

Importing Closed Caption Files into QuickScan

Revised: 1/12/07

- **Initial Setup**
- **Program Schedule Grid Fields**
- **Other Closed Caption Import Options**
- **Entering Delimiters Manually within QuickScan**
- **Enter Delimiters Manually Outside of QuickScan**
- **Alternate Import location for cc records (optional)**

Initial Setup

Important, regarding cc file naming: If you choose to use one of the automatic closed caption (cc) import methods in QuickScan, it is necessary that the CCArchive software event Titles for recording each program are named with the station's call letters as the first characters of the eight-character title (e.g. "WKRN10PM"). This event Title becomes the file name for the program's cc file, but more importantly, it becomes the first line of the closed caption document itself, which QuickScan reads during the file import. *Three-character call letters must be followed by a numeral.*

Delimiters?? Four our purposes a "delimiter" is a character or set of characters used to identify the end of a news story. Typically, triple angle brackets (>>>) or triple pound sign characters (###) are used. These delimiters are normally entered into the CC text by the broadcaster. Unfortunately, in many TV markets around the country, this is not the case.

After CCArchive events have been correctly setup the next step is to make sure that you have the corresponding broadcasts setup correctly in QuickScan's Media database. Clicking the Media button on QuickScan's main menu accesses this.

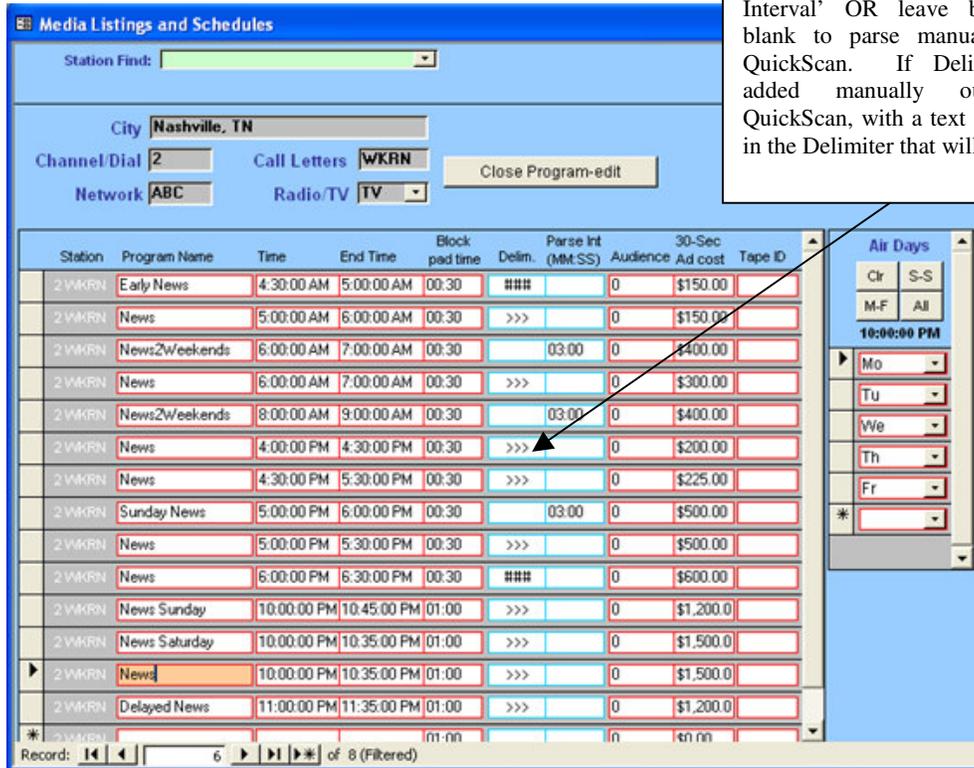
1. Enter each station by clicking the **Add** button at the top of the screen. Enter the city, channel, call letters, network and whether this is a radio or TV station. Click the Add button to enter another station. The previous station entry is automatically saved.
2. After you've finished adding stations select a station from the **Station Find** button at the top of the screen and click the **Enable Program Edit** button to view the program schedule grid for that station.

Program Schedule Grid Fields:

- **Program Name:** The official name of the program being broadcast in this time slot.
- **Time:** The actual scheduled broadcast time for the program.
- **End Time:** The time the broadcast is scheduled to end. This information is important to the cc import process and allows QuickScan to know where the program ends in order to parse it correctly.
- **Block pad time:** Additional time that will be pulled or imported from the cc file before and after the stated "Time" and "End Time". This should never equal or exceed the amount of time used for the Parse Interval (below).
- **Dlim. (Delimiter) and Parse int (Parse Interval):** First determine the delimiter characters are being used in the broadcasts you monitor. Do this by opening a cc file in a text editor (such as Notepad). Once you've determined this you can enter the delimiter character being used into the appropriate fields. For each program record there will be a field called "Delimiter" and another called "Parse Interval". What is entered (or not entered) into these fields will determine how QuickScan processes the file for a particular program. You may enter a delimiter or a parse interval or neither. *You may not enter both a delimiter and a parse interval.* Entering a *delimiter* tells QuickScan to parse the file containing this program into stories according to where the delimiter characters are located within the text. Entering a *parse interval* tells QuickScan to parse the program into random chunks each the length of time specified. If you set the *parse interval* to 02:00 (minutes) for a 30-minute broadcast, QuickScan would parse that into 15 "stories" of 2-minute length each. Leaving both fields blank tells QuickScan to ignore the file containing this program. Use this setting if you intend to manually parse the file from within QuickScan.

If the closed caption file does not contain delimiters, you'll have three choices for how to import it:

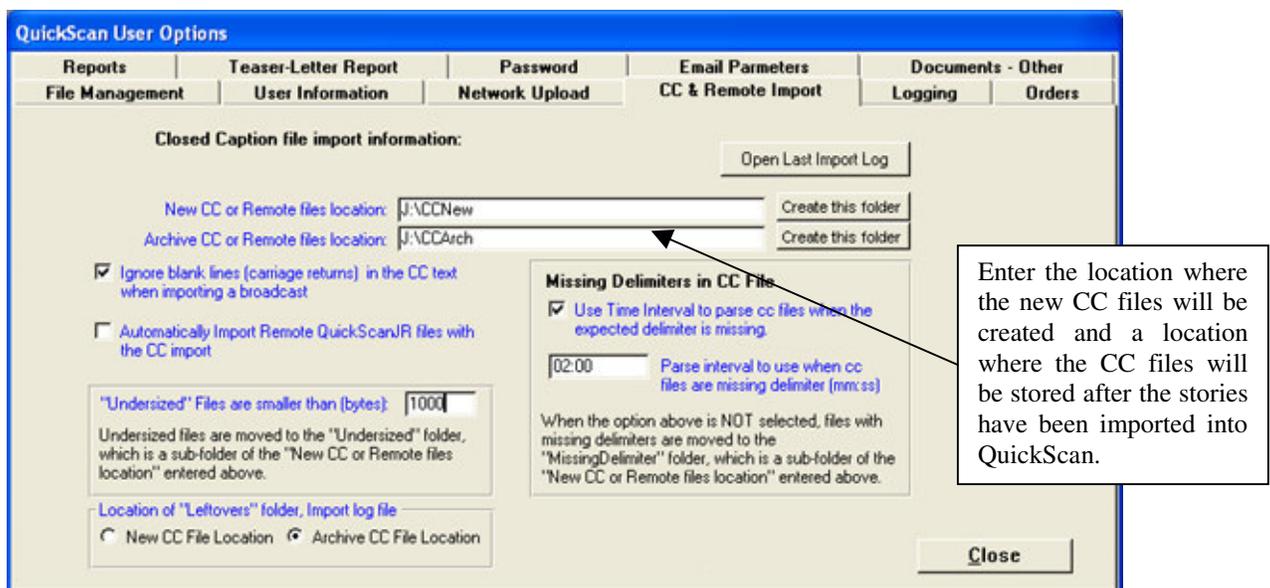
1. **Auto Parse, by Time Interval** – Described below.
2. **Manual Parse** –
 - a. **Within QuickScan** - Leave both the *Delimiter* field and the *Parse Interval* field blank (Described in detail further below).
 - b. **Outside QuickScan** - Enter delimiter characters manually into the cc file using a text editor. This method is often preferred when there are no delimiters provided by the broadcaster because you can easily enter delimiter characters into the file with a text editor and then import all your cc files at once. If you use this method, you'll need to enter the delimiter characters you're going to use into the 'Delimiter' field so QuickScan will know what to look for



- **Audience:** This is the estimated number of viewers watching this particular broadcast. This number will appear on QuickScan reports that display audience values. This number must be obtained through a research company or the TV station.
- **30-Sec Ad cost:** This is the estimated advertisement cost of commercial during this time slot. This number will be used to calculate the relative add value for each story displayed on a QuickScan report containing this information.
- **Tape ID:** The user-defined name for the tape or media where this broadcast is recorded. This will appear on subsequent orders and makes the location of tape for production easier.
- **Hidden Fields -i / -n:** Placing your cursor into the Tape ID field and clicking the Tab key on your keyboard exposes these two check-box fields. Checking the **-i** field for a program will prevent the program from being uploaded to inewsnetwork during the upload process. Checking the **-n** field for a program will prevent that program from being included in the NDS file during the NDS export.
- **Air Days:** It is important that you identify the days of the week each program will be aired. To enter the day of the week for a program you must first select that program. The program time for the selected program will then be displayed under the four selection quick-select buttons. Clicking **M-F** enters each of the weekdays; **S-S** enters Saturday and Sunday; **All** enters every day of the week and **Clr** removes all of the entries.

Other Closed Caption Import Options

- **CC File Locations:** In the QuickScan *User Options* on the “CC & Remote Import” tab you’ll enter the location from where your cc files will be imported and to where they’ll be stored once they’re imported.
 1. Create two sub-folders for your cc files. One for the newly created files and one to archive files that have been imported (there’s an option, discussed below, to delete older, archived cc files). It is recommended that you create these folders as sub-folders in the ‘common’ folder where QuickScan’s other main data files are kept. On a single-computer system this will be the c:\QSPPro folder. On a network this will be a mapped drive.
 2. Name one of the new sub-folders “CCNew” and the other sub-folder, for archived files, “CCArch” (recommended, not required naming).
- **Missing Delimiters in cc File:** You may encounter a situation where you have QuickScan setup to import cc files according to a set of delimiter characters and then, one day, the station neglects to use the delimiters. In this case QuickScan chooses one of two courses of action according to the settings entered here. Set the Missing Delimiter controls to parse the file by a time interval. If the controls are left blank QuickScan will move the file with the missing delimiters to a MissingDelimiters folder, which is a sub-folder of the NewCC folder. This gives the user the opportunity to enter delimiters manually.

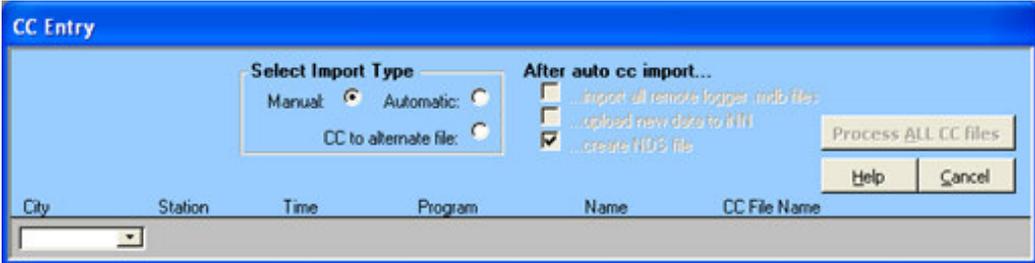


- **Undersized cc Files** - Set the file size limit. This places files that are smaller than expected into an “Undersized” folder in the New cc folder location. Leave this field blank to attempt to import all files regardless of size.
- **Location of “Leftovers” folder, Import log file** – This option allows you to choose where the “LeftOvers” folder will be, in the archive File cc location or the New cc file location. “Leftovers” are files that were not imported due to a problem. Problems as well as other import statistics are concisely described in the cc Import Log test file, which is also placed in this folder.
- **Delete Archived cc Files** – This option is found on the “Network Upload” tab of the User Options. It allows you to tell QuickScan to automatically delete archived cc files once they have become older than the number of days entered.

Entering Delimiters Manually Within QuickScan - This method will allow you to import the text from a broadcast’s cc file and manually separate it into stories using QuickScan’s easy-parse screen.

1. Be sure that you have no delimiter or parse interval entered in the program record (see schedule grid) for the program you want to import using this method.
2. Click the CC Entry button on QuickScan's main menu screen and select "Manual" from the options at the top of the screen.

What does "parse" mean??
 'To parse' simply means to separate into parts, such as in your English class when you had to parse sentences. Now QuickScan will do the parsing for you – or you can do it yourself...



3. Enter the City, Station, Time, Program and user's Name. After the Name is selected a Select File dialog box will open where you may select the cc file that you intend to import. **Important:** Be sure you select the file that correctly matches the City, Station, Time and Program that you just entered. If you don't, you'll be putting the wrong broadcast information on every story record that's imported for this broadcast.
4. After selecting a cc file the screen below will appear with lines of closed caption text from the cc file. Simply read through the story and click with the mouse to place a check in the box to the right of the last line of each story. One line or multiple lines can be kept from importing by using the 'record selectors', which appear as gray boxes to the left of each line and then using the delete key on the keyboard. You may also use the "Del Line" check box, which will eliminate the line of text later, during the story processing. Use the <Page Down> key on your keyboard to quickly move through the stories. Editing of the text may also be performed in this screen.

Use the 'record selector' boxes to quickly select single or multiple lines of CC text for deletion

Define/Edit Stories From CC File

CC file being imported: C:\QS98\CC\0308WSYX.180

Process: Auto Parse by Time Interval

Process Stories

Time Stamp	Indent	Del Line	Closed Caption text	End of Story
18:00:15	<input type="checkbox"/>	<input type="checkbox"/>	IT OUT UNLESS WE HAVE TO. BUT WE FOUND	<input type="checkbox"/>
18:00:16	<input type="checkbox"/>	<input type="checkbox"/>	STACKS OF NUMBERS-- ONE BUSINESS	<input checked="" type="checkbox"/>
18:00:17	<input type="checkbox"/>	<input type="checkbox"/>	THROUGH AWAY, AND IT WAS ALL FOR THE	<input type="checkbox"/>
18:00:18	<input type="checkbox"/>	<input type="checkbox"/>	TAKING. THAT'S TONIGHT'S BIG STORY--	<input type="checkbox"/>
18:00:19	<input type="checkbox"/>	<input type="checkbox"/>	NEWSCENTER FOUND PILES OF PERSONAL	<input type="checkbox"/>
18:00:20	<input type="checkbox"/>	<input type="checkbox"/>	INFORMATION THROWN AWAY IN THIS	<input type="checkbox"/>
18:00:21	<input type="checkbox"/>	<input type="checkbox"/>	DUMPSTER. O-S-I EMPLOYEES, A COLLECTIO	<input type="checkbox"/>
18:00:22	<input type="checkbox"/>	<input type="checkbox"/>	AGENCY IN DUBLIN, THREW AWAY WHAT COUL	<input type="checkbox"/>
18:00:23	<input type="checkbox"/>	<input type="checkbox"/>	BE YOUR VITAL INFORMATION. THIS IS WHA	<input type="checkbox"/>
18:00:24	<input type="checkbox"/>	<input type="checkbox"/>	I FOUND ON TOP OF THE PILE.	<input type="checkbox"/>
18:01:01	<input type="checkbox"/>	<input type="checkbox"/>	ANYONE CAN COME THROUGH HERE, GRAB WHA	<input type="checkbox"/>
18:01:02	<input type="checkbox"/>	<input type="checkbox"/>	THEY WANT-- AND THE INFORMATION IT HAS	<input type="checkbox"/>
18:01:03	<input type="checkbox"/>	<input type="checkbox"/>	SOCIAL	<input type="checkbox"/>
18:01:03	<input type="checkbox"/>	<input type="checkbox"/>	SERURITY NUMBERS AND PEOPLES	<input type="checkbox"/>
18:01:04	<input type="checkbox"/>	<input type="checkbox"/>	NAMES.	<input type="checkbox"/>
18:01:06	<input type="checkbox"/>	<input type="checkbox"/>	THIS STACK OF PAPERS CONTAINED	<input type="checkbox"/>
18:01:09	<input type="checkbox"/>	<input type="checkbox"/>	STUDENT LOAN PAYMENTS. WHEN I FIRST	<input type="checkbox"/>
18:01:10	<input type="checkbox"/>	<input type="checkbox"/>	QUESTIONED O-S-I	<input type="checkbox"/>
18:01:12	<input type="checkbox"/>	<input type="checkbox"/>	ABOUT THIS PRACTICE THEY REFUSED TO	<input type="checkbox"/>
18:01:13	<input type="checkbox"/>	<input type="checkbox"/>	TALK ABOUT IT ON CAMERA.	<input checked="" type="checkbox"/>
18:01:16	<input type="checkbox"/>	<input type="checkbox"/>	ARE YOU GOING TO SHRED IT? IT WILL BE	<input type="checkbox"/>

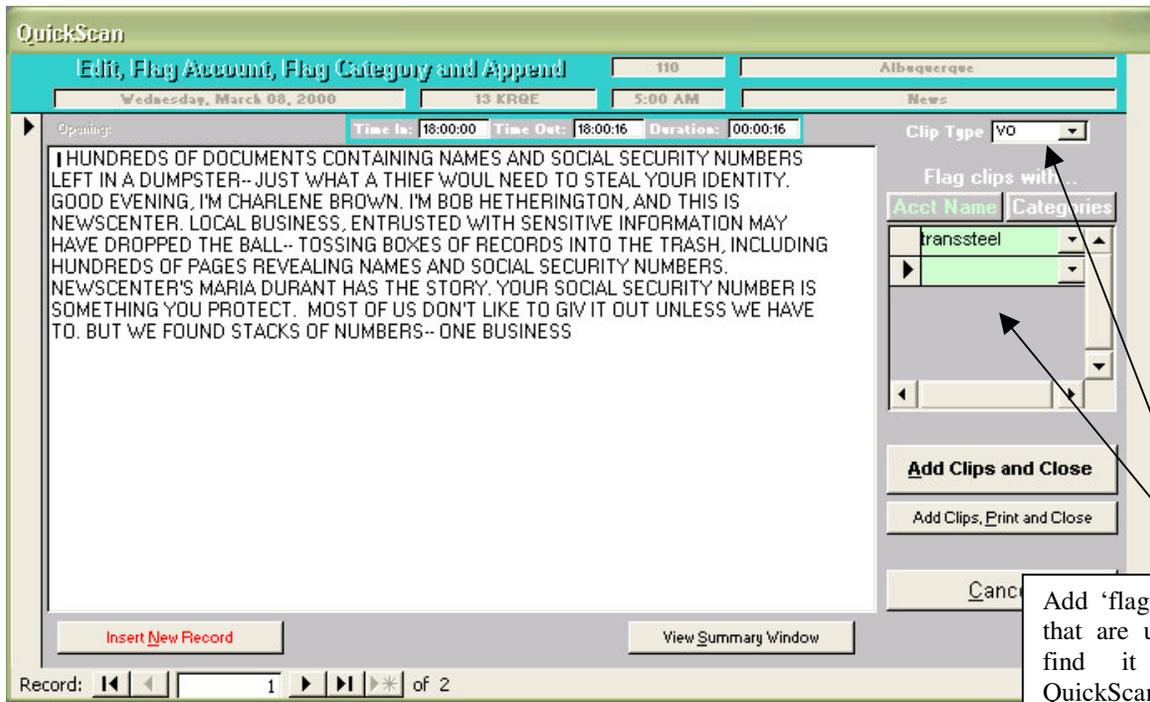
Record: 35 of 424

Place a check in the box left of the last line of the story

- When checks have been placed in all of the boxes at the end of each story click the “Process Stories” button, which parses each story into a record in the database and displays the stories in the screen shown below where they may be edited further or appended to the permanent clip file.
- When the CC Edit and Append screen, shown below, becomes visible, you can perform more edits on the cc text. In this screen you may also add ‘flags’ to each story record (‘Clip Types’, ‘Clip Categories’, ‘Account Name’), run the spell checker and then append the new clip records to the main data file.



Click to run the spell checker



Add ‘flags’ to a story that are used later to find it with the QuickScan Search Engine

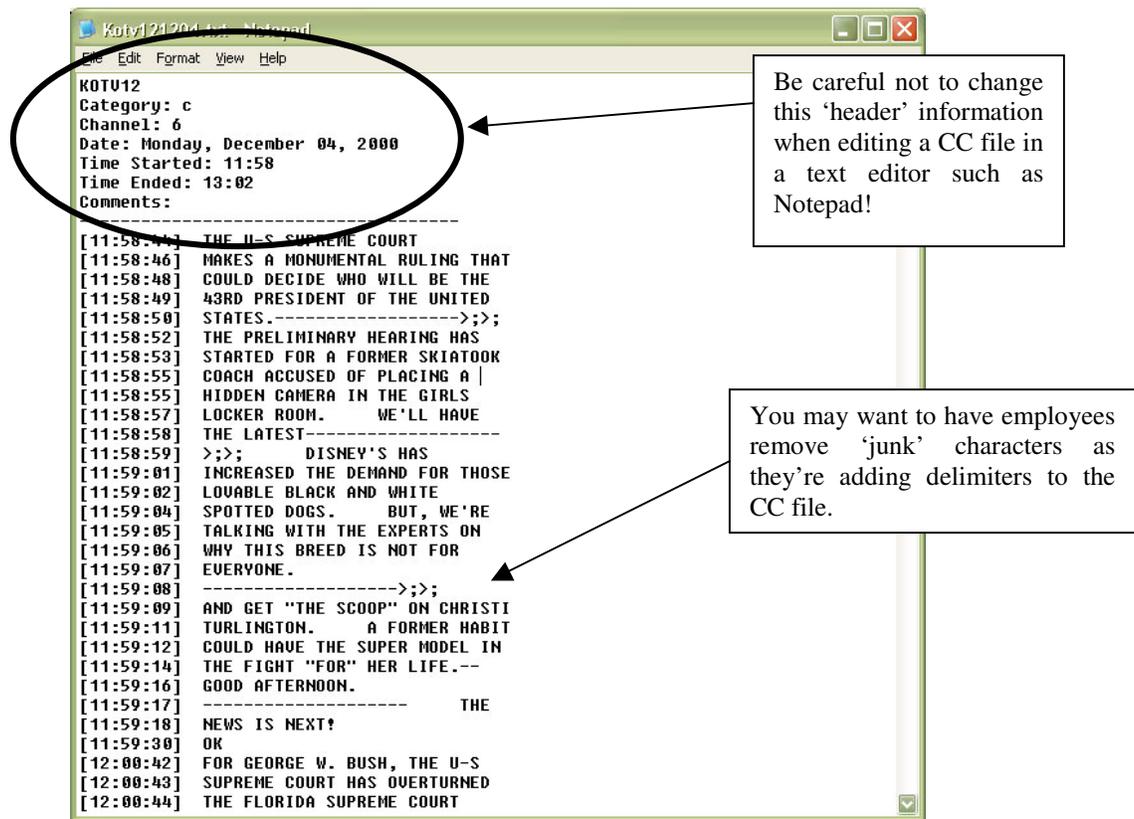
- Clicking the “Append Clips and Close” button appends the data to the main database and renames and moves the cc file being parsed to the archive folder.

Enter Delimiters Manually Outside of QuickScan

With this method story delimiter characters are entered directly into a cc file using a text editor, such as *Windows Notepad* or *WordPad*. Once the delimiters have been added to your cc files you can easily import a large number of them into QuickScan with the click of a single button. This method might be used as an alternative to entering delimiters manually within QuickScan.

- Decide on a parse interval to mark the end of each story (recommended: >>>). In the delimiter field of the program-record in the program schedule grid (shown above) type the delimiter characters. Close or minimize QuickScan.
- Open a cc file with your favorite text editor, such as Windows Notepad (if the cc file is too large for Notepad it will be opened automatically in WordPad). **Important:** If you use a word processor to edit the cc files, be sure to save the files as a *text files*. Saving a cc file as a word processor file will cause a failure during the cc file import process.

3. *Optional:* Delete any extra lines from the text that may precede the start of the newscast. These only add clutter to the database. You should also remove the time codes to the left of the deleted lines. **Be careful** not to change any of the header information that appears above the story text!
4. Go to the end of the first story and type the delimiters that you decided on in step one (such as: >>>).
5. Now, with your mouse, drag the cursor over the delimiters and Copy them (you can use the <Ctrl> and <c> keys on your keyboard for a quick copy command or go to the Edit menu and select "Copy").
6. Go to the end of the next story and place your cursor there by clicking that spot with your left mouse button. Paste the delimiters here (use the <Ctrl> and <v> keys on your keyboard for a quick paste command or go to the Edit menu and select "Paste").
7. Go to the end of the *next* story and paste delimiters there (the delimiters you copied in the last step will remain in the Windows 'clip board' so there is no need to copy again). Continue with this process until you have placed a delimiter at the end of each story in the file.
8. *Optional:* Remove any extra lines that appear after the last story. You should also remove the time codes to the left of these deleted lines.



9. When you have finished editing the cc file save the changes and open the next cc file you want to work on. Repeat the steps above for each remaining file.
10. Once the story delimiters have been added to all of the cc files it's time to import them into the QuickScan database. To do this click the "CC Entry" button on QuickScan's main menu screen and then select the "Automatic" option on the Import/Parse screen.
11. Click Continue.
12. After a short while a screen will appear with a log of all of the files that were imported into the database. For your convenience, this log may be printed, faxed or emailed and is a handy tool for troubleshooting should you need technical support.

Alternate Import location for cc records (optional)

This is an option that is selected in the CC Import/Parse screen and is intended for news monitors who primarily log the news manually but want to use closed caption text as a backup or additional

source of data. By selecting this option stories that are imported from cc files are placed into a file called CCCurrent.mdb (normally, data inputted into QuickScan goes into the Current.mdb file). This option should be used ONLY if you intend to use the cc data as a secondary data source. For example: This feature should not be used if you log most of the news manually but want to use only closed captioning for the morning programs (although it might be used to import the afternoon and evening programs that are also being logged manually).

To search in the Alternate cc database "Other" must be selected from the "Current, Archive, Other" option box of the QuickScan search engine screen. When this option is selected a select-file dialog box will open where the CCCurrent.mdb file must be selected (the CCCurrent.mdb file will be located in the same folder as all of the QuickScan data files. This is the C:\QSPro folder on stand-alone systems and it will be in the common network file folder for network users).

QuickScan automation cannot be employed when using the alternate cc import method.