The Auto-Coding Tool

The Auto-Coding tool finds text passages, selects a specified amount of text (e.g., the exact match, or spread to the surrounding word, sentence, or paragraph), and then codes the passages with a previously selected code.

Auto-coding is useful when coding structural information like speaker turns in group interviews, or other sections that can easily be identified by a text search. Auto-coding allows you to quickly collect ideas that belong to a certain concept on the basis of words or patterns found in the text.

Concepts

The Auto-Coding Dialog

The Auto-Coding tool combines the Text Search tool with an automatic segmentation and code assignment mechanism.

To open the Auto-Coding Dialog, select Codes / Coding / Auto Coding from the main menu.

Note the New Code button, which allows you to create codes on the fly for the auto-coding procedure.

Based on the GREP search expression shown in the above figure, all occurrences of string Speaker 1: located at the beginning of a paragraph in all textual PdS are found. Next, the whole speaker unit (which could contain one or more paragraphs) will be selected before a quotation is created and the code “Speaker 1” will be assigned.
The option **CONFIRM ALWAYS** is deactivated as the search is unambiguous and is likely to yield the desired results.

**Search Expression**

The top part of the Auto-Coding Dialog window resembles the Text Search tool (see “The Text Search Tool” on page 210). Search expressions can be entered or search categories can be selected. The search mode can also be set in a similar way as for the Text Search tool.

**Scope Of Search**

**Scope of Search** can be all textual PDs, the selected PD only, or it may be restricted to a set of texts from a PD family. When you choose “All current PDs,” the search starts at the beginning of the currently selected primary text.

If "Selected PD" is chosen, the auto-coding process starts at the current position of the text cursor.

**Segment Size**

When a matched string is found, the size of the segment to be coded can be specified as follows:

- The **Exact Match** only
- The **Word** surrounding the matched string
- The **Sentence** surrounding the matched string
- The matched string expanded to paragraph boundaries (**Single Hard Return**).
- The matched string expanded to one or more paragraphs up to the next blank lines (**Multiple Hard Returns**). This option is useful if you have information that should be regarded as one unit, but is separated in multiple paragraphs.
- **All Text** of the PD containing a match.

When transcribing, enter a blank line between speaker turns if you are going to use the auto-coding tool for coding speaker units.
Test Your Search Expression

The Auto-Coding tool creates a new quotation for every matched text passage that is not yet a quotation. A potentially large number of inadequate quotations could be produced by an imprecise search.

For this reason, you should always test a search pattern by using the Text Search tool first (see "The Text Search Tool" on page 210) or select the Confirm option and click Skip. This test will give an indication if the final search will yield meaningful results. The advantage of testing with the Text Search tool is that no quotations are created.

Another option is to auto code into a new code. If you later find out that the search did not yield good results, you can delete the code and all created quotations with two mouse clicks and try a different search.

How To Auto-Code

Auto-coding can be fully automatic or semi-automatic. The sections below will clarify the difference and will show the advantages of each method.

Fully Automatic Coding

1. Select the PD you want to code using the auto-coding procedure. If you only want to auto-code the current text, load the PD in question. If you want to auto-code all texts, load the first PD and position the cursor at the beginning of the document.

2. To open the Auto-Coding tool, choose Codes / Coding / AUTO CODING from the main menu.

3. If no code is selected, select a code from the drop-down list at the top of the Auto-Coding tool.

4. Enter a new search string or choose an existing search category.

5. Specify the search scope, e.g. "Selected PD"

6. Specify the segment size.

7. Start the process by clicking START.

Semi-Automatic Coding

As it is not always desirable to let the program decide whether or not to code a given text passage, you can control the process by checking "Confirm always." Specific confirmation from the user on each new code will then be requested.

An example for the usage of semi-automatic coding would be to code for the concept “distress.” Indications that a person might be distressed could be words like nervousness, tension, unease, edginess, etc. In order to capture this, you would do the following:
Create a code with the name Distress.

Open the Auto-Coding tool.

Select the appropriate code.

In the search entry field define a new category by entering:

```
DISTRESS:= nervous|tens|uneas|edg|....etc.
```

The OR operator in this example is marked in red for visibility only. The search term in ATLAS.ti does not show colors or formatting.

Depending on your data set, select a suitable **scope** of search (current PD, all PDs, or a specific PD family).

Select the **segment size** to be coded, e.g. sentence.

Check **Confirm Always** and click on the **Start** button.

Every time the program finds a piece of text that matches your search category, it stops the search and highlights the text it has found. You can then read the surrounding context and decide whether the text passage really has something to do with distress. If it does, click **Code it**, otherwise click **Skip it**. The program continues to search for the next match.

You can uncheck the **Confirm always** box at any time and let ATLAS.ti scan through the rest of your texts without prompting for further confirmation.

**Brushing-up Results After Auto-Coding**

Since no automatic search can guarantee 100% meaningful results, the quotations created and assigned to the selected code during auto-coding should be screened and modified if needed.

**CHECKING AND MODIFYING QUOTATIONS**

- Close the Auto-Coding tool.
- Open both the Code Manager and the Quotation Manager.
- Select the code you just used for auto-coding.
- In the Quotation Manager, set the quotation filter to "Selected Code".
- Click the Created column header to find all recently created quotations at the top of the list.
- Browse through the quotations assigned to this code by using the [Next] and [Previous] buttons in the Quotation Manager.
- Modify or delete misaligned or inappropriate quotations. See “Modifying Textual and PDF Quotations” on page 170.
Preparing Documents Optimized For Auto-Coding

The following instructions are useful for preparing transcriptions of focus group sessions, questionnaires, or interviews. Such data usually contains different speakers’ sections. The hints given here also apply for other documents that include sections you wish to identify for auto-coding.

It would be tedious to code speaker or section turns manually. Two things are needed: A good “marker” for which to search and, once the marker is found, a reliable identification of the unit (sentence, paragraph) to be coded.

Insert easily identifiable markup in the text to let the auto-coding pattern matcher do this for you. For proper markup, a little knowledge regarding the auto-coder’s search procedure is helpful. In combination with a few formatting rules, documents can be created that can readily support auto-coding quite a bit.

For the auto-coding tool to yield useful results, it is necessary to properly structure the source text using unique identifiers.

Simple examples are:

• P: or <Peter> for a respondent with the name Peter
• 07-01-11 letter for a letter written on the first of July 2011.

The identifier should be used exclusively to mark passages in the text that indeed relate to the person or object identified. The plain word Peter will likely also occur elsewhere in the text (for instance, when another person is referring to Peter). The markup "P:" or <Peter> however, is unlikely to occur elsewhere.

In order for the Auto-Coding tool to select a complete speaker section, a section delimiter is needed. As discussed above, a speaker or section turn will start with the speaker’s identifier markup. The end of a section is best marked by an empty line. By doing this, you can still use single hard returns to segment a speaker section into paragraphs.

When auto-coding such a document, you would choose MULTI HARD RETURNS for extending the matched text.

Using <Peter> as the search pattern would then yield a quotation including everything from <Peter> up to the empty line:
INT: Ok, so erm, so you’ve got the majority of the information anyway, I mean, primarily the most important thing really is that erm obviously you can withdraw at any time from the study, apart from when it goes on to the internet. 
TREVOR: It’s too late then yeah. 
INT: Yes. 
TREVOR: You’re in trouble if you change your mind then [laughs]. 
INT: Exactly [laughs]. But up until that point, and we will sort of tell you when that point’s gonna be. 
TREVOR: It’s alright I just won’t watch my bit [laughs].

Figure 190: Example transcript