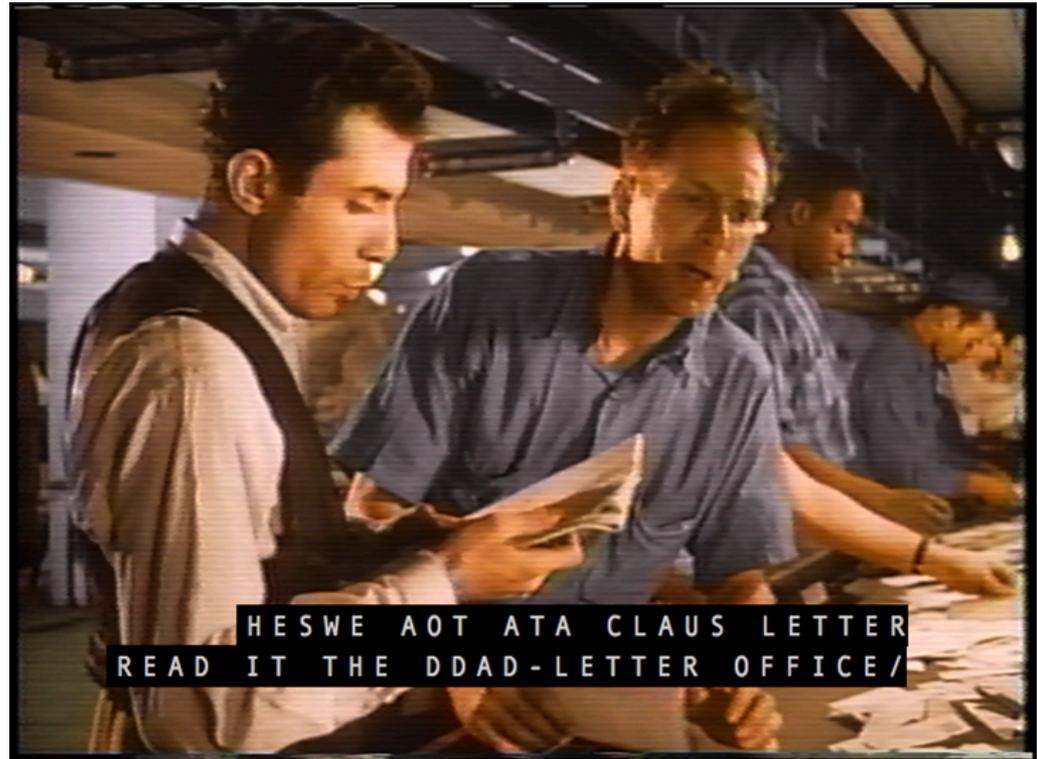
[Quick Start Guide](#) | [CaptionMaker Knowledge Base](#) | [MacCaption Knowledge Base](#) | [Troubleshooting](#)



- Very friendly
- Ludicrously handsome
- Acknowledged that this function of MacCaption is not very commonly used, and is not the primary focus of their knowledge.
- Performed diagnostics on one of our files: found that the captions were on a line other than 21.
- Advised on a way to change the line read by MacCaption.

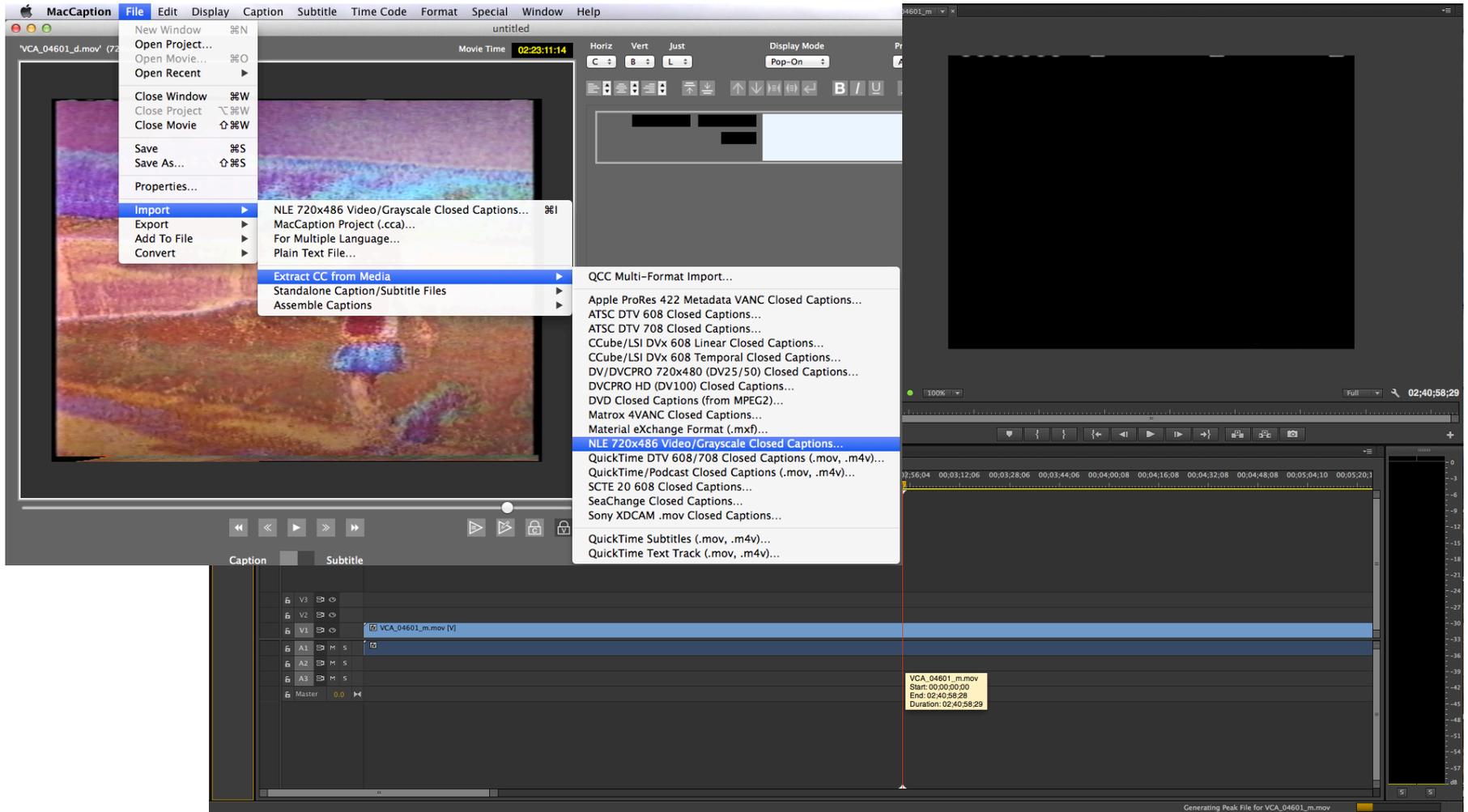
Consultation With Telestream

- Results were not reliable, and re-decoding long uncompressed files for uncertain results was too great a drag on productivity.
- Ultimately uncertain whether the technique recommended actually changes the line.
- Fortunately, we continued to experiment while in conversation with Telestream, and eventually hit on a solution.



Miracle on 34th Street (1947, colorized 1985)

Clipping the Non-Captioned Area



Movie pictured: *Until the End of the World* (1991)

Clipping the Non-Captioned Area

The screenshot shows a video editing software interface with a video player on the left and a caption list on the right. The video player displays a scene from the movie 'Mambo Mouth' with a caption '[IMITATING MONKEY]' overlaid. The caption list on the right shows a list of captions with their start and end times, and the 16th caption, '[IMITATING MONKEY]', is highlighted in light blue. The software interface includes various controls such as 'Movie Time', 'Display Mode', and 'Program'.

Movie Time: 00:00:33:25

Display Mode: Pop-On

Program: A

Decode: CCL

Horiz: 5, Vert: B, Just: C

10 00:00:21:10 00:00:23:10 YO TENGO QUE PREPARE BY MYSELF

11 00:00:23:10 00:00:24:25 CONO PUEDO HACER.

12 00:00:24:25 00:00:26:25 AY, DIOS MIO, QUE SUFRIMIENTO,

13 00:00:26:25 00:00:30:10 COMO SUFRO, MI CULPA, MI CULPA, DIOS MIO.

14 00:00:30:10 00:00:31:25 MI CULPA.

15 00:00:31:25 00:00:33:25 I'M KING OF THE AMAZONS.

16 00:00:33:25 00:00:36:10 [IMITATING MONKEY]

17 00:00:36:10 00:00:39:10 [LAUGHING]

Caption 16 of 1033
Calculation mode is 29.97 DF

Movie pictured: *Mambo Mouth* (1991)



Rudolph the Red-Nosed Reindeer (1964)



Top Rated Customer Support

Get service, support, and professional advice from our product experts

[Quick Start Guide](#) | [CaptionMaker Knowledge Base](#) | [MacCaption Knowledge Base](#) | [Troubleshooting](#)



WON'T YOU GUIDE
MY SLEIGH TONIGHT?

Rudolph the Red-Nosed Reindeer (1964)



NYU | LIBRARIES

18	00:00:41:25	00:00:44:16	Pop-On, 155 WPM H:5 V:14 J:Center	--
19	00:00:44:16	00:00:47:25	Pop-On, 145 WPM H:Left V:Top J:Center	03:09
20	00:00:56:06	00:00:57:09	Pop-On, 163 WPM H:9 V:Top J:Center	01:03
21	00:00:59:04	00:01:02:23	Pop-On, 33 WPM H:5 V:Bottom J:Center	--
22	00:01:02:23	00:01:05:10	Pop-On, 140 WPM H:Center V:Bottom J:Center	--
23	00:01:05:10	00:01:06:10	Pop-On, 60 WPM H:13 V:Top J:Center	--
24	00:01:06:10	00:01:07:25	Pop-On, 40 WPM H:13 V:Top J:Center	--
25	00:01:07:25	00:01:08:25	Pop-On, 60 WPM H:13 V:Top J:Center	--

Ripple Time Code

New starting time code

Set starting time code from current movie time

Increase time code by

Decrease time code by

Ripple Timecode!

18	00:02:51:22	00:02:54:13	Pop-On, 155 WPM H:5 V:14 J:Center	--
19	00:02:54:13	00:02:57:22	Pop-On, 145 WPM H:Left V:Top J:Center	03:09
20	00:03:06:05	00:03:07:08	Pop-On, 163 WPM H:9 V:Top J:Center	01:03
21	00:03:09:03	00:03:12:20	Pop-On, 33 WPM H:5 V:Bottom J:Center	--
22	00:03:12:20	00:03:15:07	Pop-On, 140 WPM H:Center V:Bottom J:Center	--
23	00:03:15:07	00:03:16:07	Pop-On, 60 WPM H:13 V:Top J:Center	--
24	00:03:16:07	00:03:17:22	Pop-On, 40 WPM H:13 V:Top J:Center	--
25	00:03:17:22	00:03:18:22	Pop-On, 60 WPM H:13 V:Top J:Center	--



Rudolph the Red-Nosed Reindeer (1964)



Rudolph the Red-Nosed Reindeer (1964)



THANK YOU!

Michael Grant
mhg311@nyu.edu

THE END