Sequence Viewer Manual

Version 6.2

Wil Dijkstra

June 2016

1 General

The SEQUENCE VIEWER program is intended to aid researchers in observing and coding verbal and non-verbal behavior. Such behavior is viewed in SEQUENCE VIEWER as a sequence of acts, or events.

However, not only sequences of observed behavior, but also completely different kinds of sequences, for example life careers can be analyzed. In such a case, an event refers to for example a working period.

The SEQUENCE VIEWER program is free. You are only asked to personalize your copy of SEQUENCE VIEWER, and to accept the license agreement.

If you use SEQUENCE VIEWER for your analyses, and present results in articles or papers, you should make a reference to the program like:

W. Dijkstra (2017), Sequence Viewer version 6.2.
http://www.sequenceviewer.nl/

1.1 Help

THE HELP BUTTON

Most setup windows in the SEQUENCE VIEWER program have a help button at the lower left corner of the window, that looks like this:  . If you click this button, it will bring you to a page in this help file.

If you move the mouse over buttons and fields in the setup window, a description of this object appears in the help field (unless it is self-explaining). The help field is part of the toolbar, in the upper left corner of the screen.

Moving the mouse over the help button (without clicking) provides you with information about possible errors in your setup. Sometimes you may be puzzled why the OK button (or similar button that executes a setup) remains disabled. Move the mouse over the help button, and you are (usually) told what's wrong with the setup.

THE HELP FILE

In the find field below (after 'Find:') you can type a word and click the magnifying glass, to search for the word in the help file. If you type more words, a literal search is done; e.g. 'event variable' will find 'event variable', 'event variables', but not just 'event' or 'variable'.

You can use the ampersand (&) to do an 'and' search. For example, 'mark set&name&change' will find 'mark set', provided that 'name' and 'change' are also in the text of the help field, although not necessarily in that order.

In the help texts, links are blue underlined. If you click a link, you are brought to a page with more information about that topic.

You can change the width of the topic field at the left of this window using the bar with the three dots at the left of this field. Just drag it to an appropriate position.

PRINTING THE HELP FILE

You can print the whole help file, including a table of contents, with the print button  below.

You can use **File > Page setup...** to scale the paper size, e.g. to 90 percent, if you find the font size of the print too large. The manual can also be found in the 'Documentation' folder.

COMMUNICATION WITH OTHER PROGRAMS

Quite likely you will copy results of analyses, for example tables, to your favorite word processor. Of course you can simply copy and paste (parts of) your output, or transcripts from the main SEQUENCE VIEWER window. Or you can save the output to disk. Except the output window, there are a number of other procedures that allow you to save data to disk. For example, you can save all transcripts as a text file, or export the sequence variables to use them in a statistical program. And you can import for example files with transcripts, or files with just codes into SEQUENCE VIEWER (see also Import and Export for more specific information).

When you paste text, or open a saved export file in your word processor, you may observe that some information is lost, or looks different. This is especially the case with MS Word 2011 and later. For example, 'tabs' (used to delimit values when you export for example sequence variables) are mysteriously transformed to underscores (_). The problem can partly be solved by selecting 'Keep text only' from the paste options in Word (click the  icon). Now the tabs become real tabs, but all information about text styles (like 'bold') or different text fonts is lost. A better solution is to use the previous version of Word (the one that produces the .doc files instead of .docx files), or to use TextEdit, available on every Mac. If you wish, you can next copy the pasted text from the old version of Word, or from TextEdit, and paste it into MS Word 2011: text styles and text fonts *and* tabs are now preserved...

Similar problems occur when you export files to disk and open such a file in Word 2011. Again, you better open an exported file in TextEdit or the previous version of Word. See Export for more information about the different file formats in which you can save the exported file.

1.2 Introduction

Data in SEQUENCE VIEWER concerns sequences of events. For example life careers, interviewer-respondent interactions, animal behavior, and so on.

A data file consists of a number of *sequences*. Each sequence consists of a number of (ordered) *events*.

SEQUENCES

Sequence variables

To each sequence a number of sequence variables can be added. For example, if one observes an interaction sequence between interviewer and respondent, one may add information about age and gender of both participants to the sequence as sequence variables.

A SEQUENCE VIEWER data file always has two sequence variables added: the length of the sequence, that is the number of events of the sequence ('SEQSIZE'), and the start time ('STARTTIME') of the sequence (even if the time does not play a role in your sequences).

Sequence text

In addition to sequence variables, sequences can also have *sequence text*, for example utterances in a conversation, descriptions of animal behavior, a paragraph in a newspaper article, or whatever you want to analyze. If a sequence has sequence text, two additional properties can be added to a sequence, keys and word links.

Keys

A key can be viewed as a label that is attached to some part of the sequence text. This is especially useful for more qualitative analyses, for example to explore characteristics of particular texts. Such analyses may serve to develop a coding scheme, that can be used to code the different events in a next phase. Keys can only be assigned to (a string of) whole words, not to only a part of a word.

Word links

A word link is simply the relation between two words. Two words in the sequence text can be linked by a line between both words, either without arrowheads, one arrowhead or two arrowheads. Like keys, word links are especially useful for more qualitative analyses.

EVENTS

Each event of a sequence is described at least by an *event code*. The number of event codes in a sequence defines the length of the sequence. If you delete events from a sequence, or add events to a sequence, the sequence variable SEQSIZE is automatically updated.

Events may be 'checked'. If an event is checked, a check mark ( ) appears in front of the event code.

In addition to event codes, an event can also have event variables, and event texts.

Event codes

Event codes are used to describe (or code) pieces of behavior, situations, episodes, etcetera; in SEQUENCE VIEWER these are all called *events*.

You can describe an event by up to 9 different code variables. For example, in observing verbal utterances in interviewer-respondent interaction, an elementary description would consist of two code variables: the actor (e.g. 'I' for interviewer and 'R' for respondent) and the kind of exchange (e.g. 'Q' for question, 'A' for answer and 'R' for repetition of an utterance of the other actor. A simple sequence can thus be described as:

IQ RR RA

meaning that the interviewer poses a question (IQ), the respondent repeats the question (RR) and then answers the question (RA). The full code of an event (like IQ or RR) is called the *event code* and should be distinguished from a code like 'I' or 'R', being a value on only one code variable. An event code on the other hand, is a string of codes, in a prescribed order; the first character of the string is always the code on the first code variable, the second character is the code on the second code variable, etcetera.

Code variables can be used to describe all kinds of sequences. For example, different code variables can be used to describe life events. For example, an event code like 'F0S0' may describe that a person follows fulltime education ('F' on the first code variable), has no job ('0'), is single ('S'), and has no children ('0'), whereas 'PPM1' describes that the person follows a part-time education ('P'), has a part-time job ('P'), is married ('M'), and has one child ('1').

Event variables

In addition to event codes, a number of numerical event variables can be added to each event.

For example, in case events consist of verbal utterances, you can add the number of words of the utterance to each event. You can also add time information to events, i.e. the onset time of an event and the offset time of that event.

In the example of life events, onset times may be the number of years (or months) when a transition takes place.

Event times

'Event times' is the short name for both 'Onset time' and 'Offset time'. These two event variables always appear together; you cannot have 'onset time without 'offset time' or vice versa. To add or delete both variables, you should use **Event variables > Add event time**, or **Event variables > Delete event time...**. If you add event times, the 'Onset time' and the 'Offset time' are always the first two event variables.

Event text

If a sequence has sequence text, each paragraph of the sequence text can be associated (or *linked*) with an event code: the *event text*. For example, the text of a question by the interviewer in the example above, is associated with the event code IQ. An event text should always consist of exactly one paragraph. If the number of event codes in a sequence does not equal the number of paragraphs of the sequence text, they cannot be linked. See also Linked codes and text.

In summary:

A SEQUENCE VIEWER file consists of sequences. Each sequence consists of a number of ordered events. Sequences may have sequence variables and sequence texts, whereas sequence texts may have keys and word links.

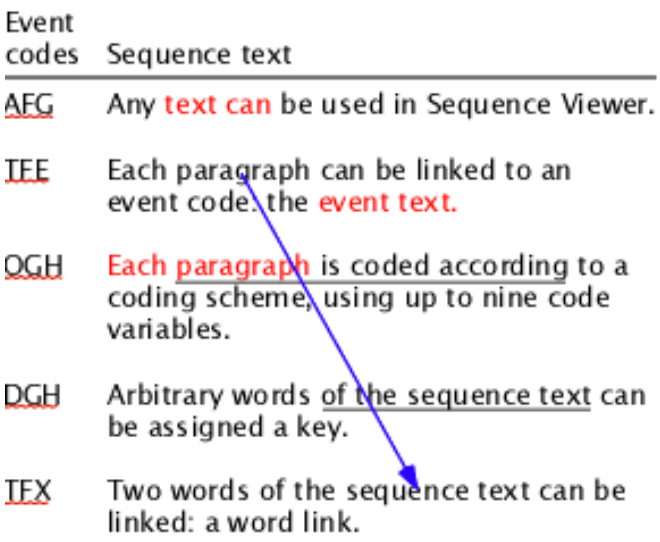
Events consists of event codes and may also have numerical event variables. Each paragraph of the sequence text may be linked to one and only one event code.

STRUCTURE OF A SEQUENCE

Consider the coded sequence:

AFG TFE OGH DGH

Because this sequence has four events, the value on the sequence variable SEQSIZE is 4. Each event code in this example consists of three code variables. For example, the first event code (AFG) is coded 'A' on the first code variable, 'F' on the second code variable, and 'G' on the third code variable. Suppose the sequence text consists of four paragraphs; thus code and text can be linked; see picture below.



The event text of the first event of this sequence, that is, the first paragraph of the sequence text ("Any text can be used in Sequence Viewer."), is coded as AFG, the second event text as TFE, etcetera.

In addition to the event codes, two different keywords are used to characterize parts of the texts, 'red', and 'underscore' in the example above. The 'red' key occurs two times in the sequence text. You may note that 'event text' and 'Each paragraph' belong to the same key, although they are part of different events. The 'underscore' key also occurs two times. You may also note that keys may overlap: the word 'paragraph' belongs to both keys. (In the actual sequence text, keys are not displayed with colors and underscores, but in a different way.)

Finally, a word link is added from 'paragraph' to 'sequence'. In the actual sequence text, word links are also visible as arrows, like in the example above.

Now suppose you have also added an event variable, describing the number of words of each event text. This variable would have the values 8, 12, 15, 11, and 12 on the respective events.

SEQUENCE VIEWER FILES

A SEQUENCE VIEWER file contains all information with respect to values on sequence, event and code variables, sequence texts, keys and word links. In addition, sequences can have other types of information, which are helpful for analyses. For example, sequences can be marked, can have extra information (in five information fields), they can be linked to audio or video files, and much more. Such information, that can be different for each sequence, is called *file data*.

In addition to data, a file also contains general information about the file, e.g. the number and names of sequence, event and code variables, the time unit, predefined sequence definitions, and much more. Such information is called *file properties*.

1.3 Numerical data

INTRODUCTION

Values of sequence variables and event variables are numerical data. A distinction is made between three variable types:

- numbers
- dates
- time stamps

NUMBERS

The maximum number of characters for numbers (including decimal point and/or minus sign) is 12. Whole numbers (or integers) thus range between -99999999999 and 999999999999.

The range of floating point numbers, depends on the number of decimal digits. For example, if the number of decimal digits is 3, the variable can range from -9999999.999 to 99999999.999. The maximum number of decimal digits for floating point variables is 6.

If as a result of calculations, a value will be assigned that exceeds these limits, that value obtains a missing value, or 'M'. A missing value is also assigned if a calculation cannot be performed, e.g. a division by 0.

All internal calculations within SEQUENCE VIEWER are performed with so-called double precision (64 bit floating point) numbers. By and large, if (temporary) internal results exceed 16 digits, the eventual result may not be exact. As a simple example, calculating $9800^4 + 1 - 9800^4$ doesn't yield 1, but 0 instead (although $9800^4 + 2 - 9800^4$ will give 2).

DATES

Dates are just whole numbers, but are assumed to represent the number of days since January 1, 1970. For example, 15393 represents February 23, 2012. The main advantage of the date type is that you can view and enter values on a date variable in the date format, e.g. as 'Feb 23, 2012' instead of '15393'. Because dates are whole numbers, it is easy to calculate for example the difference between two date variables, which will give you the number of days between both dates.

The minimum numerical value of a date is 0 (Jan 1, 1970), whereas the maximum is 115740 (Nov 20, 2286). Dates exceeding these values (e.g. if you try to subtract 10 days from Jan 3, 1970) are assigned a missing value.

Date formats

SEQUENCE VIEWER can recognize a number of different human readable date formats:

mm/dd/yyyy (e.g. 12/17/2011)
yyyy-mm-dd (e.g. 2011-12-17)
month dd, yyyy (e.g. dec 17, 2011)
dd month yyyy (e.g. 17 dec 2011)

'dd' means the number of the day in a month, 'mm' the number of the month, and 'yyyy' the year. 'month' means the name of the month, either the whole name (like 'February'), or a three character abbreviation ('feb', or 'Feb'). The year should preferably consist of four digits, to avoid confusion, but also a year of two digits is recognized; if a two-digit year is less than 70, it is recognized as a year from the 21st century. For example: '12/17/69' is recognized as 'December 17, 2069', and '12/17/70' as 'December 17, 1970'.

Days and months less than 10, e.g. 8, can be written either as '8', or as '08'. For example 'May 8, 1989' is equivalent to 'May 08, 1989'.

TIME STAMPS

Time stamps represent the number of seconds since January 1, 1970 00:00:00 GMT. Like dates, the main advantage is that you can view and edit time stamps in human readable format, for example 'Feb 23, 2012 14:15:30'.

The minimum time stamp is 0, or January 1, 1970 00:00:00, whereas the maximum is 9999999999, or Nov 20, 2286 17:46:39. Time stamps exceeding these values are assigned a missing value.

Time stamp format

The general format of the time part is:
hh:mm:ss [AM, PM]
(that is hours:minutes:seconds). For example:
15:35:12
The time can be followed by 'AM' or 'PM' (preceded by a space). A time like '03:35:12' is the same as '03:35:12 AM', whereas '15:35:12' is the same as '03:35:12 PM'.
A time like '15:35:12 AM' is equivalent to '15:35:12', but '15:35:12 PM' will yield an error. The number of hours should be less than 24, e.g. 24:00:00 will yield an error.
If the number of hours, minutes or seconds is less than 10, e.g. '6', you can write either '6' or '06'.
If you don't specify the seconds, ':00' is assumed; for example, '15:35' is equivalent to '15:35:00'.

The date and the time part are separated by a space, hence the format of date & time may look like:
2011-12-17 15:35:12
It is legal to specify the date only; for example, 'Mar 14, 2007' is equivalent to 'Mar 14, 2007 00:00:00'. Date and hours only (e.g. 'dec 12, 2011 03') will yield an error however.

CONVERSION BETWEEN VARIABLE TYPES

In the DEFINE SEQUENCE VARIABLES and DEFINE EVENT VARIABLES setup windows, you can set the variable type of a particular variable. You should be aware of the fact that if you change a variable type, the values in your file themselves are not changed. Suppose sequence variable SEQVAR4 is of type 'date', and has the value 15351 (Jan 12, 2012) on a particular sequence. If you change the variable type to 'time stamp', this value will not be changed, but will now represent 'Jan 1, 1970 4:15:51' (which is 15351 seconds since January 1, 1979 00:00:00). Worse, if you change the variable type from 'time stamp' to 'date', chances are good that all values will exceed the maximum value for dates, and are hence assigned a missing value.
So how to convert for example a 'time stamp' variable to a 'date' variable? This is the safest solution:

Create a new sequence variable of type 'date'.
Open the TRANSFORM SEQUENCE VARIABLE setup window (**Sequence variables > Transform...**).
Select 'date' from the 'Transform' pop-up menu.
Enter the time stamp variable you wish to convert in the field below 'Transform variable'.
Enter the new variable in the field below 'Assign to variable'.
Press the Transform button.

If you have event times ('Onset time' and 'Offset time'), it is not possible to change the variable type of these two event variables and the sequence variable STARTTIME in this way. The reason is that the event times and STARTTIME are related, and should never have a different time unit. You can change the variable type of all these three variables at once, with the **Time** tab of the FILE SETTINGS. With this method not only the type of these three variables is changed, but also their values, to preserve the date information. For example, if you change a date of 16883 (March 23, 2016) to a time stamp, the new value will become 1458691200 (March 23, 2016 0:00:00).

MINIMUM AND MAXIMUM VALUES

A distinction should be made between the minimum and maximum values allowed by SEQUENCE VIEWER, and the minimum and maximum values defined by the user (see

Define sequence variables and Define event variables). The user defined minimum and maximum are primarily used for data entry: if you enter a number exceeding the user defined maximum, you are warned. On the other hand, if for example because of the result of a calculation, a particular value exceeds the user defined maximum, but not the maximum allowed by SEQUENCE VIEWER, it will just be stored. Only values that exceed the limits of the minimum and maximum allowed by SEQUENCE VIEWER, will obtain 'M' for missing value.

1.4 Sequence window

INTRODUCTION

The main data of your sequence file (the sequences) appear in the main sequence window. This window has three different modes, showing your data in three different ways.

CODES AND TEXT MODE

If your primary data consists of texts (like transcribed interactions), these texts are made visible (and can be edited) in the *codes and text* mode. This mode is activated if you click at the  button in the *Layout* section.

In this mode the main field of the window consists of two parts: at the right are the texts, and at the left the codes that are assigned to these texts. To the left of the code field there is some space to check particular event codes with a check mark (✓; see Checking event codes).

CODES ONLY MODE

A second mode is the *codes only* mode, which can be activated with the  button. In this mode the main field of the window shows just the codes (preceded by a check mark, if the event is checked). This mode is especially useful, if you are only interested in analysing codes, because it shows a large number of codes at once. It is also useful to study the effects of a sequence definition.

CODES AND EVENT VARIABLES MODE

The third mode shows *codes and event variables*. Click at the  button to activate it. This mode shows the codes and all event variables belonging to an event.

In the *codes and text* mode and the *codes only mode* one event variable can be shown in addition to the codes. You can select this event variable with the 'Show event variable' pop-up menu in the *Layout* section.

1.5 Audio and video

You can link audio or video files to the sequences in your data file. Each sequence can have one or more linked files. For example you may have linked both an audio and a video file to a sequence. The name(s) of the linked file(s) appear in the 'Select AV file' pop-up menu of the

Linked files section in the main sequence window. Use this pop-up menu to select a particular file.
You can link separate audio or video files (AV files for short) to each sequence, but you can also link a long AV file to a number of sequences. In that case, a particular segment of the AV file is linked to a sequence.

1.5.1 Linking

LINKING AN AV FILE

To link audio or video (AV) files to sequences, you should check **Linked AV files** in the **File info** tab of the FILE SETTINGS window (**File > File settings...**).

To add a linked file, click at the 'add new link' button in the *Linked files* section and select the file from the Open file dialog box.

Tip
Usually you will only link audio or video files. If you select 'Any file' in the 'Open file dialog box', you can also open different types of file, for example pdf files, or pictures, provided the file can be read by Quicktime. If QuickTime cannot read the file, you obtain an error message, and the linked file will be closed.

You can add up to four different AV files to a sequence.

AUTO LINK FILES

To automatically link the linked files of a sequence to a set of other sequences in your file, click the 'paste link' button. The linked files of the present sequence are copied and pasted to a set of sequences.

You have four possibilities:

Select **All sequences** if you want to set the linked files of all sequences to the linked files of the present sequence.
Select **All marked sequences** if you want to set the linked files of all marked sequences to the linked files of the present sequence. This radio button will be disabled if there are no marked sequences.
Select **Sequence from** to specify a range. By default, the first number is the number of the next sequence, whereas the second number is the number of the last sequence.
Select **Sequences if** to assign the linked files to the sequences with a particular value on a sequence variable. To select this variable, click at the name of a sequence variable in the sequence variables drawer of the main sequence window. The value will be the value from the present sequence on this variable.

Please note that all existing links are replaced.

UNLINK FILES

To unlink a file, first select the name of that file from the 'Select AV file' pop-up menu in the *Linked files* section. Then click the 'unlink' button.

Now you have five possibilities:

Select **Sequences #** (where # is the number of the presently open sequence) if you want to delete the linked file just from this sequence.
Select **All sequences** if you want to delete the linked file from all sequences.
Select **All marked sequences** if you want to delete the linked file from all marked sequences. This radio button will be disabled if there are no marked sequences.
Select **From Sequence:** to specify a range. By default, the first number is the number of the next sequence, whereas the second number is the number of the last sequence.
Select **Sequences if:** to delete the linked file from the sequences with a particular value on a sequence variable. To select this variable, click at the name of a sequence variable in the main sequence window. The value will be the value from the present sequence on this variable.

See also:
Edit event time.

1.5.2 Player window

If the name of a linked file is selected from the pop-up menu in the *Linked files* section of the main sequence window, a common QuickTime window (the player window) appears. You can play (and stop) the sound or the video with the triangle at the left of the slider.
If you click at the grey triangle ▶ before 'more' in the lower left corner, a number of additional options become available.

You can fine tune the position of the playhead, using the precision slider.
You can use the stepper control below 'speed:' to play the sound or video at different speeds. If the speed is negative, the sound or movie plays backwards.
Below 'play from:' you find two buttons, that enable you to quickly start the movie at different start points. Click to start the movie at the beginning of the sequence, and to start the movie at the previous start.
With the two buttons below 'move slider' you can move the slider a particular amount of time backwards ( ) or forwards ( ). You can change this amount of time with the 'move slider' pop-up menu. The time unit of the present SV file is indicated with a bullet.
Click Find sequence to find the sequence, corresponding to the position of the playhead (if the playhead points to a movie or audio fragment that is linked to a different sequence). SEQUENCE VIEWER finds this sequence using the STARTTIME of each file).

You can resize the window to increase or decrease the rectangle wherein the movie is playing.

1.5.3 Playing sound

You can play sound or movies of linked files in a number of ways.

First, you can use the player window as described in the previous section.
Second, you can click somewhere in the text field of the main sequence window, provided that text and codes are linked, and the player window is open (the cursor will change into a loudspeaker).
If onset and offset times are added, clicking at a particular event in the text field, will play the

sound corresponding to that event.
To stop the sound, you may either click at the event text of the same event, or just somewhere in the text field (depending on the setting in the **Playing** tab of the FILE SETTINGS).

See also:
USB device.

1.6 Output

Results of analyses, etc. appear in the output window. The contents of the output window are lost when you quit the program.
If you increase the width of the output window, you will observe that the right part of the output window is gray. The white part represents the page when you print the output. If you are using SEQUENCE VIEWER the first time, or want to print the output file with a different paper type (e.g. changing from A4 to US letter), you are advised to select **File > Page setup...** before you perform analyses, because the 'page' in the output window is adjusted to the size of the paper type.
Some analyses may produce output that exceeds the right margin of the 'page', e.g. the 'Tree' analysis. In such a case you can use 'Page setup...' to set the orientation to 'landscape' and/or set the 'scale' to a lower percentage.
You can also reduce the left and right margins of the print settings to increase the width of the print area.

To find a particular text in the output, type this text in the field after 'Find' and click the magnifying glass.

1.6.1 Save

You can save the output as a RTF (Rich Text Format) file with the Save button  at the top of the window. This format preserves text font, text size, etc. and can be read by most word processors. However, graphics (like the graph created with Transition graph) cannot be saved within the output file. You can save them as separate picture (PNG) files to your hard disk and then copy them into the saved output file.

1.6.2 Print

The content of the output window can be printed using **File > Print output...**.
Alternatively you can directly give a print command from the output window itself, using the Print button  at the top of the window.
Before you print, you should use **File > Page setup...**, e.g. to adjust the type of paper.
Please note that the page selection (e.g. "from page 1 to 12") in the standard Macintosh Print dialog box, does not have effect; all pages will be printed.

1.7 Linked codes and text

INTRODUCTION

Events may consist of event codes, event variables and event texts. Event codes appear in the code field at the left of the main sequence window. The number of event codes defines the number of events.

Event codes and event variables are always 'linked'. That means that the values on the event variables always belong to a particular event code. If you delete an event code, all corresponding values of event variables are deleted too. If you add an event code, all event variables for this new event obtain a missing value.

All texts of a sequence appear in the text field of the main sequence window. SEQUENCE VIEWER uses the carriage return as the delimiter for event texts. That means that each paragraph belongs to a different event, provided event texts and event codes are linked: the first paragraph of the text in the text field belongs to the first event code; the second paragraph to the second event code, and so on. Editing the text in the text field (e.g. by pressing the return key), may break this correspondence.

LINK MODES

Three different modes of the link between code and text of a sequence are discerned. The mode can be different for different sequences in the same file.

The unlinked mode

In this mode the link button looks like:  . You can edit the text field in this mode: enter text, delete text or split and join paragraphs (by typing or deleting a carriage return). If a sequence is unlinked, some types of analyses are not possible on this sequence, or will result in missing values. For example, if you want to assign the number of words of the event text to an event variable, this is not possible for sequences with unlinked text and codes (as it is unclear to which event a paragraph belongs).

The linked mode

To link event code and event text of a sequence, just click at the link button. The link button will change too:  , provided that the number of event codes equals the number of paragraphs in the text field. You cannot edit the text field in this mode.

The bad link mode

If the number of event codes does not equal the number of paragraphs in the text field the link button will look like:  . You can edit the text field (e.g. to correct the cause of the bad link).

You can link codes and texts of all sequences at once, with **Sequences > Link code and text...**. You can then mark all sequences with bad links using **Sequences > Mark links...**, and then find the marked sequences to find out what's wrong with the link.

1.8 Marking, selecting and locking

There are a number of ways to restrict your analyses or data manipulations to a subset of sequences.

MARKING SEQUENCES

First, you can mark sequences and specify that analyses or data manipulations are only performed on the marked sequences. Eight different sets of marked sequences are available. In this way you always have up to eight different subsets of your sequences immediately available for analyses or data manipulation.

You can select a particular subset using the 'mark set' option menu in the upper right corner of the main sequence window. The name of a mark set can be changed with **File > File settings... > Customize**.

You can mark or unmark a sequence manually, by clicking at the mark checkbox in the upper left corner of the main sequence window. In addition there are quite a few other number of options available:

- You can use **Mark all** or **Unmark all** from the **Sequences** menu
- **Toggle marks** from the **Sequences** menu marks the unmarked sequences and unmarks the marked sequences.
- **Mark locked** marks the locked sequences (see below).
- **Mark up** marks the sequences from the present sequence until the last sequence, whereas **Mark down** marks the first until the present sequence.

In addition a number of other mark options are available from the **Sequences** menu. See Mark random, Mark links and Copy marks.

Finally, you can mark and unmark sequences using particular values on code variables, event variables, sequence variables or strings of text in the event texts.

SELECTING SEQUENCES

Secondly, you can specify a particular subset of sequences, based on the values of sequence variables and/or a particular mark set. Such a selection is especially useful if you want to perform a series of analyses on the same subset of sequences. See select sequences for more information.

LOCKING SEQUENCES

Thirdly, you can lock or unlock sequences. If you lock a sequence, its main data cannot be changed, that is, the codes, the event texts, the sequence variables, the event variables, the keys, the word links and the five information fields. For example, if you use **Sequence variables > Compute value...**, values of sequence variables will only be changed if the sequence is unlocked.

All other characteristics of a sequence can be changed if a sequence is locked. For example, you can mark or unmark a locked sequence, or link an audio file to a locked sequence.

You can lock a sequence using the lock at the upper left corner of the main sequence window; just click at the lock to lock or unlock the sequence. You can also use **Lock** from the **Sequences** menu to lock a number of sequences at once.

1.8.1 Select marked sequences

You can select a set of sequences using the marked property of a sequence. To perform analyses or operations on the marked sequences only, click at the select marked sequences icon  . This icon (if available) will always be found at the bottom of 'setup windows' (windows to specify a setup for analysis or data manipulation). The icon changes into  , to signify that a selection is applied. The selection icon  (if available; see select sequences) becomes disabled: you cannot both select a set of sequences using select sequences and select marked sequences at the same time.

1.8.2 Select sequences

You can select a set of sequences using particular values on sequence variables and the marked property of a sequence. To open the SELECT SEQUENCES window, option-click at the selection icon, in the lower part of the window. The selection icon looks like this:  and is only available, if selecting sequences makes sense. It will always be found at the bottom of the setup window. You can use values on one or two sequence variables and/or a particular mark set to select sequences. You should first check or uncheck the appropriate checkboxes.

If  **Use sequence variables for selection** is checked:

Click in the field after 'Select if:' to select this field.
Click on a sequence variable in the left field. The name of the selected variable appears in the selected field.
Select '=', '>', '<', '≥', '≤' or '≠' with the option menu and type a value after this menu. Instead of typing a value, you can also select a label with the pull down menu 'labels'. The corresponding value is automatically inserted.
To select a second sequence variable, first select 'and', 'or' or 'xor' from the option menu.

If  **Use mark set for selection** is checked:

Select  **Marked sequences** or  **Unmarked sequences** if you want to restrict the selection to these sequences. These radio buttons are disabled if no sequences, or all sequences are marked.
You can use any mark set you want, using the pop-up menu after the checkbox. The default mark set (shown by the 'mark set' pop-up menu in the upper right corner of the main sequence window) is not affected.

In addition:

If you change the selection criteria, the number of sequences to be selected is automatically calculated, unless the number of sequences is very large (and the calculation will take some time). In that case you can click on the calculator button  to calculate how many sequences will become selected.
The Select button applies the selection criteria and closes the window.
The Apply button applies the selection criteria, but does not close the window. So you can use this button to select sequences with different selection criteria for next analyses without having to reopen the Select sequences window.

The Clear button clears the selection criteria. It does not affect the actual selected sequences (that is, all sequences will be selected) until you click at the Apply or Select button.

To restore the previously applied selection criteria, you can click at the  icon. Again, to effectuate the criteria you should click at the Apply or Select button.

Applying the selection criteria assigns an invisible selection mark to each selected sequence, which is used to select these sequences for analyses. In this way you can use the same selection for a number of subsequent analyses, whereas the program does not need to perform the selection calculations.

You can infer from the appearance of the selection icon whether or not a selection is made. If all sequences are selected (no selection criteria are applied), the small arrows are blue and somewhere in the lower part of the window the text "All sequences selected" appears. If a selection is made, the arrows are red:  and you are informed about the number of selected sequences. Just click on the selection icon to switch between both states.

The program keeps track of changes in your data file that possibly affect which sequences will be selected. For example, you may have changed the value of a sequence variable that is used as selection criterion. If possible, the program adjusts the selected sequences. If not, the selection is reset to all sequences. The small red arrows of the selection icon become blue and the text stating the number of selected sequences is changed to "All sequences selected". You have to open the SELECT SEQUENCES setup window and make the selection again.

1.9 Checking event codes

Events can be checked in order to perform analyses on checked events only. Just click in front of an event code to check that event. Checked event have a check mark (✓). Click on a check mark to uncheck the event.

With **Sequences > Check events**, you can check all events, uncheck all events, or toggle the event checks.

In the main sequence window, you can see the number of checked events of a sequence at the left side of this window. For example '3 ✓' mean that three events are checked.

To perform analyses on checked events only, click at the 'Checked events only' icon  (if available) at the bottom of the setup window. The icon changes into  , and you are told how many checked events are selected for analysis in the lower left corner of the setup window.

1.10 Editing

You can manually change the values of code variables, event variables and sequence variables, or the sequence text in the main sequence window.

1.10.1 Texts

To edit the event texts in the main sequence window, you should first unlink the event codes and event texts (). Don't forget to link them again after you are done with editing.

To find and replace event text, see Find text.

ENTERING SYMBOLS

You can insert a number of symbols that are common for transcriptions of conversations. The easiest way to insert symbols is by means of the 'Format text' part of the toolbar. If the 'Format text part' is not visible, choose **Edit > Toolbar > Show format text**. For example, to insert an up arrow, just click at ↑. Arrows can also be entered with the arrow keys on the keyboard: hold the option key down and press one of the arrow keys.

TEXT STYLE

With the 'Text style' pull down menu from the toolbar, you can change the text style of a piece of the sequence text: plain, bold, italic, underline or strike out.

To remove a text style, either select 'plain' or select the text style you wish to remove. For example, suppose a particular word is both italic and underlined. To remove the underline, but keep the italics, just select 'underline'.

To apply a text style, it is not necessary to unlink event codes and event text first. If event codes and text are linked, you cannot change the text, but you can select pieces of text, and apply a text style.

Text styles can only be applied to the sequence text. All other texts, including the information fields, only allow plain text.

WORD LINKS

- Word links are connections between two different words of the sequence text. To create a word link:
- Check  **show word links** in the word links section of the main sequence window.
- Click at the Add word link button.
- Click the mouse at the first word, and move, with the mouse still down, to the second word. An arrow connects both words.
- Leave the mouse at the second word.
- Move the mouse over the arrow. The arrow obtains an outline. Now double click at the arrow.
- A drawer appears that allows you to change the color and type (0, 1 or 2 arrow heads) of the arrow, and to add annotations to the word link.

You can check the annotations button  at the left of the main sequence window to show a sticky. If you move the mouse over the arrow of the word link, the annotations of the word link (if any) appear on this sticky.

1.10.2 Codes

See Edit sequences.

1.10.3 Sequence variables

You can quickly edit the values of sequence variables, using the *Sequence variables drawer* (click at the ▼ icon before 'vars' in the lower left corner of the main sequence window to show the sequence variables and their values). Just click at the value of a sequence variable and edit the value.

Another way to replace a particular value of a sequence variable by another value, is to click at the name of a sequence variable in the Sequence variables drawer. The CHANGE SEQUENCE VARIABLES setup window appears.
The value of the sequence variable for the present sequence appears after 'Change value': (you cannot edit this value). Enter the new value in the field after 'of [name of sequence variable] to:'. You can choose to replace the value of the present sequence only, or to replace the selected value of all sequences.
If you go to a different sequence, the value of the selected sequence variable is automatically adjusted.

See also replace values of sequence variables for even more options.

Sequence variables of type 'date' and 'time stamp can be viewed and edited in date or time stamp format (e.g. 'Jan 12, 2012'). If you select such a variable, a checkbox ☑ **Edit as date/time stamp** becomes available in the CHANGE SEQUENCE VARIABLES setup window that, if checked, turns the numerical value into human readable date format. See also Numerical data.

1.10.4 Event variables

You can edit the values of event variables using the *Codes and event variables* mode of the main sequence window. Just click at the value of an event variable and edit the value.
If the event variable is of type 'date' or 'time stamp', the value of the event variable also appears in date, resp. time stamp format in a separate edit box in the lower right corner of the main sequence window (in the 'Codes and event variables' mode). This allows you to view and enter these values as dates or time stamps.

Values on event variables can also be changed with Edit sequences.

Event variables of type 'date' and 'time stamp can also be viewed and edited in date or time stamp format (e.g. 'Jan 12, 2012') in common setup windows. To this end a checkbox ☑ **Edit as date/time stamp** is available in the setup window that, if checked, turns the numerical value into human readable date format.

1.11 Enter event times

INTRODUCTION

SEQUENCE VIEWER offers a number of options to ease entering onset and offset times. As an alternative to the procedure described in this section, see also Assign event times, using waveforms to enter event times.

You should have transcribed and coded the texts (but the codes may consist of empty codes, that is dashes) and you should have added onset and offset times to your file (**Event variables > Add event time**). Be sure that event codes and event texts are linked and that the player window of the linked sound file is open. The mode of the main sequence window should be codes and text.

Add offset times

Select 'Offset time' from the 'Show event variable' popup menu in the *Layout* section.

- Check the 'Assign event time' button ().
- Play the sound from the player window.
- At the very moment the sound is at the end of the first event text, click the mouse at the first event code in the codefield of the sequence window. The offset time is changed to the present time of the sound.
- If the sound is at the end of the second event text, click at the second event code; and so on.

You may observe that the cursor changes into a clock as soon as you enter the code field. This warns you that as soon as you click at an event code, the event time will be changed to the present position of the sound file.

The offset times you entered in this way are usually not perfect yet, but it is easy to fine-tune them. This will be discussed later. First enter all offset times as described here.

Add onset times

The onset time of an event will be more or less equal to the offset time of the previous event. We can use this to quickly assign approximate onset times.

- Select **Event variables > Compute event variable...**.
- In the COMPUTE EVENT VARIABLE setup window, enter 'Onset time' in the 'Compute variable' edit box and type 'Offset time[EVTNO-1]' in the equation field.
- Click Compute.

Now the offset time of each event is assigned to the onset time of the preceding event ("EVTNO-1"); see Compute event variables for more information about indexing.

The very first event of a sequence will have a missing value, because there is no preceding event. The onset time however should equal the STARTTIME of the sequence if the event time is absolute, or 0 (zero) if the event time is relative (see Absolute and relative time).

Fine-tuning onset times

Be sure that event codes and event texts are linked and that the player window of the linked sound file is open.

- Select 'Onset time' from the 'Show event variable' popup menu in the *Layout* section.
- Check the 'Assign event time' button ().

- Click at the text of an event. Don't click event codes in the code field! The sound will now start playing. It is easy to hear whether the onset time is to low, to high, or exactly right. Use the up and down arrow keys to adjust the onset time. For example, if the sound starts a bit too late, press the down arrow key two times and click once again at the text of the event.

Note: You cannot use the up and down arrows while the movie is playing.

Fine-tuning offset times

Be sure that event codes and event texts are linked and that the player window of the linked sound file is open.

- Select 'Offset time' from the 'Show event variable' popup menu in the *Layout* section.
- Check the 'Assign event time' button ().
- Hold the option-key down and click at the text of an event. Don't click codes in the code field! The sound starts playing and will end at the offset time. Use the up and down arrow keys to adjust the offset time.
- In addition to holding the option-key down, you can also hold the shift-key down. Now only the last ten time units of the event will play.
- You can use the **Playing** tab of the FILE SETTINGS window to use the capslock key instead of holding down the option and/or the shift keys.

The 'Assign event time'  button becomes unchecked if you go to another sequence. To prevent this, you may check  **Keep auto time checked if new sequence is opened**, available from the **Times** tab of the FILE SETTINGS window. Be sure to uncheck this button if you stop editing event times, to prevent that you change the event time if you click at an event code.

1.12 Quick keys

The following short cuts are available in SEQUENCE VIEWER.

To move between sequences:

- shift-left arrow: go to previous sequence.
- shift-right arrow: go to next sequence.

Change onset and offset times:
(if the 'assign event times' button  in the main sequence window is checked)

- Up arrow: subtracts one unit from the event time.
- Down arrow: adds one unit to the event time.
- Click at an event code: assigns present time of sound file to onset or offset time (depending on which one is selected with the 'Show event variable' pop-up menu in the *Layout* section).

To play sound and movies:

(if event times are available and event texts and codes are linked)

- Click at an event text to play linked sound.
- Option-click at an event text to stop the sound at the end of the event.
- Shift-click at the event text to play the last 10 time units of the event.
- Press option-spacebar to play or stop the sound, whether or not event texts and codes are linked or event times are available.

To enter STARTTIME:

- Option-command click at the time (lower right corner of the player window) to put that time in the STARTTIME field of the sequence variables drawer.
- Alternatively you can type a 't' holding both the option and command keys down.

If you want to change the value of the STARTIME, and event times are relative to the STARTIME, see the warning at the discussion of absolute and relative event times.

To open 'EDIT SEQUENCE' setup window"

- Shift-click at a code in the code field.

1.13 Preferences

Preferences concern general settings that apply to the application. Preferences are stored in the preferences file and are saved as soon as you make any changes to your preferences. The preferences file is called 'SV5prefs', and is stored in the folder 'Sequence Viewer 5 preferences' (the same names are also used for SEQUENCE VIEWER 6 files). This folder can be found in the folder 'Preferences' within your Home library folder. If this file becomes corrupted for one reason or another, you can safely trash it (if SEQUENCE VIEWER is not open). If you restart SEQUENCE VIEWER, a new preferences file is created (and you are asked to personalize SEQUENCE VIEWER).

You can open the PREFERENCES window with **Sequence Viewer > Preferences...**.

1.13.1 View

If ☑ **Live scrolling through sequences** is checked, moving the slider in the main sequence window immediately displays the sequence. If it is unchecked, the sequence is displayed as soon as you leave the mouse. Live scrolling may slow down scrolling through the sequences. For this reason, live scrolling is not effective if the waveform window is open.

The ☑ **Show grids** checkbox governs whether or not grids are shown in the codes only mode and the codes and event variables mode of the main sequence window.

Use the ☑ **Show tooltips** checkbox to turn the tool tips on or off. Tool tips are small descriptive texts that appear as soon as you move the mouse over an object. Tool tips are only available in the text keys and time keys windows, and with some graphics, to show additional information.

1.13.2 Highlight

You can select a highlight color that is used to highlight texts, or lines of a list (for example a list with sequence variables, as usually appear in setup windows concerning analyses with sequence variables). Select one of the two corresponding radio buttons and click at a color to see the effect. Click 'other' to select any other color you prefer.

1.13.3 Files

OPEN FILE OPTIONS

When you open SEQUENCE VIEWER, and if  **Show open file dialog at startup** is checked, an open file dialog is shown as soon as the startup screen is clicked away.

Recently opened SEQUENCE VIEWER files are stored in the preferences file, for quick access to these files (**File > Open recent**).

SAVE OPTIONS

You can automatically save your data file after a particular amount of time is elapsed since the last save.

FILE CHANGES

If you make changes to your data file, the program keeps track of the kind of changes you made. If you close your file (or the application) you are informed about the most recent changes you have made, in order to help you to decide whether or not you should save the changed file. You can select the number of changes the program should remember.

A distinction is made between minor and major changes. Major changes always concern changes in the primary data, that are definitely lost if the file is not saved and cannot easily be restored.

Generally, major changes are generic changes made by the program on a number of sequences at once as the result of some SEQUENCE VIEWER procedure. For example recoding a code variable, computing a sequence variable or deleting a code variable. Major changes can always be undone (see undo).

Minor changes usually concern changes you made manually to your data file. For example changing the event text by typing, or changing a particular code in one sequence manually. Most minor changes can be undone (see undo).

Most analyses can be aborted during execution (press command-period). The progress bar tells you whether or not an analysis can be aborted. If no progress bar is shown, you can usually use command-period to abort the analysis too.

If an analysis is interrupted, you are informed by a message. This is especially important in case the analysis changes your data file, like a compute command. If no message appears, the interrupt was not successful (the analysis was already performed on all sequences), and you can use **Edit > Undo [command]** to undo the last change, where [command] tells you which change will be undone. You can change the number of successive undo's with the pull down menu after 'Number of undo's'.

1.13.4 Analyses

PROGRESS BAR

If analyses take some time, a progress bar shows how the analysis proceeds. Use the slider to decide whether the progress bar should be shown if the analysis is expected to exceed a particular amount of time. Showing the progress bar takes some extra analysis time.

TABLE OPTIONS

Analyses are often presented in the form of a table. The format of a table can be changed with the table options. The following options are available:

- Adjust the spacing of columns.
- Adjust the number of decimal digits in the output.
- Add value labels or code labels.
- Abbreviate value labels. This option is useful for example in tables, when value labels are shown at the top of a column, to prevent that the column widths become too wide.

If appropriate, setup windows have a table option button, which looks like  that can be found at the bottom of the window. This button brings you immediately to **Analyses** tab of the PREFERENCES window.

1.13.5 Print

You can use the sliders to change the print margins. The range of the margins is between 36 and 144 points (0.5 inch to 2 inch).

To print page numbers (and the name of the file and the date in case you print the output), check  **Page numbers** and enter a margin for the header (between 36 and 131 pts). The top margin should always be at least 13 points higher than the header margins.

1.13.6 Editor

The **Editor** tab of the PREFERENCES window provides you with a number of options for the Command Editor.

SCRIPT FOLDER

You should store your script files in a separate folder. Select this folder with Browse. If you click Load in the COMMAND EDITOR, the file selection dialog box will show the contents of this folder.

RECORDING DISABLED CONTROLS

You can choose whether or not to record the setting of disabled checkboxes and edit fields. Not recording disabled checkboxes and edit fields reduces the number of setup lines in the script. Recording disabled checkboxes makes it easier to adjust existing scripts.

RECORDING CUSTOM OR DEFAULT NAMES

Mark sets, sequence definitions and information fields have default names (like 'Mark set 1') that you can change into custom names. If you record a script using a SEQUENCE VIEWER file that has custom names, these names may not be recognized if you use that script for a different SEQUENCE VIEWER file. To prevent this, you can choose to record either the custom names, the default names, or both.

DELAY IN TRACE MODE

In the Trace mode, the setup line that is executed, is highlighted. You can change the delay (in ticks, that is 1/60 of a second) between the execution of successive setup lines.

HIGHLIGHT SETUP LINES IN RUN MODE

In the Run mode the execution of the script is not visible to the user (except that the progress bar may be shown for the more tedious analyses). Checking **Hilite setup lines in Run mode** provides you with feedback about which part of the script is executed. This may be informative in case of long scripts.

1.13.7 USB device

INTRODUCTION

To ease transcription, you can connect your computer to a USB device, i.e. foot pedals. USB devices usually come with a so-called driver, which communicates between the external device and the computer. Such a driver may for example steer the behavior of sounds (play, stop) played via QuickTime player. Or the driver contains software that enables to play the sound by itself.

However, you may find it more convenient to transcribe recorded conversations when playing the sound from within SEQUENCE VIEWER. To this end, the driver should communicate with SEQUENCE VIEWER instead of QuickTime Player. A suitable driver is USB overdrive, downloadable from www.usboverdrive.com (18 euros shareware, but you can test it before buying). If appropriately configured, this driver can communicate signals from an external USB device (like foot pedals) directly to SEQUENCE VIEWER. The driver appears to work fine with a number of external devices, including usually much cheaper game devices.

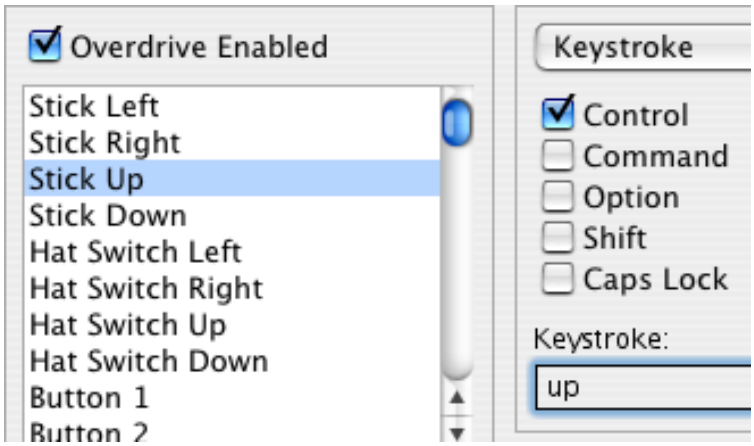
CONFIGURE AN USB DEVICE

We will illustrate how to configure USB overdrive, to use a game device consisting of foot pedals (and a steering wheel) within SEQUENCE VIEWER. The game device we used is the Predator GM 3100R steering wheel (www.trust.com/products/) bought by us for only 30 euros.

For different drivers and/or devices, the procedures may be a bit different. The description below is just for illustrative purpose. The Predator GM 3100R is no longer available, whereas a newer version of USB overdrive is available then the one tested by us. Nevertheless, with the description below, you should be able to set up your own device.

- Connect the external device to your computer.
- Download USB overdrive and install it.
- Open System Preferences in the Finder and click 'USB overdrive' (to be found in the 'Other' row).

- A window appears that can be used to buy USB overdrive. After about ten seconds you are allowed to continue without buying. Click 'Later'.
- With the pop-up menu in the upper left corner, you can select 'mouse' or 'Joystick'. Select 'Joystick'.
- Check the **Overdrive enabled** checkbox.
- Below this checkbox you will find a larger number of possible inputs from game devices, e.g. 'Stick left' or 'Stick right'. You should now figure out which of these signals correspond to the left and right foot pedals.
- Press the right footpedal. 'Stick up' becomes bold (with the Predator GM 3100R steering wheel ; with another device some other signal may become bold).
- Select 'Stick up' (or the signal that became bold) and select 'Keystroke' from the pop-up menu in the upper right corner. Check the **Control** checkbox. Select the field below "Keystroke" and press the 'up' arrow key. See picture below.



- Press the left pedal (in case of the Predator GM 3100R steering wheel the 'Stick down' signal should become bold).
- Select the signal that became bold, check 'Control', select the 'Keystroke' field, and press the down arrow key.
- Be sure that all other signals are set to 'Do Nothing'.

You're done! You can now close USB overdrive and the system preferences. Each time you start up your computer, you will be reminded if you did not pay the shareware fee.

You may now go to SEQUENCE VIEWER, open **Sequence Viewer > Preferences**, and click the **USB device** tab to adjust the actions taken by both foot pedals.

USB DEVICE OPTIONS

You should have a sound (or movie) file open in SEQUENCE VIEWER to use the foot pedals. If you press the right pedal (down and up) the sound starts to play. If you press the right pedal once again, it stops. If you press it again, the sound starts playing, but some time units earlier than you stopped it. You can adjust the number of time units with the slider after 'Move sound pointer backwards before replay'. If you keep the right pedal down, this generates a fast forward; if you leave the pedal, the sound will stop. You can adjust the time you should keep the pedal down before 'fast forward' starts, with the slider after 'Delay before fast forward'. The left pedal is used for fast backward.

2 Files

SEQUENCE VIEWER files consist of sequences, consisting of texts, codes, event variables, sequence variables, keys and word links (see the Introduction for more information).

Within a SEQUENCE VIEWER file, not only the event texts, the codes, the values of the event variables and the values of the sequence variables are stored, but also all additional information about code, event and sequence variables (like labels, etc.). Also the File settings are stored within the file.

The maximum size of a data file is about 4 GB.

2.1 New

To create a new file, select **File > New**. A small window appears asking you after the number of code variables and whether or not you want to add event times to your file.

A SEQUENCE VIEWER file should always have at least one code variable. You can always add additional code variables with **Codes > Add code variable...**, but it is recommended to decide upon the number of code variables you are going to use for your coding scheme, at the moment you create a new file.

You should also decide whether or not you are going to add event times to your file. You can always add event times with **Event variables > Add event time...** afterwards, but it is recommended to do this at the creation of a new file.

After having created a new file, it contains one (empty) sequence. It is further recommended that after you have created the new file, you:

- Define properties (like the allowed codes) of the code variables with **Code variables > Define event codes...**.
- Add at least those sequence variables that permit you to identify sequences (e.g. in case of interview transcriptions, interviewer number, respondent number and question number) with **Sequence variables > Add variable...**.
- Define the sequence variables, e.g. minimum and maximum values, or value labels with **Sequence variables > Define variables...**.
- Select **File > File settings...**, to adjust the settings according to the characteristics of your file (see File settings).

2.2 Open

To open a SEQUENCE VIEWER file, select **File > Open...**. You can choose between SEQUENCE VIEWER 6 files, any SEQUENCE VIEWER file, or any file. The latter choice is added because it may occur that SEQUENCE VIEWER file icons are disabled in the open file dialog window (this sometimes occurs if files are copied using a USB stick, or are sent as an attachment via e-mail). This enables you to open any file (also non-SEQUENCE VIEWER files; in that case you will obtain an error message if you try to open such a file).

You can use the preferences if you want SEQUENCE VIEWER to show an Open File dialog box at start up. Select **Sequence Viewer > Preferences...**, click the **Files** tab and check ☑ **Show open file dialog at startup**. You can also use **File > Open recent**. SEQUENCE VIEWER stores up to ten files that are opened by the program in previous sessions. You can change this number with the **Files** tab of the SEQUENCE VIEWER preferences.

2.3 Save

You can save your file in the usual way with **File > Save**. You can use the **Files** tab of the Sequence Viewer preferences, to save your data file automatically after a particular number of minutes.

The program keeps track of changes you have made, and warns you if you close the file or quit the program, and your changes are not saved yet.

Use **File > Save as...** if you want to save your file with a different name. Note that if you make changes to your data file after you saved the file with a different name using **Save as...**, and then select **File > Save**, the changes are saved to the file you saved with **Save as...**.

In case your computer crashes during a save, the saved file quite probably becomes corrupted. However, before saving a file, SEQUENCE VIEWER always first renames that file on your hard drive with a temporary name, thus keeping the original file untouched. If the file is saved correctly, the original file (with the temporary name) is deleted. Temporary files have the same name as the file that is saved, but this name is preceded with 'SVtmpxxxxx' (where 'xxxxx' is some number) and will be in the same folder as your original file.

In case of a crash during the save, the original file with the temporary name still resides untouched on your hard drive. The new (maybe corrupted) file quite probably resides on your disk too. You can check whether this file is corrupted with **Check data...** from the **Sequences** menu. If you experience any problems, we recommend that you trash the new file and use the original file with the temporary name (you may use the 'any file' option to open the file; see Open).

2.4 Duplicate files

'Duplicate files' reads a file from disk, and makes a number of copies of this file. For example, suppose you want transcribing and/or coding to be done by a number of different people. You can first make a template file, containing the code variables, the allowed codes, dependencies between codes, etcetera. Now you can make a number of copies of this template file, with for example the file names containing the names of the different coders (just type the names of the different coders in the 'Names for copied files' field; see below). This ensures that the different files produced by the coders have the same basic properties, and thus can easily be combined into one file (see Import Sequence Viewer files).

The file that will be copied has to be read from disk. You can read the presently open file from disk, but be sure to save it first, if you wish to incorporate any changes you made after you have opened the file. Select the file that should be copied with the button Select file.

The different copies obtain their names from the list of 'Names for copied files'. Hence, the number of copies equals the number of names in this list. Just type the different file names in the field below 'names for duplicates', each name on a separate line.

It is not necessary to add the file extension 'sv6' to the filenames; these are added by SEQUENCE

VIEWER.

The copied files are all stored in the same folder. You have to select this folder with Get folder.

Tip
If you select a non-SEQUENCE VIEWER file (select 'Any file' after 'Format' in the 'Copy file' dialog box you obtain after clicking Select file), you are warned that the file is not a SEQUENCE VIEWER file. You can nevertheless continue copying. So you can make a number of copies of any type of file (e.g. Word documents) with this command.

2.5 Import

You can:

- Import a file with just codes to create a new SEQUENCE VIEWER file.
- Import a file with values of event variables. The series of event variables of each event may be preceded by a code, which is imported too. This also creates a new file.
- Add SEQUENCE VIEWER files to the presently open file.
- Add sequence variables from a tab delimited data file to the presently open file.
- Add event variables from a tab delimited data file to the presently open file.
- Import texts from a text file and add these to the sequence text of the presently open file.

The first two options are only available if no SEQUENCE VIEWER file is open, whereas the latter four options are only available if a SEQUENCE VIEWER file is open.
If you select **File > Import...**, you obtain a window with these six options. Depending on whether or not you have a SEQUENCE VIEWER file open, some of these options are disabled. Click the appropriate radio button and press OK. Depending on the selected radio button, you obtain a setup window with different options.

2.5.1 Codes

INTRODUCTION

You can read a file with just codes. The program will create a SEQUENCE VIEWER file from these codes. If you already have a SEQUENCE VIEWER file open, you should first close it. The import file should be of type text ('txt'), but SEQUENCE VIEWER can also import other file formats, for example '.docx' files, and convert them to plain text.
Event codes should be space delimited, and each sequence should be on a different line, that is, sequences should be CR delimited. So it may look just like this:

```
Az Cs Xs Cd Hz
Ps Qd Rz Sq Ae Fd Qs
Gc Fs Rz
```

In this case, SEQUENCE VIEWER assumes there are three sequences, the first with 5 events, the second with 7 events and the third with 3 events, whereas the number of code variables is 2.
If the file you wish to import does not fulfill these requirements, SEQUENCE VIEWER offers you a number of options to transform the file. For example, sequence and event delimiters may be different from CR and space; or your 'event codes' consists of strings of different length. For

example:

ABX□AC→DEFG¶
GA→HIJ□KMNJK¶
ACVB□DF¶

where '□' designates a space, '→' a tab and '¶' a carriage return. In this case you may want to use the '¶' as sequence delimiter and both '□' (or space) and '→' (or tab) as event delimiters. SEQUENCE VIEWER allows you to specify this, and transforms the file accordingly. It is very well possible to import any kind of text, and create a file with codes from it!

SETUP

To import a file with codes, do this:

- Select **File > Import...**.
- Select **Import codes** and click OK.
- A large setup window appears, allowing you to open a file, and make the necessary adjustments to this file.
- Click Get file (in the lower right corner) to open an import file. Select the appropriate file format and click Open.
- The content of the file is shown in the large field.
- if the Import button is disabled, you have to make adjustments in order to fulfill the requirements for import. You have two main options to adjust the file.
  - With **Edit text** (the popup menu at the upper part of the setup window) you can edit, find and replace text. These changes are permanent (unless you reload the file of course).
  - With **Use conversions** you can apply a number of transformations to the text. These conversions are non-destructive; e.g. if you check **Remove diacritical marks**, diacritical marks are removed, but reappear when you uncheck this checkbox. Each transformation is immediately applied, so you can inspect the results.
  - If the transformations fulfill the requirements for import, the Import button will become enabled.
- Click Import to actually import the file.

If the Import button remains disabled, move the mouse over the help button. This will tell you (in the help field of the tool bar) which adjustments have to be made.

For a more detailed discussion of import codes, see section 7.3 of the tutorial.

NOTE

Because (after successful import) you imported a file with just codes, the codes and text mode is not available. To make it available, select **File > File settings...** and check **Sequence text** in the **File info** tab.

2.5.2 Event variables and codes

INTRODUCTION

If you have access to continuous data that have already been extracted from another source, such as time stamps, you can import them into SEQUENCE VIEWER as event variables. To import a file with event variables, the format is quite similar to the 'Show codes and event variables' mode from the main sequence window. SEQUENCE VIEWER creates a new SEQUENCE VIEWER file, containing the imported event variables.

In the file you wish to import (which should be in '.txt.' format, although it is possible to import any type of file), the values of the event variables for a particular event should be separated by tabs, whereas events are delimited by a carriage return. A row of event variables may be preceded by an event code (see below), and should be followed by a tab. Sequences should be separated by a character like a '@'. For example (→ means a tab, ¶ a carriage return):

```
13→547→38¶
28→3→572¶
265→23→21¶
@
287→3→8712¶
42→46→98¶
(etc.)
```

Or in case the event variables are preceded by event codes:

```
AVc→13→547→38¶
BCz→28→3→572¶
XAs→265→23→21¶
@
XSd→287→3→8712¶
BCd→42→46→98¶
(etc.)
```

To import the file, select **File > Import...**, than select **Import (codes and) event variables**, and press OK. In the setup window that now appears, a number of options are available to specify the format of the import file. These are described below.

THE SETUP WINDOW

Sequence delimiter

Different sequences should be distinguished by some special symbol (character). You can use either one of the following characters: @, # and &. Select the appropriate delimiter with the 'Sequence delimiter' pop-up menu.

Event codes

A row of event variables can be preceded by a string of one or more codes (the event code). Check **Include event codes** if you have event codes. In that case another option appears, allowing you to specify the number of code variables. All code strings should have the same length, equal to the number of code variables.

If no event codes are imported together with the event variables, SEQUENCE VIEWER creates one code variable, and assigns a dash ('-' for 'not coded yet') to the code of each event.

If you want to import only event codes, without event variables, you better use the 'Import codes' option from the IMPORT FILE setup window, which is much more flexible and allows you to correct for errors in the event codes.

Event variables

You can import up to nine event variables (in addition to the event codes). You should specify the number of event variables in your import file), by selecting a number between 1 and 9 with the pull down menu before 'event variable, including:'.

You are also asked if the event variables either include onset and offset times, only one of these, or no event times at all. If both event times are included, the first event variable (after any event codes) *should* be the onset time, and the second one the offset time. If you import only onset times or only offset times, it should be the first event variable. SEQUENCE VIEWER will add the other time variable (with all values set to 'M' for 'missing value').

In case the first variable is the onset time, it is strongly recommended to check ☑ **Set STARTTIME of each sequence to onset time of first event**.

Importing onset and offset times

If event times are included, you have to decide on the format of the event times. In SEQUENCE VIEWER onset and offset times are always whole numbers. However, in the import file, onset and offset times can also be time stamps, e.g.:

Dec 17, 2011 14:16:13

If the onset and offset times are just whole numbers, and don't represent time stamps, select the radio button ◉ **values with a time unit of:**. The time unit pop-up menu becomes enabled, thus enabling you to select the correct time unit.

Importing time stamps

If onset and offset times represent time stamps, you should select ◉ **time stamps**, giving you another set of options.

A time stamp is just a value, denoting a point in time. A common format is 'Unix time' (or 'POSIX time'), which is the number of seconds elapsed since January 1, 1970 00:00:00 GMT (Greenwich Mean Time, also called UTC, or Coordinated Universal Time).

If the onset and offset times that are imported are integer values that represent this number of seconds, select 'seconds since January 1, 1970' after 'Format of time stamp:'.

The second format (or rather, set of formats) consists of a human readable date and time, like 'August 23, 2008 4:50 PM'. SEQUENCE VIEWER converts such a date and time to the number of seconds since January 1, 1970 00:00:00 GMT, and stores this number in the SEQUENCE VIEWER file. If the import file contains such human readable times, select 'human readable date and time'.

In the import file, a line representing an event with event variables, with onset and offset times and one additional event variable, but without codes, may thus look like this:

12/17/2011□14:16:23→12/17/2011□2:16:28□PM→38¶

where → is a tab and □ is a space.

Time zone

As explained above, SEQUENCE VIEWER converts date and time to the number of seconds elapsed since January 1, 1970 00:00:00 GMT.

You should be aware of the fact that the local time is different from GMT. For example, in The Netherlands in the winter it is one hour later than GMT, whereas in the USA it will be (depending on the time zone) several hours earlier.

This fact poses a problem with the conversion. For example, if you want to convert December

17, 2011 14:16:32, the number of seconds that elapsed since January 1, 1970 depends on the time zone (as it will be December 17, 2011 14:16:32 at different times, at different places in the world). If your time stamps are created by a computer, it is very well possible that local times are used. If you are in time zone -6:00 (neglecting DST, Daylight Saving Time), December 17, 2011 14:16:32 local time, equals December 17, 2011 20:16:32 GMT. You can correct for these differences by subtracting the time zone (including DST) from the imported time stamps, by entering the time zone in the edit box after 'Adjust for local time zone'. Use the format hh:mm, e.g. -05:00. If you leave the edit box empty, it is assumed that the time stamp is GMT.

Usually, if time stamps are generated by a computer, the 'seconds' format is the number of seconds since January 1, 1970 00:00:00 GMT, whereas time stamps in a human readable format (like Feb, 19, 2008) are local times (but this is not necessarily always the case).

Please note that the time zone adjustment does not necessarily equals the time zone you are in at the moment; it is the time zone of the location of the computer at the time (DST should be included in the correction) the data was created.

When to correct for time zone

The advantage of correcting for time zone is that the correction yields an unambiguous result. For example, for an airplane pilot confusion about the exact time may be disastrous. As another example, suppose you want to know the time of the first step on the moon. It's a bit awkward to say "that depends on where you are". Instead, 'July 21, 1969 02:56 GMT' provides an unambiguous answer to such a question.

Nevertheless, adjustments for time zones may sometimes yield results that are a bit confusing. For example, suppose data is collected on a computer in time zone -8:00 (e.g. California), at December 17, 2011 17:20:00. This evaluates to December 18, 2011 01:20:00 GMT, which is a different day.

Moreover, it may be important to know the local time for particular purposes. For example, if the study is concerned with the time of phone calls, it may be important to know the time of the call in local time. On the other hand, if persons in different time zones are called, GMT times may be better suited, for example to preserve the order of the calls.

It is also possible that data is gathered in different time zones. If you want to correct for time zones, it may be better to use different datasets, each with its own correction. The eventual SEQUENCE VIEWER files can easily be merged afterwards (see Import Sequence Viewer files).

Importing the file

If you're done with the settings, click the Get File button, to read the import file.

The file contents are shown in the large text box at the right of the window. Be sure your settings are correct, and click the Check and convert data button. This checks the file for errors. Any errors are shown in the field below 'errors'. If this field remains empty, there are no errors, and you can click the Import button to create a new SEQUENCE VIEWER file from the data.

If errors are encountered, a message appears in the error field, describing the error, highlighting the offending value. Some errors may be caused by wrong settings, which can be corrected (e.g. a wrong number of event variables), whereafter you can press the Check and convert data button once again. Other errors need to be corrected in the import file itself.

In addition to checking the data, pressing this button also triggers the necessary conversions, e.g. converting dates and times to the number of seconds since January 1, 1970, or adjusting for a local time zone.

If checking and conversion are successful, you can move the mouse over converted time stamps for a last check; in the upper right corner of the window, the corresponding (GMT) dates and times will be displayed.

After you have pressed Import, the new file will be created, and opened in the 'codes and event variables' mode.

Time stamps in Sequence Viewer

Time stamps are stored as whole numbers, and thus represent pretty big numbers. Many computer systems use signed 32 bit integers to store dates, that is, a series of 32 ones and zeros. The maximum is 32 ones, or 2147483647. This number of seconds after January 1, 1970 will be reached at January 19, 2038 03:14:08, and thus sets a limit on the range of dates that can be stored on some computer systems. Negative numbers are used for dates before January 1, 1970.

Although SEQUENCE VIEWER can use much larger numbers, onset and offset times cannot be negative, thus event times before January 1, 1970 are not possible. The maximum is set to 9999999999 (November 20, 2286 17:46:39).

See also Numerical data and the discussion about Absolute and relative time in the 'File settings' section.

2.5.3 Sequence Viewer files

SETUP

To add SEQUENCE VIEWER files to the presently open file, do this:

- Select **File > Import...**.
- Select **Import Sequence Viewer file(s)** and click OK.
- Another window appears, allowing you to select one or more SEQUENCE VIEWER files for import.
- Select a SEQUENCE VIEWER file with Get file. The file name appears in the upper field.
- To add a second SEQUENCE VIEWER file for import, click Get file once again. The file name is added to the file list.
- Alternatively, you can use the Get Folder button. All SEQUENCE VIEWER files in that folder (irrespective of the version) are added to the file list.
- Click Add to start the import.

Don't forget to save your file using **Save...** from the **File** menu.

ADDITIONAL INFORMATION

You can only add another SEQUENCE VIEWER file to the presently open file, if both files have essentially the same structure, e.g. the same code variables, sequence variables and event variables. If you select a file for import using the Get file button, this is checked. If both files have a different structure, you are informed about the difference and the file will not be added to the list of files to be added.

If you use the Get Folder button to import all SEQUENCE VIEWER files from a particular folder, all SEQUENCE VIEWER files will appear in the list. However, those files that do not have the same structure, are shown in red. You can select a file name to see the differences between the presently open file and the selected file, in the file info field. To import the files in the file list, the list should not contain 'red' file names; you have to remove those files from the list.

Usually it is easy to change one of both files, to make them similar, e.g. by adding a sequence variable. If the program encounters non-essential differences (e.g. the minimum and maximum values of a sequence variable), the file will be added, and the program keeps the information from the presently open file.

The largest problem arises when the order of variables in both files is different. For example, in file A (the open file) you have event variables VARA, VARB and VARC, whereas the import file has

event variables VARA, VARC and VARB, in that order. In case of sequence and event variables, the solution is simple. You can use **Rearrange...** from the **Sequence variables** and the **Event variables** menus to change the order of the sequence variables or the order of the event variables.

In case of code variables the only way to solve this is to change the order of the variables manually. For example, suppose you have three code variables, VARA, VARB and VARC, in this order. To change the order to VARA, VARC and VARB, do this:

- Create a new event variable (VARD) with **Codes > Add code variable...**. Now you have VARA, VARB, VARC and VARD.
- Copy variable VARB to VARD using **Codes > Copy code variable...**.
- Delete VARB with **Codes > Delete code variable...**.
- Change the name of VARD into VARB with **Codes > Define event codes...**.

2.5.4 Sequence variables

You can add sequence variables to the presently open file from a tab delimited data file. Such files can be created by, for example a statistical program. The first line of the data file should consist of (tab-delimited) variable names. Each next line should represent a different case (sequence) with tab-delimited values.

To add sequence variables do this:

- Select **File > Import...**.
- Select **Import sequence variables** and click OK.
- Open the data file with Get file. The variables field shows the names of the variables from the data file.
- Click at the field below 'Add variables' to select it.
- Click at a variable name in the variables field. The name will be added to the list of sequence variables that are added. If this name already exists in your SEQUENCE VIEWER file, you are asked to change the name.
- To remove a variable name from the list, just click once again at the name in the variables field at the left of the window.
- You can click the Select all button to add all variables from the data file at once. If one or more names already exist, you can change the variable name, skip the variable, or cancel the addition of variables.

Your next actions depend on which of the following situations applies.

Number and order of cases are the same

You are sure that the number and order of cases from the data file correspond exactly to the number and order of the sequences in your SEQUENCE VIEWER file.

- Select **Order of cases equals order of sequences**.
- Uncheck **Use check variables to check order**.
- Click Add.

Number of cases is the same, but order may be different

The number of cases from the data file equals the number of sequences in your SEQUENCE VIEWER file, but you're not sure if the order of the cases in the import file, is equal to the order of the sequences. This problem can be solved if both the import file and your SEQUENCE VIEWER file have one or more variables that should have the same values for corresponding cases and sequences.

Suppose for example that both files have the variables age and gender. If the values on these (check) variables are the same for successive cases, respectively sequences, you can be pretty sure that the order is the same.

- Select **Order of cases equals order of sequences**.
- Check **Use check variables to check order**.
- Click at the field below 'Check variables' to select it (if it is not selected yet). You may note that the variables field changes: variables already added to the 'Add variables' list are removed.
- Click at a variable name in the variables field. The name appears in the field below 'Check variables', followed by an equal sign. In the variables field the names of the sequence variables from your SEQUENCE VIEWER file are now displayed.
- Click at the sequence variable in the variables field whose values should equal the values of the variable from the data file. This name appears after the equal sign. You can add up to three such check variables.
- To remove a check variable, be sure the field with check variables is selected, and click at the name of the variable you want to remove in the variable list.
- Click Add.

The program first checks if the values of the check variables from the data file, equal the corresponding values in your SEQUENCE VIEWER file. If not, the program informs you about the mismatch.

Number and/or order of cases and sequences are different

The number or order of the cases from the data file does not correspond to number or the order of the sequences in your SEQUENCE VIEWER file. This problem can be solved if both the import file and your SEQUENCE VIEWER file have one or more variables that uniquely identify a case and the corresponding sequence. Suppose for example that sequences concern question-answer sequences, than each sequence is uniquely identified by the number of the respondent and the number of the question. If both the import file and your SV file have both (identification) variables available , you solve the problem as follows.

- Select **Use identification variables to identify sequence**.
- Click at the field below 'Identification variables' to select it (if it is not selected yet). You may note that the variables field changes: variables already added to the 'add variables' list are removed.
- Click at a variable name in the variables field. The name appears in the field below 'Identification variables', followed by an equal sign. In the variables field the names of the sequence variables from your SEQUENCE VIEWER file are displayed.
- Click at the sequence variable in the variables field whose values should equal the values of the variable from the data file. This name appears after the equal sign. You can add up to three such identification variables.
- To remove an identification variable, be sure the box with identification variables is selected, and click at the name of the variable you want to remove in the variable list.
- Click Add.

For each sequence, the program tries to find a case from the data file, which has equal values on all identification variables. If such a case is not found, the sequence variable(s) that is (are) added, obtain a missing value.

If the program finds two or more cases (lines) in the data file with the same values on the identification variables, you will get a warning dialog. If you click the Continue button, the import will continue without further warnings in case more lines have the same values on the identification variables, and the program assigns the data from the *latest* case in the file with identical values.

You may note that it is possible that a *number* of sequences all have the *same* corresponding case in the data file. A common example is that your SEQUENCE VIEWER file consists of question-answer sequences, whereas the cases in the data file are respondents (hence the

number of cases is much less than the number of sequences). Suppose you want to add the variable GENDER (the respondent's gender) from the import file to your SEQUENCE VIEWER file, and your identification variable is RESPNO (the respondent's number). This number will appear only *once* in your data file (because cases are respondents). If this respondent is posed 20 questions, you may have 20 sequences with the same RESPNO. For each of these 20 sequences, SEQUENCE VIEWER finds the same case in the import file. Hence, the value on the variable GENDER will be assigned to all these 20 sequences, which is most likely exactly what you want.

NOTES

The program assigns the values of the imported sequence variables also for locked sequences.
If the value of a sequence variable exceeds the limits set by SEQUENCE VIEWER (e.g. larger than 999999999999), it obtains the value 'M' for missing.

2.5.5 Event variables

You can add event variables to the presently open file from a tab delimited data file. This import facility is different from importing Event variables and codes:

- Event variables are added to an existing file (instead of creating a new file).
- Only event variables can be imported, not any accompanying codes.
- No facilities are available to convert dates or time stamps.
- Variable names should be included.

The first line of the data file should consist of (tab-delimited) variable names. Each next line should represent a different case (the event of a sequence) with tab-delimited values. The very first variable of each dataline should correspond to a sequence variable in the open file, in order to identify the sequence to which the values of the event variables should be added. The second variable should correspond to an event variable, in order to identify the event of that sequence. Hence, the combination of the first two values should uniquely identify the event in the SEQUENCE VIEWER file.

To add event variables do this:

- Select **File > Import...**.
- Select **Import event variables** and click OK.
- Open the data file with Get file. The variables field shows the names of the variables from the data file.
- You can check with Check file if there are any errors in the import file. For example, the combination of the first two values may occur in more than one line.
- Click at the field below 'sequence variable', and select the sequence variable that should identify the sequence.
- Click at the field below 'event variable', and select the event variable that should identify the event.
- Click at the field below 'Add variables' to select it. The variables field now shows the names of the event variables of the import file (the first line of the import file).
- Click at a variable name in the variables field. The name will be added to the list of event variables that are added. If this name already exists in your SEQUENCE VIEWER file, or if the name is too long, you are asked to change the name.
- To remove a variable name from the list, just click once again at the name in the variables field at the left of the window.
- You can add the variables in any order you wish.

- You can click the Add all button to add all variables from the data file at once. If one or more names already exist, you can change the variable name, skip the variable, or cancel the addition of variables.

If the combination of the first two values occurs more than once, the values from the last of these lines are added.

Values that are not a number (e.g. a period), or exceed the limits set by SEQUENCE VIEWER (e.g. larger than 999999999999), obtain 'M' for missing value.

The number of cases in the import file needs not be equal to the total number of events in your SEQUENCE VIEWER file. For example, if there are less imported cases than events, the events in the SEQUENCE VIEWER file that cannot be identified, obtain 'M' on the new event variables.

NOTES

The program assigns the values of the imported sequence variables also for locked sequences.

2.5.6 Sequence texts

This option is especially useful if you already have transcripts or other texts available. If you select this option, the IMPORT FILE setup window is closed and another window (IMPORT TEXT) is opened, with some import options.

First you should decide on whether you want the imported text to appear as only one sequence text, which is put in the text field of only one sequence.

Alternatively, you can divide the text into different parts and put each part into the text field of a different sequence. To this end, you should divide the text into parts using some delimiter. For example, you can put the character '&' between each part in your text file to serve as a delimiter for different sequence texts. You should prepare the text file to be imported beforehand.

Next you open the text file to be imported with the Get file button in the IMPORT TEXT setup window.

Type a string of maximally three characters in the edit box after 'sequence delimiter', to serve as delimiter between sequence texts; text separated by sequence delimiters are put in different sequences. If you type a carriage return as delimiter, a '¶' appears in the delimiter field. If you delimit the text by blank lines, you should enter two carriages returns as delimiter string ('¶¶'). If you leave this field empty, the imported text is not divided into parts and the whole text will be put into the text field of only one sequence. You can use a string of up to three characters that serves as a sequence delimiter, for example something like '[+]'. Be sure that the delimiter string only occurs in the text at places where you want to divide the text, and not as text you want to preserve as part of the sequence text.

The program places each part into the text field of a different sequence. Existing sequence texts are replaced.

You have three options from which sequence you want to start the import: the first sequence, the presently open sequence, or after the last sequence. If there are fewer sequences than text parts, new sequences will be added to the file (which is always the case if you select the third option of course).

You may note that as soon as you type a delimiter, the program calculates the number of text parts, and provides information about the number of sequences that are going to be added.

Click Import to import the new sequence texts. The program does not check if the number of lines of the sequence text equals the number of codes (if any) in existing sequences. After the

import you can select **Sequences > Link code and text...**, and then mark the sequences with bad links (with **Sequences > Mark links...**). Then you can inspect the marked sequences (if any).

NOTE

When you import text, you have the choice between plain text files, RTF (Rich text format) files, doc files, or docx files. If the imported file is different from plain text SEQUENCE VIEWER tries to convert such a file to plain text, except that SEQUENCE VIEWER tries to preserve text styles like underlines and symbols like the ↑. If the imported file has more complex formatting (like a table of contents), the imported text may contain 'tags' (like 'PAGEREF _Toc'). You better get rid of such formatting before importing a file, that is, save it first as plain text. However, in that case text styles are lost too.

2.6 Export

INTRODUCTION

You can export data for analysis by a statistical program (e.g. values of sequence variables), or to open them as a word processor file (e.g. the event texts). The available options are different for both possibilities.

The field below 'preview' shows a preview of the file as it will be exported. Inspect the preview file to check if the file to be exported is what you expect. To show returns and tabs, check the ☑ ¶ checkbox (in the upper right corner). In case of large files, or tedious transformations, not all sequences are displayed in the preview field.

- Select **File > Export... t**o open the FILE EXPORT setup window.
- Decide whether you want to export data for analysis by another program, or export data to read them in a word processor program.
- Select the appropriate radio button below 'export for' (in the upper left part of the setup window).

EXPORT FOR STATISTICAL ANALYSIS

Data for analysis are exported as a plain ASCII file, with tabs between values and each case on a new line. Values can be numerical (e.g. sequence variables) or alphanumerical (e.g. event codes). Not all combinations of different data can be exported for statistical analysis, because the unit of analysis of two different data types may be incompatible. For example, an event text may contain two or more text keys, whereas a text key may concern two or more event texts.

The first line of the export file may contain variable names, depending on whether or not you check the ☑ **Include variable names** checkbox, in the export options section. In the preview field the variable names remain visible if you scroll vertically. You can abbreviate variable names to maximally 8 characters. In that case:

- spaces are deleted
- leading numerical characters are deleted
- characters after the eighth character are deleted.

Please note that this may lead to identical variable names. You can also choose to keep the original names, as appear in SEQUENCE VIEWER.

You can choose to add an end-of-line character after the last line of the file, because this is required by some statistical programs. A number of other options are available, depending on the type of data.

Export sequence variables

If you export only sequence variables, or only sequence variables in combination with one or more information fields, the cases are sequences.

- Check ☑ **sequence variables**
- Select 'sequence variables' with the option menu in the select variables section.
- Click at a variable name in the variables list to select it for export. The name becomes highlighted and appears in the selected export variables, below 'sequence variables'.
- Click at another variable name to add it to the export list. The order of the exported variables will correspond to the order in the 'export variables' field.
- To remove a variable from the export list, click its name in the variables list.
- Check ☑ **Include variable names** if you want the names of the exported variables to appear on the first line. If you want abbreviated variable names, check ☑ **Abbreviate names**.
- Check ☑ **Add EOL to end of file** if you want to add an EOL character to the end of file.
- Enter a numerical value or a period for the missing values (designated with 'M' in SEQUENCE VIEWER).
- Click Export. You are asked for a file name. Enter a file name and click Save.

Export event variables

If you export event variables, the cases are always the events. If you export both sequence and event variables, the values of each sequence variable are repeated for each event of a sequence.

You can also check 'Checked events only'  , if you wish to export the event variables of checked events only.

To export event variables, you follow the same steps as with exporting sequence variables, except that you now select 'Event variables' with the option menu in the select variables section.

Export event codes

To export event codes, check ☑ **event codes**. Event codes can be exported in a number of different ways (check the preview field for the effect of the different options).

Events or sequences as cases

You can export event codes either with events or sequences as cases. Check or uncheck ☑ **Use sequences as cases**. If event codes are cases, each event code is written on a separate line. If you include variable names, the name of the variable will be 'codes'.

If you check ☑ **Use sequences as cases**, all event codes of each sequence are written on one line, delimited by a tab. Because the number of event codes is usually different for different sequences, either empty events (just dashes), or missing values should be added to each sequence, to make all sequences of equal length. If you include variable names, the name of the variables will be 'code1', 'code2, 'code3', etcetera.

If you want to export event codes together with event variables and/or event texts, only events can be cases.

Delimit codes

If event codes are exported with events as cases, you have the option to delimit the codes. This means that the codes of the event codes are separated by a tab; thus, each column simply represents the codes on a code variable. If you check ☑ **Include variable names**, variable names are the names of the code variables.

Export event texts

If you export event texts, the cases are always events. Each event text constitutes an alphanumerical value on a variable that is called 'event text' (or 'evttext', if abbreviated).

Export text keys or time keys

If you export a text or a time key, each instance of a particular keyword in the sequence, is a separate case. Exported are the start and end of each key. Variable names are 'keyStart' and 'keyEnd'. You can only export one text key or one time key. You should select the name of the key with 'Select text key' or 'Select time key' in the key options section.

If text keys are selected, you have the option to export the text of the text key too. Because the text of a text key can contain carriage returns (because two or more event txts may be involved in a text key, these returns should be replaced by another character. In the recode options section you are offered the choice between '/', '|', '\' and a space.

Only sequence variables and/or information fields can be exported together with text or time keys; the values of these variables are repeated for each key in a sequence.

Export word links

If you select word links, the positions of the start word and end word of the key are exported as two separate variables. In addition you can export both words themselves.

Export information fields

Like the text of text keys, information fields may contain carriage returns. You should replace them by another character (in the recode options section).

EXPORT FOR WORD PROCESSOR

The data for a word processor are exported as a RTF file. Sequence and event variables are selected in the same way as with export for statistical analysis. To export text or time keys and word links, you should also export the event texts. There are no options available for how the different pieces of information are exported; see the preview field to observe how the eventual file will look like.

2.7 Compare

INTRODUCTION

The 'Compare files' command is most useful if you have a number of copies of the same file, and different persons independently make adjustments to each file. For example, one user of the file detects some errors in the text of the text field and corrects these errors. Of course it is useful to make the same corrections in the other files.

The 'Compare files' command detects differences in two files, thus allowing you to adjust both files and to make them similar, if necessary.

The program does not compare everything in the file. Comparisons are only made with respect to the event codes, the event texts, the values on sequence variables and the values on event variables. You will be told however if both files are exactly identical in all respects.

SETUP

- Select **File > Compare files...** to open the COMPARE FILES setup window.
- Click the Get file button.
- Open the file you wish to compare with the presently open file.

After you opened the file you wish to compare, the program first checks how compatible both files are, that is, file properties are checked first. For example, if one user added or deleted a code variable, comparing this type of data is not straightforward. The user is informed about differences with respect to the number of code, sequence and event variables, the number of sequences and events, and properties of event times.

Use the option menu at the top of the window to select the type of data (event texts, event codes, etc.) you wish to compare between both files.
If you select 'Codes', 'Sequence variables' or 'Event variables', you can in addition select a variable name. If you select a variable name, the properties of that variable are compared, for example the variable type.
Click the Compare button.

Now the values of event variables or sequence variables, or the (event) codes in case you selected 'Codes', are compared. If you selected 'Text', the sequence text of the sequences are compared.
If the number of code variables is different in both files, for example 4 and 6, the program only compares the first four code variables; it is assumed that these four code variables concern the same variables.
If the number and/or names of sequence variable are different, you can select the name of a sequence variable that occurs in both files, to compare the values of the selected sequence variable.

If a difference is found, the program stops and the differences are flagged. The sequence were the difference is found, is shown in the main sequence window, so you can adjust the data, if you wish. Information about the difference is found in the information field. To continue with the comparison, click the Compare once again. You may change the sequence from which to start the comparison. For example, if you adjusted the data, you may prefer to select **Start at present sequence**.

If you compare (event) codes, and the number of events in both sequences is different, the program tries to figure out which event codes are different. Also in case of comparing text, the program tries to detect were exactly the text differs.

It may happen that one user added or deleted a sequence from the file. In that case you can check one of both **Ignore sequence** checkboxes. This sequence will be ignored, and hence the sequences will remain synchronized in the compare process. If you select another data type, the 'ignored' sequences remain ignored.

You can use the blue arrow buttons to go one sequence backwards or forwards. If for example, the presently open file shows sequence 14, whereas the other file shows sequence 16, after clicking the Compare button, sequence 15 will be compared with sequence 17, sequence 16 with sequence 18, and so on. You can quickly set the sequence number of the one file to the sequence number of the other file, by using one of both Go sequence buttons.

If **Synchronize scroll** is checked, scrolling in one field, scrolls the other field with an equal amount.

2.8 Settings

'File settings' concern options that are specific for a particular file. File settings are stored within the file, as soon as you save the file. Open the FILE SETTINGS window with **File > File settings...**.

2.8.1 File info

The **File info** tab gives you some basic information about your file (like name, location, size or modification date), but you can also enter information about the file yourself. Just type any text in the scrolling field. You can use hard returns in this field (the OK button will not be activated by pressing return). The information is stored within your file, and saved as soon as you save your file.

It is assumed that your data file usually has codes, because the main analyses of SEQUENCE VIEWER concern codes. However, it is very well possible that all your sequences are empty (that is, have no codes), and that you only have sequence variables for each sequence. In such a case you may use SEQUENCE VIEWER as a simple statistics program.

In addition to codes, you may have:

Sequence texts. If sequence texts are separated by carriage returns, each paragraph is called an event text.
Audio files or video files (AV files). Audio and video files are always linked to a sequence.

Depending on the types of data in your file, use the following settings:

- If you do have sequence texts, or want to add sequence texts later on, check ☑ **Sequence text**, else uncheck ☐ **Sequence text**.
- If you have linked audio or video files, or want to add such files, check ☑ **Linked AV files**, else uncheck ☐ **Linked AV files**.

The effect of these different settings is that the main sequence window is adjusted according to these settings, whereas some menu options become available. You can always change these settings later on.

2.8.2 Time

VARIABLE TYPE

SEQUENCE VIEWER discerns between three different variable types, 'numbers', 'dates' and 'time stamps'. For a discussion of these types, see Numerical data.

The settings in the **Time** tab of the FILE SETTINGS, concern the variable type of three variables, the sequence variable 'STARTTIME', and the event variables 'Onset time' and 'Offset time'. All these three variables should always be of the same type and time unit.

TIME UNIT

In case of audio, a good time unit is 6 ticks, or one tenth of a second. If you want to observe video fragments, select the corresponding frames per second as time unit (2 ticks if the frame rate is 30, 2.4 if the frame rate is 25). Onset time and offset times, as well as the start time of a sequence, are displayed in the chosen time unit. If you change the time unit, all onset times, offset times and start times (unless they have missing values of course), and the start and end of time keys, are automatically recalculated at the moment you click OK. For example, if you change the time unit from 10 ticks to 6 ticks, an onset time of 20 time units will become (10/6) * 20 = 33; the values will be rounded to whole numbers.

If you have time keys, these are adjusted too. You should be aware that if you change your time unit upwards (for example from 1/10 of a second to 1/6 of a second), you may loose information (precision).

In measuring the time, numbers are rounded to the nearest integer. That means that if the time unit is for example one second, after 0.5 seconds, the movie will show 'time: 1' (1 second); after 1.5 seconds it will change to 'time: 2'; and so on.

Time units can very well be of a completely different magnitude. For example, if you use SEQUENCE VIEWER to study life events (and don't use AV files), an appropriate time unit may be years or months.

ABSOLUTE AND RELATIVE TIME

If the linked AV file concerns a number of sequences (for example, the AV file is a complete interview, and the sequences are the utterances belonging to each question that is posed in the interview), the start time of each sequence will be different; for example, the start of the second sequence (when the second question is posed to the respondent) may be 20 seconds after the start of the interview. Onset and onset times of the events (utterances) of a sequence may either refer to the very start of the AV file (Event time is *absolute*), or to the start time of each sequence (Event time is *relative* to STARTTIME).

For example, suppose that the STARTTIME of a particular sequence is 150, and the absolute onset time of the second event is 170. The relative onset time of this event will be 20, that is, 20 time units later than the STARTTIME of the sequence.

In general, the onset time of the first event of a sequence should always be the same as the STARTTIME of the sequence. Thus, the *relative* onset time should be 0, whereas the value of the *absolute* onset time should equal the value of the STARTTIME.

If you change 'absolute' to 'relative' or vice versa, onset and offset times, as well as STARTTIME are automatically adjusted at the moment you click OK.

Time stamp and date variables are always absolute, and you cannot change them to 'relative'. The reason is that the numerical values of 'relative' dates no longer represent the number of days or seconds since January 1, 1970. For some types of analyses however, 'relative' dates or time stamps are more useful than 'absolute' dates or time stamps; for example if you want to compare the average duration of a sequence for different groups. It is easy to create pseudo relative onset and offset times however. For example, in the COMPUTE EVENT VARIABLE setup window, enter respectively:

RelOnset = Onset time - Onset time[1]

and

RelOffset = Offset time - Onset time[1]

which has the effect that the onset time of the first event is subtracted from each onset an offset time. 'Onset time[1]' is the onset time of the first event in a sequence; see Compute event variables for more information.

The ☑ **Keep auto time checked if new sequence is opened** checkbox is useful in case you are

assigning onset or offset times to events. For more information, see Enter event times.

In the case of life events, each sequence describes the life course of an individual. Quite likely the start time of each sequence is always zero (the birth of the individual). In that case relative event times make no sense, and you should set the event to 'absolute'.

WARNING

If your onset and offset times are relative to STARTTIME, changing the value of the STARTTIME, makes all existing onset and offset times invalid. If you need to change to STARTTIME nevertheless and prefer to use relative event times, first change relative event times to absolute event times, than edit the values of STARTTIME, and finally change the absolute event times back to relative event times.

2.8.3 Playing

The options on the **Playing** tab of the FILE SETTINGS window only have effect if the player window is open.

If ☑ **Stop sound at end of event** is checked, the sound stops playing at the end of the event text (provided the offset time is available for the event).

If ☑ **Play last ten units of event** is checked, the sound plays the last ten time units of the event (provided the offset time is available for the event). This option is useful to fine tune offset times (see Enter event times).

If the sound (or movie) is playing, you can stop it, either by clicking anywhere in the sequence text field (select ◉ **Text field**), or at the very event text that is playing (select ◉ **Event text**).

If ☑ **Open linked file if new sequence is opened** is checked, the linked AV will be automatically opened if you go to another sequence.

If ☑ **Show next sequence if movie exceeds next STARTTIME** is checked, the next sequence is shown in the main sequence window, as soon as the counter of the movie exceeds the STARTTIME of the next sequence. The checkbox has no effect if:
the linked file to the next sequence is different from the present one,
the position of the counter when playing is started does not refer to the present sequence, or
the TIME KEYS or TEXT KEYS setup window, or the WAVE FORM setup window is open.

See also:
Audio and video: Linking.

2.8.4 Coding

The coding options concern additional options of the EDIT SEQUENCE window and only make sense if this window is open.

If you go to a different sequence, depending on which one of the radio buttons ◉ **Select first**

event code and **Select first uncoded code** is selected, either the very first event code of a sequence is selected, or the first event code that is not (fully) coded yet (that is, an event code with at least one '-').

The option **Use first character of event text as first code** is especially useful if you code conversations, were each utterance is preceded by a one character code to identify the speaker. For example "I:" or "R:" for interviewer and respondent respectively in an interview (you can use any character, of course).

Assume, the first code variable designates the speaker, with the codes 'I' or 'R'. Because the first character of each utterance now corresponds to the code of the first code variable, it is possible to automatically 'code' the first code variable. To this end check **Use first character of event text as first code**.

If you start coding using the EDIT SEQUENCE window, and click Auto fill in this window, each utterance obtains and 'empty' event code, consisting of dashes, but the first dash is replaced by the first character of the utterance, e.g. 'R----'. Be sure to uncheck this checkbox if you don't want to assign the first character of the event text to the first code variable! See also Edit sequences in text mode.

The minimum agreement between text strings is used for suggestions (see Edit sequences in text mode) of codes and finding event texts that resembles a specified text (see Auto code).

For more information about how to code sequences and edit codes, see Edit Sequence.

2.8.5 Customize

You can change the names of your mark sets and the info fields of the main sequence window to more meaningful names. Select **File > File settings...**, and click the **Customize** tab. Select the name of an existing mark set or info field, type a new name after 'New name' and click Rename.

3 Edit

The **Edit** menu contains the common edit operations like undo, cut, copy and paste.
In addition to these options, you find some more options in the **Edit** menu, allowing you to edit your SEQUENCE VIEWER file. These are discussed in the next sections

3.1 Undo

The **Edit** menu has two different 'Undo' commands. The first one concerns minor changes to your file, for example editing the sequence text, or adding a word link. Such changes can usually be undone. The text of the Undo command changes in order to inform you what can be undone, e.g. 'Undo typing'.

The second Undo command concerns major changes made to you file (see files). Major changes are generic changes made by executing a particular command, for example deleting a sequence variable
A major change can be undone, using **Edit > Undo [type of change]**. The text of the Undo command informs you about the action that will be undone (for example 'Undo delete sequence variable'). The data file will be set to the situation immediately before this action (e.g., before the sequence variable was deleted). All (minor) changes you made to your data file after this action will be lost. You can undo up to nine successive major changes. In case of large data files, a high number of undo's, may slow down performance a bit.

3.2 Paste

If you paste text in a particular field, the pasted text may be a bit different from the text you copied, because SEQUENCE VIEWER sometimes adjusts the text according to the requirements belonging to the field were you paste the text.
If the text to be pasted contains diacritical marks, whereas the field where you want to paste the text does not allow for such marks, these marks are removed; for example, "señor" becomes "senor", or "Mr. Müller" becomes "Mr. Muller".
If the text to be pasted contains a carriage return, whereas this is not allowed, the carriage return is usually replaced by a space. Two ore more spaces in succession are usually replaced by one space.
If you paste a number preceded by "+" in a field expecting a number, the "+" is usually removed.
Generally, characters that are not allowed for a particular field are removed when pasting. For example, if you copy "(my NewVar)" and paste it into a variable name field, the parentheses are removed (parentheses are not allowed for a variable name, because they may interfere with parentheses in an equation).
Because variable names cannot have more than 12 characters, only the first 12 characters of a string are pasted in a variable name field - after removing characters that are not allowed for a variable name (like the parentheses, or a carriage return).
The same is true for values: because values cannot have more than 12 characters (digits, minus

sign, decimal point), only the first 12 characters are pasted.

If such small corrections are applied in pasting text, the original clipboard remains unaffected. This means for example, that if you copy "(my NewVar)" and paste it successively into a field for variable names, and into the sequence text field, in the variable name field "my NewVar" is pasted, whereas the next paste in the sequence text field gives you "(my NewVar)".

In other cases, strings that don't fulfill the requirements of the format are not pasted at all. For example, suppose you copy "123a" and paste it into a field where a number is expected, then nothing will be pasted. Also pasting a negative number where a positive number is expected is (usually) not possible. Neither is it possible to paste a value for a variable, if the corresponding (legal) variable name is not entered yet.

Finally, in some fields you cannot paste at all; for example into a one-character field, like the field for the value (one-character code) of a code variable.

3.3 Clear sequences

You can clear particular fields of a sequence: the text, the codes or each of the five information fields. In addition you can clear the marks of all eight mark sets, that is, unmark all sequences according to each mark set. And you can set all sequence variables to 'M' (except sequence variable 'SEQSIZE', because the value of this sequence variable depends on the number of event codes).

Please note that if you clear the codes, also the values of the event variables will be deleted.

- Select **Edit > Clear sequences...**.
- Check the fields you wish to clear.
- Click Clear.

See also:
Select marked sequences.

3.4 Paste sequences

INTRODUCTION

You can copy sequence information (the codes, the event variables, the sequence text and the five information fields), from a particular sequence. The copied information can then be pasted into the corresponding fields of other sequences, i.e. all other sequences, the marked sequences or the selected sequences.

You can also copy all information of a sequence and paste it into other sequences. These sequences will be exact copies (including information about marks, linked files, sequence variables, etcetera) of the original sequence.

SETUP

- Go to the sequence you wish to copy.
- Select **Edit > Paste sequence...**.

- To copy the whole sequence, select **Copy whole sequence**. This includes all information, like values of sequence variables, linked sound files, etcetera.
- To copy only selected parts, select **Copy part of sequence**. Check one or more checkboxes.
- The event variables (if available) can only be pasted, if the event codes are pasted too, and vice versa. See 'Additional information' below.
- If **destination field is empty** is checked, the information will only be pasted if the destination field is empty. For example, if **Copy event codes and event variables** is checked, event codes are only pasted into the code field of another sequence, if this sequence has no event codes yet.
- Check **sequence variable** to paste data only to the sequences with a particular value on a sequence variable. To select this variable, click at the name of a sequence variable in the main sequence window. The value will be the value from the present sequence on this variable, but you can change it.
- Before you click Paste, you can go to a different sequence. You may note that the sequence number is reflected in the field after 'Copy information from sequence:'.
- Be sure to specify to which sequences the present sequence should be pasted (to prevent that the present sequence is pasted to all other sequences). You can copy to the marked sequences (check , see Select marked sequences), or to a selection of sequences (with , see Select sequences) or use the sequence variable checkbox, as described above.
- Click Paste. It will paste the information from the presently open sequence.

ADDITIONAL INFORMATION

You can neglect this explanation if you don't have event variables.

If you copy event codes, you have to decide whether or not to paste and copy the values of the event variables belonging to each event code. This may cause a problem if the number of event codes of the copied sequence is different from the number of event codes of the sequence you paste to (the 'pasted' sequence). Hence this explanation.

You have three options.
- First, you can copy and paste the values of the event variables.
- Secondly, you can assign a missing value to all event variables in the 'pasted' sequence.
- Thirdly, you can keep the original values of the pasted sequence.

For example, suppose you copy the event codes from sequence 1 with three event codes to sequence 2 with five event codes. Sequence 2 will have values on the event variables for each of these five events. If you select the first option, there will be no problem: both event codes and values on the event variables of sequence 2 will be replaced by the three event codes and their corresponding values on the event variables from sequence 1.

If you select the second option (Replace values with 'M'), the values of the event variables for each of the (now) three events of sequence 2 all obtain a missing value.

If 'Keep old values' is selected, the old values are kept. However, sequence 2 has five events, whereas sequence 2 has only two events. In this case the values for event 1 to 3 are kept, whereas the values for events 4 to 5 of sequence 2 are deleted. Now suppose sequence 2 has only two event codes. In this case, the values for event 1 to 2 are kept, whereas the values on the event variables for event number three obtain missing values.

3.5 Edit sequences

To add, delete and change codes (values of code variables), you should use the EDIT SEQUENCE window (you cannot edit code variables in the main sequence window itself). Select **Edit > Edit Sequence...** to open the EDIT SEQUENCE window. The available options depend on the mode of the sequence window. Most options are available in the 'codes and text mode'. You should use this mode (click at  in the *Layout* section) if you are going to assign codes to texts, like transcriptions of conversations. If you don't have sequence texts, use the codes only mode () if you want to edit codes.

Event variables can both be edited using the EDIT SEQUENCE window, and in the *Codes and event variables mode* of the main sequence window (click  in the *Layout* section).

Sequence variables can be edited in the Sequence variables drawer.

3.5.1 Text mode

INTRODUCTION

Texts can be anything, e.g. articles of newspapers or conversations. We will use interview transcripts to illustrate the coding procedures.

You can view a whole interview as a sequence of successive utterances. Different interviews constitute different sequences in that approach. This approach is most appropriate for open interviews.

For standardized interviews, when the same questions are posed to a number of respondents in a prescribed order, it is more appropriate to view each interview as a number of question-answer sequences, each question-answer sequence constituting a different sequence. Each question-answer sequence starts with a particular question, and the events are the utterances of interviewer and respondent belonging to that question. We use the latter approach for our illustration.

BASIC EDITING

Start coding (coding field is still empty)

Click the Auto fill button at the bottom of the EDIT SEQUENCE window. In the code field of the main sequence window, each event text is now 'coded' with an 'empty' event code, consisting of dashes. You may note the effect of checking  **Use first character of event text as first code** (see 'Coding' in the FILE SETTINGS).

Changing codes

Click at the event code in the code field of the main sequence window you want to change. This event code now appears in the code field of the EDIT SEQUENCE window. You can now type the new codes, or you can click at a code in the field showing the allowed codes. If you have defined code dependencies, some codes may appear in red. These are codes that belong to the code variable, but are not allowed because of particular earlier codes in the event code.

After you have entered the codes of an event, click at the Replace button in the EDIT SEQUENCE window. The next event code becomes selected.

If ☑ **Replace if last code of event is typed** is checked in the EDIT SEQUENCE window, the codes in the code field of the main sequence window are automatically replaced, if the last code of an event code is entered.

Inserting event codes

To insert an 'empty' event code, click at an event code in the code field of the main sequence window and click the Insert button in the EDIT SEQUENCE window. The empty event is inserted before the selected event.

To add an empty event after the last event, click the Add button in the EDIT SEQUENCE window.

If you hold the option key down when clicking the Insert or Add button, the codes of the selected event (and the value of the event variable, if applicable) are copied to the inserted or added event.

Deleting event codes

To delete a coded event, click at an event code in the code field of the sequence window and click the Delete button in the EDIT SEQUENCE window.

LINKED CODE/TEXT OPTIONS

The Add, Insert and Delete buttons also affect the associated event variables. For example, if you delete the event code of the second event, the values of the event *variables* belonging to the second event will be deleted too. This is not the case for the event *texts* however; these will *not* be deleted. If codes and text are linked (see Linked codes and text), the number of event *codes* and the number of event *texts* will become different after the deletion or insertion, and code and text are no longer linked. If you want to delete the event text that belonged to the deleted event code, you should delete it manually in the main sequence window. Alternatively, you can use the Linked code/text options, described below.

Delete link

If codes and text are linked, and you want to delete both the event code and the corresponding event text, you can use the Delete link button.

Combine links

If you want to delete a code from a sequence with linked codes and text, you may want to add the event text to the preceding or subsequent event. You can use the 'Combine links' pull down menu to this end. With this pull down menu you can join two events, both the event codes and the event texts. If you combine the event code with the next event code, the event text of the next event will be added to the event text of the selected event. The event code after the selected event will be deleted, including its event variables. However, if you have onset and offset times, the offset time of the selected event will be replaced by the offset time of the deleted event. If you combine the selected event with the previous event, the onset time of the selected event will be replaced by the onset time of the deleted event.

Insert link

Similarly, if you insert an event using the Insert button, codes and text will no longer be linked. Instead, you can use the 'Insert link' pull down menu. This adds an event text (you can enter this text in a separate dialog box) to the inserted event, thus preserving the link between codes and text.

Split link

To split the event text into two parts when codes and text are linked, you can use the Split link button. First click this button (the cursor will change into scissors), then click at a word in the event text (if you decide at that moment to not split the event, just click somewhere outside the sequence text field). The event text is split into two new event texts; the word that is clicked will be the first word of the second event.

You can assign the event code to either the first, second, or both parts of the event after the split, with the pop-up menu below 'assign event to:'. If you select 'first' or 'second', the other event will be coded with dashes ('-' for 'not coded yet'). However, if in the **Coding** tab of the FILE SETTINGS, **use first character of event text as first code** is checked, this event will obtain the same first code as the other event, whereas it will be preceded by this code plus a colon.

The event variables of the old event are copied to the new event (if you have onset and offset times, the offset time of the first event and the onset time of the second event will be set to 'M').

UNDO AND RESTORE

You can undo the last change you made with the Undo button. The Restore button restores the original sequence, so all changes are undone. Restore itself is also a change and can be undone with the Undo button.

See also:
Suggestions.

3.5.2 Codes only mode

In the *Codes only* mode ( ) only codes are shown. If an event variable is selected with the 'Show event variable' pop-up menu in the *Layout* section, the value of this event variable, separated by a slash, appears after each code.

To enter new event codes, in the 'codes only' mode, open the EDIT SEQUENCE window (**Edit > Edit sequence...**). Changing codes, adding event codes or deleting event codes, is essentially the same as described in the section Basic editing of editing codes in text mode. The 'Linked code/text' and 'suggestions' options are not available.

3.5.3 Codes and event variables

In the *Codes and event variables* mode ( ) the codes and all event variables are shown. Changing codes, adding event codes, or deleting event codes with the EDIT SEQUENCE window (**Edit > Edit sequence...**), is essentially the same as described in editing in the Codes only mode and the section Basic editing of editing codes in text mode.

However, using the EDIT SEQUENCE window, you can only edit the very first event variable. To edit the other event variables too, you should close the EDIT SEQUENCE window. Now you can click on a value of any event variable in the main sequence window, and change the value.

If one or more of your event variables are of type Date or Time stamp, then you can also edit the value in human readable format. In the lower right part of the main sequence window you find an edit box, where the numerical value of the selected date or time stamp variable is shown. You can edit the date or time stamp in this edit box, and click the Accept button.

3.5.4 Suggestions

INTRODUCTION

Especially in standardized interviews, it may occur that the same or nearly the same event texts repeatedly occur. The same questions are posed to a large number of respondents. All these questions will usually obtain the same event code (provided the interviewer worded the question as scripted). Also many answers of respondents will be more or less the same and may be coded with the same event code.

To ease the task of coding, you can associate particular event texts with a particular event code. Each time an event text is coded that is similar to such a text with associated code, the most likely code is that associated code.

CREATE SUGGESTIONS

The easiest way to create such associations is to use the Learn button of EDIT SEQUENCE window:

- Select **Edit > Edit Sequence...** to open the EDIT SEQUENCE window.
- Be sure the main sequence window is in 'codes and text' ()mode.
- Text and codes should be linked.
- Click at an event code in the code field of the main sequence window.
- Click at the Learn button in the EDIT SEQUENCE window.

The event text from the main sequence window that is coded in the EDIT SEQUENCE window, is now associated with that event code as a suggestion. This event code will be suggested by SEQUENCE VIEWER as the most likely event code for that event text.

Please note that the associated event code, is the event code from the EDIT SEQUENCE window, not the event code in the main sequence window. If a particular event text already has an associated event code, the Learn button is disabled.

HOW TO USE SUGGESTIONS

To use such suggestions:

- Click at an event code in the code field of the sequence window that is not (or not fully) coded yet.
- Click at the Suggest button in the EDIT SEQUENCE window.
- Click Replace to accept the suggested event code.

SEQUENCE VIEWER compares the to be coded event text with all texts that have associated event codes. The event code associated with a text that resembles the to be coded event text most, will be suggested.

In the lower part of the EDIT SEQUENCE window, the actual agreement between both texts, as a number between 0 and 1 is shown.

The to be coded event text may already be partially coded. For example, if you code interviews, you may already have assigned the code for interviewer and respondent, whereas the remaining codes consist of dashes (for "not coded yet"). In finding the event code that resembles the associated event text best, existing codes are respected. For example, if an event text is already coded as "I----", the program only looks to event texts in the database, with an associated code that has an "I" at the first position of the event code.

EDITING SUGGESTIONS

Alternatively, you can add, delete or edit suggestions manually. To this end, click at the Edit suggestions button in the EDIT SEQUENCE window (the main sequence window should be in 'codes and text' mode). You obtain the SUGGEST EVENT CODES setup window.

- In the field below Event t*ext*, type or edit the text that is associated with a particular event code.
- In the field below *Suggested event codes* you can type or edit the event code that should be associated with the event text.
- You can enter additional information for a particular association in the field below *Additional info:*. This may be useful if there are less obvious reasons to use a particular event code for the event text.
- Check ☑ **Ignore suggestion** to temporarily ignore a particular suggestion (alternatively you can of course delete the suggestion, using the Delete button, but then it will be lost).
- After each change regarding a suggestion, you should click the Store button.
- The suggestions can be sorted, according to the code, text, additional information, or whether or not they should be ignored.
- To enter a new suggestion, click New ; to undo the changes with respect to a particular suggestion, click Revert .

In addition there are a number of general options to improve the likelihood that a correct event code is suggested.

Suggested event codes are based on how much the event text to be coded resembles one of the event texts in the database of suggestions. The event code belonging to the event text from the database that resembles the to be coded event text most, is suggested, provided that the resemblance exceeds some lower limit. You can adjust this lower limit with the 'Suggest if resemblance ≥' popup menu. If you obtain many wrong suggestions, set this limit to a higher value. If (nearly) all suggestions are correct, but in many cases no suggestion is given at all, the limit may be too high.

It may be useful neglect particular words in determining the agreement between the to be coded event text and the event texts from the database; for example words like "uhm". You can type such words, separated by a space, in the field below *Neglect words*.

Finally there are a number of common options that are similar to options used for finding text; see Text adjustments for an explanation of these options.

If you save your data file, all changes to the suggestions are saved too.

3.6 Assign empty codes

You can fill the code field of empty sequences (that is, without codes) with empty event codes. An empty event code only consists of dashes, for 'not coded yet', for example '-----' in case you have five code variables. The number of coded events equals the number of paragraphs in the text field; if the text field is empty, no empty events will be added. You can select to fill all empty sequences or to fill only the marked empty sequences. The texts and the empty event codes are automatically linked.

• Select **Edit > Assign empty codes...** to open the FILL EMPTY SEQUENCES setup window.

• Select **Fill all empty sequences** or **Fill marked empty sequences**.
• Click OK.

3.7 Waveform

INTRODUCTION

A waveform is a graphical representation of a sound. The most important characteristics of a sound are loudness (which is related to the amplitude or magnitude of the sound wave) and pitch (which is related to the frequency or oscillation of the sound wave).

A waveform can aid you in assigning onset and offset times to events. In addition, you can assign basic characteristics like pitch and (relative) loudness of voices to event and sequence variables. This makes it possible to investigate relations between these variables and other features of events and sequences. As a simple example, it is believed that pitch is related to the attractiveness of a (female) voice. Attractiveness of the interviewer's voice in turn, may be related to the success in persuading sample persons to interviewed.

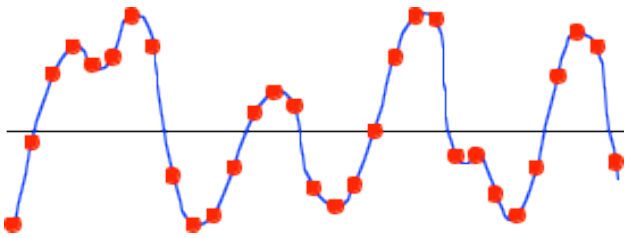
The analysis of waveforms in SEQUENCE VIEWER cannot replace results from sophisticated programs like the much used program 'Praat' (www.praat.org). SEQUENCE VIEWER however gives you the possibility to directly assign simple voice characteristics to event variables.

Below we will give you a short (and simplified) overview of sound waveforms, in order to better understand the available information and options in SEQUENCE VIEWER. If you plan to work with waveforms, and especially if you plan to do pitch analyses and assign sound characteristics to event variables, we urge you to do the exercises in chapter 11 of the Tutorial. Before doing these exercises you should read this introduction first.

Sample Rate

Vinyl records represent the waveform directly, but sound files are digitized, which means that at particular time intervals a sample of the (analogous) sound is taken. A common time interval is 1/44100 second (or a sample rate of 44100 per second), which is used for audio CD's.

The picture below clearly shows that the waveform can be more or less reconstructed from the samples (or sample frames: the red dots); and that the higher the sample rate, the better the reconstruction. Especially fast oscillations (higher pitch) are inadequately represented, as the picture below makes clear. The rule is that the sample rate should be twice the sound frequency to be represented. Because the maximum frequency that can be heard by human beings is about 20000 Hz, this (partly) explains the choice for a sampling rate of 44100 for audio CD's.



Sample format

In a sound file the different sample frames are simply stored as a sequence of numbers, representing the deviations from the horizontal axis. These numbers are usually stored as 16 bit integers (the sample format). This format is used for audio CD's. Other sample formats are 8 bit integer, 24 bit integer, or 32 bit floating point. The higher the number of bits, the more precise the sample points can be located at the vertical axis of the waveform. This affects the dynamic

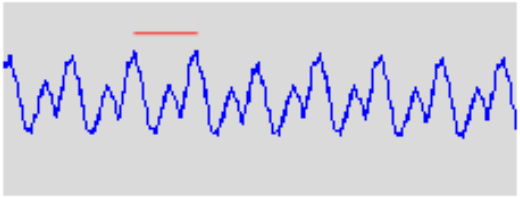
range of the digitized sound, which is the difference between the loudest and softest parts of the sound.

Channels

Stereo sound simply consists of two such sequences of numbers, or channels. More than two channels are common in recording audio signals, and are also used for sound reproduction by some surround sound systems.

Frequency and pitch

The waveform of a human voice is pretty complex. If you look at the picture below (taken from the waveform of a female voice), you may wonder what the frequency of this sound is.



If you take a closer look at the picture you can see that there is a clear repetition, with a length equal to the red bar. The number of such intervals per second is the fundamental frequency (or F_0). The fundamental frequency determines the perception of pitch. The picture also tells you that this is not the whole story; there are also a kind of subcycles (or formants), which in addition affect the perception of voices. The fundamental frequency is usually used as an indication for the perception of pitch, and that is the measure we also use in SEQUENCE VIEWER.

WAVEFORMS IN SEQUENCE VIEWER

File size and amount of data

The size of a sound file depends on the duration of the sound, the sample rate, the sample format, and the number of channels. For example, one second of sound on an audio CD consists of 44100 * 2 (16 bits or 2 bytes) * 2 (channels) = 176400 bytes. This is quite a large amount of data to convert to a waveform. There are a number of ways to reduce the amount of data to be processed in SEQUENCE VIEWER.

First, sound files in SEQUENCE VIEWER usually consist of speech only. The frequency of human speech ranges from about 60 to 7000 Hz; the range of telephone speech is even lower (about 300 to 3400 Hz). For this reason, a sample rate of about 8000 Hz is sufficient to obtain perfectly understandable speech, and reduces the amount of data to be processed considerably.

A further reduction in data with only a slight loss in quality (hardly noticeable), can be obtained by using a sample format of 8 bits instead of 16 bits.

Finally, instead of two channels (stereo), one channel (mono) is usually sufficient.

File format

Although SEQUENCE VIEWER can read and play many audio file formats, only uncompressed AIFF files (Audio Interchange File Format) are suitable for creating waveforms. Unfortunately this also means that you cannot directly create a waveform of the sound of a video file. However, you can convert any sound file that can be read by SEQUENCE VIEWER, or the sound track of a video file, to an AIFF file, and at the same time reduce the size of the file according to the options described in the previous section. Any professional sound program (e.g. Adobe Audition) can do the job. Alternatively, you can also use the excellent freeware program Audacity, available at audacity.sourceforge.net.

The Waveform window

The WAVEFORM setup window is available via the 'Edit' menu, but to see the waveform, you should first open a sound file attached to a sequence. If the sound file is not an AIFF file, this is reported in the information field at the top of the window. If the file is an AIFF file, the left part of this field provides you with a number of properties of the sound file, as explained in the Introduction section above.

In the right part of the information field, you will find information, about the part of the sound file corresponding to the present sequence (in case the same sound file is attached to a number of successive sequences).

Play buttons

Below the information field, you will find three play buttons at the left. With the  button, you play the sound from the position of the time slider: the slider with the red vertical line. The  button plays the selected part of the waveform. The  button plays the sound from the position of the time slider to the end of the selected part of the waveform.

Downsampling

To the right of the three play buttons is a slider, allowing you to resample the sound. Suppose the sample rate is 22050. If you set the resample rate to for example 50 percent, the sound will be resampled at 11025 frames per second. Resampling (or more precisely, downsampling) increases the speed of analyses. For most analyses a resampling rate of 8000 frames is sufficient (in case of human speech).

Normalize

If you check  **Normalize**, the amplitude of the sound is increased to the maximum possible values, given the sample format.

Channels

If you have two sound channels (stereo), you can select either channel, or mix both channels with the option menu after 'Channels:'. If only one channel is available, the setting of this option menu does not have effect.

Time slider

Below these controls is a slider. This slider or *time slider* allows you to move the playhead to a particular position of the waveform. A red vertical line moves together with the slider.

Magnify waveform

In the lower right corner you will find some buttons that allow you to magnify the waveform, either horizontally, or vertically (if **Normalize** is checked, vertical magnification is disabled).

Waveform selection

In the lower right corner you will also find some buttons to select parts of the waveform. Move the mouse over these buttons to obtain information about these buttons in the help field.

Finally you have the choice between a number of tabs, for particular analyses. These tabs are explained in the next sections.

Tip
Because you can only do waveform analyses on AIFF files, you cannot analyze the audio part of a movie. However there is a simple trick to accomplish this nevertheless.
Create an audio file from the movie. You can do this with audio applications like Audacity, Adobe Audition, or most video editing software. Just open the movie in such an application, and use the Export function to save the audio channel(s) as a separate (audio) file. Be sure to export the whole file, from start to end: it is important that the audio and video file are perfectly synchronized.
Link both the audio file and the video file to the sequence.
Open the Waveform window.
Select the video file as the linked movie. SEQUENCE VIEWER now looks if there is any AIF file linked to the sequence. If such a file is found, its waveform is created and drawn in the waveform field, and you can do all kind of analyses on the waveform.

3.7.1 Assign event times

To assign the STARTTIME, and the onset and offset times of a sequence using the audio waveform, first follow these steps.

- The event texts should already be coded. If not, assign empty codes.
- Be sure text and codes are linked.
- Open the AIFF sound file corresponding to the sequence.
- Check **Normalize**.
- Select the **Assign event times** tab.

To enter (or adjust) the STARTTIME:

- Select 'Time slider' with the option menu below 'Use:' and before the STARTTIME button.
- Move the time slider to the beginning of the sequence.
- If necessary (that is, if the width of one pixel represents less than one time unit), click the horizontal magnify button one or more times, in order to position the slider exactly at the point where the sound starts. But remember that you don't need more precision than allowed by the chosen time unit.
- Click the STARTTIME button.

To enter or adjust onset and offset times:

- Select 'Selection' with the option menu before the Onset time button.
- Select 'Selection' with the option menu before the Offset time button.
- Click at an event code in the main sequence window.
- Select **SV time unit**.
- Click at the button in the 'waveform selection' section in the lower right corner.
- Roughly select the part of the waveform of the sound corresponding to the selected event. Because you selected **SV time unit**, time units instead of pixels are used for the left and right borders of the selection.
- Click at the play selection button. The selected sound is played.
- If necessary (that is, if the width of one pixel represents less than one time unit), magnify the selected part of the waveform (the time slider should be somewhere in the middle of the selection).
- Adjust the left side of the selection, with the and buttons, and check the adjustment with the play selection button. You may note that in the 'waveform selection' section (the section with the selection buttons), the selection is shown in time units.

- Click the Onset time button.
- Position the time slider five or ten time units before the right side of the selection, and use the  button to play the sound from the time slider to the right side of the selection.
- Adjust the right side of the selection with the  and  buttons, and check the adjustment with the play selection  button.
- Click the Offset time button.

Below 'new value', you will see the values that will be assigned to STARTTIME, onset time, or offset time. These values may be shown in red. In that case there is some inconsistency in the times. For example, the new value for STARTTIME may be higher then the (already entered) lowest onset or offset time. Or the new value for the offset time of an event, is lower then the already entered onset time for that event. Move the mouse over a red 'new value', and the help field tells you what the problem is. You can assign a 'red' value nevertheless.

As soon as both onset and offset time of an event are assigned, the position of this event on the waveform is indicated with a green bar at the top of the waveform field. If there is overlap between two events, this is indicated with a dark green color (there is also overlap if the offset time of an event equals the onset time of the next event). If only the onset time or only the offset time is assigned (whereas the other has still the value 'M'), the position of the onset time or offset time is indicated with a small red bar. If the onset time of an event equals or exceeds the offset time, this is indicated with a red bar. A black bar means that there is overlap between three or even more events, or that a 'red' event overlaps a 'green' event.

Alternatively, you can also use the time slider to enter onset and offset times, in a way similar to entering the STARTTIME. If you first enter all onset times, and then all offset times, you may check  **select next event after assigning**, and select 'Onset times', respectively 'Offset times' with the popup menu below this checkbox. If checked, the next event code will be selected after you have assigned the onset (or offset) time.

If successive sequences use the same sound file (for example is the sound file concerns a whole interview, whereas the sequences concern question-answer sequences with respect to the different questions in the interview), it is recommended to first assign the STARTTIME to all sequences:

- Be sure the sound file is already linked to the sequences.
- Set the resample rate to the lowest value (10 percent) to speed up things.
- Go to the first sequence.
- Also be sure that the whole file is loaded (the scale should start with 0). If this is not the case, press the Load whole file button.
- It may occur that after you pressed the Load whole file button, the start or the end of the waveform is 'grayed'. Don't mind. Position the time slider at the start of the sequence, and click the STARTTIME button.
- Go to the next sequence, position the time slider at the start of the second sequence, and click the STARTTIME button.
- Continue until you're done.
- If you now go for example to the first sequence, only that part of the sound file is loaded, that is, from the STARTIME of that sequence, until the STARTIME of the next sequence. You may now set the resample rate to a higher value and enter the onset and offset times as described above.

3.7.2 Sound analyses

INTRODUCTION

To perform sound analyses, first select the part of the waveform you want to analyze, using the buttons in the 'waveform selection' section in the lower right corner of the WAVEFORM setup window.

You can create a graph of the sound characteristic and calculate a number of sound statistics.

Graphs

First select the particular sound characteristic you want to analyze with the corresponding tab. Then select the part of the waveform you want to analyze. To create the graph and calculate the sound statistics, just click the Get [characteristic] ('[characteristic]' depends on the selected sound characteristic).

The shape of the graph and the values of the statistics, depend on the settings of the downsample percentage, whether or not ☑ **Normalize** is checked, and on the selected option for 'Channels' (if two channels are available). Vertical magnification does not affect the graph and the statistics. Shape and statistics also depend on the additional settings, available for silence and pitch analyses.

You can change the color and line size of the graphs you create. For example, you can create a pitch graph with particular settings, change the settings and the color and create another pitch graph. The difference in color makes it easy to compare the effects of the different settings of both graphs.

To hide or delete a graph of a particular characteristic, first select its color, then click Hide or Delete. Click Show to show a hidden graph.

You can also change the color and or line size of existing graphs. First select a particular color, and than select a new color with the 'new colour' pull down menu, and/or a new line size with the 'new line size' pull down menu.

Loudness analysis

The loudness analysis does not have additional settings and is pretty straightforward. Downsampling usually does not have much effect on the shape of the graph and the statistics. Checking ☑ **Normalize** usually increases the values. To compare the loudness of different audio files, it is recommended to check this checkbox. To create the loudness graph and/or calculate the loudness statistics, click Get loudness.

Silence analysis

The silence analysis detects parts of the waveform that fall below a particular threshold. You can change this threshold and the minimum duration a silence should have with the sliders in the 'settings' section.

It is recommended that before you perform a silence analysis, you first create a loudness graph. If you change the silence threshold, the parts of the silence graph that are below the threshold, become hidden. In this way it is very easy to see the effects of different thresholds.

The minimum duration of a silence varies between 0.2 and 2.5 seconds. Only silence durations equal to, or exceeding this duration will be viewed as 'silences'.

If you click Get silence, the 'graph' is shown as a number of bars at the bottom of the graph section of the waveform. To make the bars better visible, it is recommended to select a line size of 2.

Pitch analysis

The pitch analysis is much less straightforward than the other two sound analyses. First of all, downsampling affects the results of pitch analyses considerably. You should not use a resample rate of less than 7000 frames per second.

Pitch correction

The reliability of the results of the analyses highly depends on the quality of the sound. If the sound is less than perfect, you can adjust for noise correction. The less the noise correction, the less fragments of the graph are assigned a pitch: only the fragments with a very clear fundamental frequency are detected, If you set the noise correction to 0, quite probably nothing will be detected at all, except fragments that approach a pure sine wave. If you set the noise correction to100, quite probably some fragments are assigned a wrong fundamental frequency.

In addition you can correct for high pitch and low pitch. If you are sure there are no high pitches in the sound file (e.g. in case of adult male voices), you can set the high pitch correction to a high number, e.g. 100. This prevents that a sound fragment is incorrectly assigned a too high pitch.
In other cases, if the sound is less than perfect, a too low value may be assigned to a sound fragment, and you should use the low correction.

The pictures above illustrate the principle. The fundamental frequency in both simplified waveforms is indicated by the red bar. It is easy to imagine however that you can have a waveform that is in between both waveforms. At what moment should the program decide whether the fundamental frequency is high (the first picture) or low (the second one)? This decision is affected by the settings of the high and low pitch corrections.
To decide upon which setting gives the most reliable results, follow these steps.

- Select a reasonable setting (e.g. the default settings: click Default settings).
- Select a color, and line size 3 for the graph.
- Click Get pitch to create the graph. Usually the graph is not continuous. Parts without a clear pitch (e.g. silence or noise) are not displayed.
- Listen to the sound. Different sound fragments may sound as having the same pitch, but show different pitches in the graph; or different sound fragments may sound as having clearly different pitches, but have similar values for the pitch in the graph.
- If this is the case, adjust the settings, depending on whether the pitch in one or more parts of the graph is lower or higher than the perceived pitch.
- If necessary, select a smaller part of the waveform to listen to.
- Select a different color, and a line size of 1. This allows you to see the differences between both graphs.
- Click Get pitch .
- Inspect both graphs. Quite probably you will see that some parts of the waveform now obtain a clearly higher or lower pitch (usually half or twice as high).
- Select fragments of the waveform with such differences and listen to the sound. This will tell you which graph represents the pitch best.
- Make further adjustments if necessary.

This may be a bit tedious, but once you have decided upon reasonable settings for a sequence, you can safely apply the same settings for other sequences that use the same sound file, or for different sound files, if these are recorded in a similar way.

Exclude sound fragments

Three other settings may also prevent that problematic parts of the sound are assigned a pitch value. These settings simply exclude such problemetic parts.

First, the speech of different speakers may overlap. For example in case the offset time of a particular event is higher than the onset time of the next event. If ☑ **Exclude overlap** is checked, overlapping parts in event times are excluded.

A second setting is the minimum duration of a sound fragment. If you look at the picture above, the first picture has six 'tops'. If part of the sound file contains a large number of such successive tops, you can be pretty sure that the fundamental frequency can be determined reliably. However, if there are only two or three such tops (preceded and followed by silence, or noise), it is very well possible that such tops only represent noise. If you set the minimum duration to a low number, the graph may contain a number of short parts that only represent noise, or are otherwise unreliably measured. If you set the minimum duration to a high number, you prevent such errors, but you may miss relevant parts of the waveform.

A final setting concerns the minimum loudness. If the sound is very soft, it is much more difficult to determine the pitch in a reliable way. You can set the minimum loudness to a higher number to exclude soft parts of the sound file.

Pitch statistics

Among the pitch statistics is also the 'variation' calculated. The variation is the standard deviation of the log 10 transformed pitch values. This is a common measure in linguistics for intonation of spoken pitch.

In interpreting the pitch statistics, you should take into account of course whether or not the speech concerns only one, or more persons. If for example, the statistics concern both a male respondent and a female interviewer, these statistics don't make much sense.

Usually however, you want to assign sound characteristics to event texts, uttered by only one person. In the next section is described how you can assign sound characteristics to event variables.

3.7.3 Assign sound characteristics

INTRODUCTION

Once you have decided upon a number of reasonable settings, you may want to assign sound (or speech) characteristics to event variables.

You can add up to 10 different characteristics: the median, mean and standard deviation of loudness' the number of silences, the average duration of the silences and the standard deviation of duration of the silences; and the median, mean, standard deviation and variation of the pitch. For the interpretation of these measures, please consult the linguistic literature.

Create an event variable for each characteristic you want to assign. Most of these characteristics are represented by floating point variables, so be sure to select an appropriate number of decimal digits for these variables. Use appropriate names for these variables.

Be sure the settings for calculating the sound characteristics are correct. You may also enter the values of the settings as a variable label in the DEFINE EVENT VARIABLES setup window.

Because the analysis uses the linked AV files, it is important that the links between sequences and AV file(s) are still valid. For example, if you moved an AV file to another location on the hard disk, the link is no longer valid. You can check whether all links point to an existing file, with Check data (**Sequences > Check data...**).

SETUP

- Click the Assign... button.
- You obtain a sheet where you can specify which characteristics are assigned to which variables.
- Check the appropriate checkboxes, e.g. ☑ **Assign median to:**, and enter or select an event variable in the usual way.
- Click the Assign button.

ADDITIONAL INFORMATION

The program uses the onset and offset times of each event, to identify the corresponding part of the audio file. This part is used for the calculation of the sound characteristics. In case of large files, be prepared that the analyses can take a long time. Especially the pitch analysis is time consuming. If you only want to assign loudness and silence variables, you can downsample to for example about 2500 frames per second, to speed up calculations, without much loss of precision.

If an event has no onset or offset times, you can choose to assign a 'M' for missing value, or leave the present value unchanged.

If a sequence has no linked AIFF audio files, you can also choose to assign a 'M' for missing value, or leave the present values unchanged. If more audio files are linked to a sequence, the program selects the first AIFF file for that sequence. However, if an AIFF file was selected for a particular sequence, the program selects the same AIFF file for the next sequence (if the same audio file is linked to different sequences).

3.8 Command editor

INTRODUCTION

The COMMAND EDITOR allows you to execute a series of commands in succession. A series of commands is called an execution script, and can be saved to disk. Such a saved script can be reloaded, and executed again.

The simplest way to create an execution script (or simply 'script'), is to record the setup of one or more setup windows. If you are new to the COMMAND EDITOR, we strongly recommend you to do the exercise described in section 10.1 of the tutorial, first.

SCRIPTS

Scripts consist of three different types of lines. The first type is called (menu) *command* lines, the second *setup* lines and the third *comment* lines. Command lines appear in bold face in the script field, and are indented (that is, preceded by a tab). Setup lines are in plain face, without indentation. Comment lines are italic and preceded by '>'.

Command lines

Command lines usually directly refer to menu items available from the menu bar. Execution of such a command is essentially a call to a menu item from the menu bar. These command lines consist of the name of the menu (without '...'), a colon, and the name of the menu item. In case of subitems, another colon with the name of the subitem is added.

The exception is a number of commands that are executed via the main sequence window, like going to a next sequence, or marking a particular sequence. These command lines start with 'Sequence window', followed by a colon and a description of the action; for example:

Sequence window : Go next sequence.

A distinction can be made between menu commands that open a window, allowing the user to specify a setup, and commands that perform an action immediately (like marking all sequences).
Commands that open a window have a number of associated setup lines, describing the state of checkboxes, radio buttons, edit boxes, and so on.

Setup lines

Setup lines always consist of three elements, delimited by a colon. The first element describes the kind of control, e.g. a 'Radio button' or an 'Edit field'. The second element designates the name of the control, and the third one its value. The different control types are described below.

Checkbox

The name of a checkbox always equals the name as appears in the setup window. Two values are possible: 'on' and 'off'. Example (from TREE):

```
Checkbox:One branch per sequence:on
```

Radio button

The name of a radio button always equals the name as appears in the setup window. Although a radio button can be on or off, only the value 'on' is possible. The reason is that radio buttons always appear in groups; if one radio button of a group is on, the others are always off. To reduce the length of the script, it is only necessary to specify the radio button that is on; SEQUENCE VIEWER automatically sets the other radio buttons of the same group to off. Example (from ADD CODE VARIABLE):

```
Radio button:Number:on
```

The radio buttons used for recoding codes in the SEQUENCE DEFINITION window, behave a bit different from the radio buttons in other setup windows. Usually radio buttons are grouped, and the selected radio button represents the 'value' of that group. In the SEQUENCE DEFINITION window, the highlighted radio button only tells you which code variable is recoded, i.e. to which code variable the original codes and the new code applies. The value of each radio button in this window is the recode specification, for example:

```
Radio button:Recode code variable 3:134=G
```

Tab

A tab menu lets you choose between a number of different setups in a window. For example, the FIND VALUE OF SEQUENCE VARIABLE setup window allows you to choose between **Find**, **Replace**, **Mark** and **Unmark**. Each tab provides you with a different set of controls, hence in the script of a window, the 'Tab' line should always precede the other setup lines. The name of a tab is always 'Select'. Because there will never be more than one tab in a setup window (and because the 'Tab' line will always be the first line), no confusion is possible. The value of a tab menu is the (name of the) selected tab, except in case of the information windows; these may either have the standard names 'info 1', 'info 2', etcetera, or custom names provided to the information fields (see customize), or both names, separated by a slash (see also recording custom or default names). Example (from Find value of sequence variable):

```
Tab:Select:Replace
```

Pop-up menu

Pop-up menus (or option menus) usually provide you with a number of different options. In most cases a description of the pop-up menu appears in front of the pop-up menu. This description (sometimes in abbreviated form) serves as the name of the pop-up menu. The value is in most cases the selected option. Like 'Tabs', this is only partly true for pop-up menus whose menu items you can change (customize), i.e. the names of sequence definitions, or the names of mark sets. These either have the default names 'Sequence definition 1', 'Sequence definition 2, 'Mark set 1', Mark set 2', etcetera, or the custom names you have assigned to them, or both the default name and the custom name, separated by a slash. See also recording custom or default names and the section 'Custom and default names' below. Example (from TEXT ANALYSIS):

Pop-up menu:Count type:literal

Pull down menu

Pull down menus usually let you select a particular value (which of course is also the value in the script line of a pull down menu). Like pop-up menus, the name of a pull-down menu usually appear in front of the menu. Example (from TREE):

Pull down menu:Tree depth:5

Edit field

Edit fields (or edit boxes) can contain any kind of information. In most cases the name of an edit field appears before, or above the edit field. The value is simply the content of the edit field. Values of edit fields should have double quotes. This allows you to enter a colon in the text of an edit field, without this colon being viewed as a delimiter between the different parts of a setup line.

Edit fields are also used to display event codes like 'IQ1'. The number of characters (codes) should correspond to the number of code variables. If the number of characters of an event code in the setup line exceeds this number, the characters at positions after the number of event codes are deleted. If the number of characters is less than the number of event variables, the event code of the setup line is filled up with question marks or empty codes ('-'), whichever is more appropriate. Example (from TREE):

Edit field:Start:"????????"

Icon

Many windows have a number of icon buttons at the bottom of the window. The most important ones are 'Define sequence', 'Selected sequences', Marked sequences' and 'Checked events only'.

- 'Define sequence' can have the value 'No sequence definition' or the name (default name, custom name, or both) of the selected sequence definition. The specification of a sequence selection or sequence definition, should always be made in a separate script, preceding the script that uses the selection or definition.
- 'Marked sequences' can have either the value 'off', or the name of the selected mark set. The name can be the default name, the custom name, or both, separated by a slash.
- 'Selected sequences' can have the values 'Yes' and 'No'.
- 'Checked events only' can have the values 'Yes' and 'No'.

Example:

Icon:Marked sequences:Female/Mark set 4

Push button

Although a setup window may contain more than one push buttons, the script describing the setup of a window, always contains only one push button: the button that executes the setup (the OK or similar button). In the COMMAND EDITOR the corresponding script line executes the script. The name of a push button is always the name as shown in the push button of the setup window. Only one value ('Press') is possible.

Executing the push button setup line is identical to pressing the button in the setup window, except that no messages (e.g. warnings) are shown. For example, if you press 'Delete' in the DELETE SEQUENCE VARIABLE setup window, you first obtain a warning, asking you if you are sure to delete the variable. You will not obtain this warning if you execute the delete command from the COMMAND EDITOR. Similarly, if you have made changes to the data file, and execute the **File > Close** command by means of the COMMAND EDITOR, you are not asked if you want to save the changes; the file is just closed, without saving the changes (you can precede the 'Close' setup line with a 'Save' setup line of course).

Comment lines

Comment lines can be inserted at any place in the script. They should be preceded by the character '>'. Comment lines are not executed.

Comments can be entered using the Edit field (below the Script field. Click at an empty line in the script, type any comment you want (but be sure to precede it with '>') and press Change. You can insert an empty line with  .

SYNTAX

In the script field, colons are always preceded and followed by a space. This is only for readability. A colon without accompanying spaces does not cause an error.
Script lines are not case sensitive; a line like 'Push button : OK : Press' is equal to 'push button : ok : press'. Of course, event codes that appear in setup lines, are case sensitive.

Custom and default names

Mark sets, information fields and sequence definitions have default names like 'Mark set 1' or 'Info 2', etcetera. These default names can be changed into custom names (with FILE SETTINGS, tab **Customize**). For example, changing 'Mark set 1' into 'women' makes it easier to remember which type of sequences are marked with a particular mark set.

In setup lines you can use both default names and custom names. Custom names have the advantage that they make setup lines more readable and understandable. However, if the custom name does not exist in the file that is executed (for example because the script is created by a different SEQUENCE VIEWER file), you will obtain an error message. Default names are always recognized, making it easier to use saved script files with different SEQUENCE VIEWER files.

To allow for more flexibility and better exchange of scripts between different files, it is allowed to enter both the custom name and the default name.

If you record a setup, you can choose if you want default names or custom names in your setup. You can also choose to record both the custom name and the default name (they will be separated by a slash in the setup line). In case the command editor does not recognize the custom name, it will use the default name. To select one of these three options, open PREFERENCES, and click the **Editor** tab. Select the appropriate item with the 'Record custom or default names' pop-up menu.

Invisible and disabled controls

Controls (checkboxes, radio buttons, etcetera) may be invisible in setup windows. For example, the FIND VALUE OF SEQUENCE VARIABLE setup window shows or hides controls, depending on the tab setting (**Find**, **Replace**, **Mark** or **Unmark**). Invisible controls are not recorded. A setup line that specifies a value for an invisible control is not executed, but generates an error message instead.

Controls in a setup window may also be disabled, depending on the setting of other controls (which is one of the reasons that the order of the setup lines is important). Disabled controls don't need to be recorded. However, you can choose to record disabled checkboxes and edit fields nevertheless. To this end, open PREFERENCES, click the **Editor** tab and check ☑ **Record disabled checkboxes and edit fields**. This makes it easier to edit existing scripts (as relevant controls are already incorporated in the script). A setup line that sets the value of a disabled checkbox or edit field is neglected; the setting is not changed.

If the OK or similar button (the button that executes the setup in the setup window) itself is disabled (for example because of an error in the setup), you will get an error message, because the OK button should be enabled to perform the command.

RECORDING SCRIPTS

It is very easy to create a script with the COMMAND EDITOR. As an example we will create a script for a tree analysis.

Select **Sequences > Command** editor to open the COMMAND EDITOR window.
Be sure that an empty line is highlighted in the script field.
Use the first pop-up menu after 'Select' to select 'Codes'.
Use the second pop-up menu after 'Select' to select 'Tree'. As you see, the menus correspond to the items in the menu bar.
Click the Enter button in the upper right corner. A command line (Codes : Tree) is added to the script field. In addition the TREE setup window is opened.
In the TREE setup window you can now enter any setup you want, in the usual way.
When you are finished, click the Record button, in the upper right corner.
That's it. Your script is ready to execute.

You can enter scripts for different setup windows below the tree script. Or you can add commands from the menu bar without an own setup window like **Sequences : Mark sequences : Mark all**. Or you can select 'Sequence window' with the first pop-up menu, and than 'Go next sequence' with the second pop-up menu.

EDITING SCRIPTS

Editing script lines

Command lines cannot be changed within SEQUENCE VIEWER (but you can write a whole script with a word processor, save it as 'text only', and load it into the COMMAND EDITOR). Command lines are defined by the names of menus and menu items, and there is no need to change (edit) such a line: it is faster and easier (without the risk of making syntax errors), to enter a command using the pop-up menus in the COMMAND EDITOR.

This is different for setup lines. It is often easier to change for example one single value in a script than to record a whole script. To change a setup line, click at that line; the line appears in the change field below the script field. Edit the script line; usually you should only change the value part of the line. Click the Change button to replace the old script line with the edited one.

However, you should be careful with editing setup lines. For example, checking or unchecking a checkbox, or selecting a radio button, may automatically change the values of other buttons, or enable or disable other buttons. It is safer to re-record the script. To re-record a particular

script, double-click on the menu command line in the script field. The window related to the command line will open. Execute the script, except the last line (the 'Push button' line). The setup of the window is changed according to the script. Change the setup in the window, and press the Record button. The script in the script field is automatically replaced.

Order of setup lines

The order of the setup lines in a script is in many cases very crucial. Generally, the order corresponds to the order of the controls in the setup window, i.e. from top to bottom. If a setup is recorded, the order will always be the correct one. Never change the order in editing the script.

You can however change the order of command lines (including the related setup lines) with the  and  buttons. Click at a command line to select it. Then click the  button to move the command line, including its setup lines, up. Click the  button to move the command line with associated setup lines down.

Deleting and inserting script lines

To insert an empty line before the selected line, click the  button. To delete a script line, select it and click the  button.

If you delete a command line, also the associated setup lines are deleted.

RUNNING A SCRIPT

To run a script:

- Select a command line
- Click the Run button

The script, starting with the selected command line, and ending with the last line in the script field is executed.

Instead of the Run button, you can also click the Trace button or the Step button. The Trace button does the same as the Run button, except that the script lines are executed with some delay, allowing you to follow the execution of the script. To change the delay time, open PREFERENCES, click the **Editor** tab, and adjust the delay with the 'Delay between executing setup lines in Trace mode' slider.

The Step button executes the script lines one by one. To use the Step button, you should always start with a command line.

Testing a script

You can test a script by checking the  **Test** checkbox and run the script. In the test mode:
The relevant setup window is opened (if the command line refers to a setup window).
The execution line (the 'Push button' line) is not executed.
Command lines without setup window (that is, command lines that immediately perform an action, like marking all sequences) are not executed.
Command lines that refer to the main sequence window (that start with 'Sequence window', for example to go to a next sequence) are not executed.

To test a script, you best use the Step button. Select a command line, and proceed step by step through the script. If a command line uses a setup window, this window will be opened, and you can observe all adjustments made to the setup window.

If an error is reported it is very well possible that this is not a real error. Suppose for example, that a particular script line adds a sequence variable to your file. In the test mode such a change

is not executed however. If a script line after this script refers to this new variable (e.g. changing the value of the sequence variable you have added), in the test mode you will obtain the error message that the sequence variable does not exist.

Errors in scripts will usually lead to errors in the window setup. In such a case the OK (or similar) button of the setup window that executes the setup, will be disabled, and the COMMAND EDITOR gives an error message. Move the mouse over the help button of the setup window to see what exactly is wrong with the setup (shown in the help field of the toolbar).

Executing a script

To execute the script, uncheck **Test**, and click the Run button. If (despite testing) an error is encountered, the script stops and an error message is displayed. If the script concerns a setup window, this window is opened, allowing you to figure out what exactly caused the error. After you have changed the setup window, you can record the new setup, and continue execution; just select the menu command line related to the setup window, and run the script.

Executing part of a script

It is easy to execute only part of a script. Select a command line, and set a break point ( ) in front of the line were you want to stop execution. To set a break point, just click before the line were you want to set the break point. To remove a break point, click on the break point itself. Click the Run button to execute the script from the selected command line to (not including) the line with the breakpoint.

SAVING AND LOADING A SCRIPT

You can save a script for later use, or to share it with other users. Scripts are stored in a separate folder on your hard disk. First of all you should specify the script folder; SEQUENCE VIEWER should know where this folder is located to be able to run scripts using the toolbar (see Running scripts), without the need to open the COMMAND EDITOR. Go to PREFERENCES, click the **Editor** tab, click the Browse button, and choose your default script folder. Click OK and return to the COMMAND EDITOR.

To save a script, click Save, enter a name for the script, and click Save in the 'Save file' dialog box. Be sure the default script folder is selected.

To load a script, you should first decide whether or not to replace any existing script in the script field. To insert a saved script into the existing script, select an empty line in the script. This empty line should either be the last line in the script, or an empty line before a command line. Check **Insert**, and click Load. The selected script is added to the existing script.

If you want to replace the existing script, just leave **Insert** unchecked, and click Load.

See also:
Editor (Preferences), Auto recording, Running scripts (toolbar).

3.9 Toolbar

The toolbar can be found in the upper left corner of the screen, and consists of six different parts:

- The auto recording box
- The running scripts box
- The edit text box
- The text layout box

- The convert date box
- The help field

In the next subsections, these different parts will be discussed in more detail.
The Toolbar menu item of the **Edit** menu, makes it possible to show or hide the different parts of the toolbox. For example, if you never auto record scripts, you can hide this part of the toolbar with **Edit > Toolbar > Hide autorecording**.

3.9.1 Auto recording scripts

You can automatically record execution scripts without having the COMMAND EDITOR open.

Click the **Record** button  on the toolbar in the upper left corner of your screen.
The 'record light'  turns into red, whereas the 'stop' button  becomes enabled.
Setup windows that allow auto recording can be recognized by a red dot to the left of the OK (or similar) button. As soon as you click this button, the execution script will be automatically recorded.
In addition to setup windows, also a number of actions from the main sequence window will be auto recorded (i.e. those commands that are available in the COMMAND EDITOR to manually create a setup script).
With the Neglect button you can suppress auto recording of commands from the main sequence window. Clicking this button gives you a list of these commands. For example, if you check  **Go next sequence** from this list, this command will not be auto recorded if auto recording is on.
SEQUENCE VIEWER stores the checked commands in the preferences file, and thus the list of commands that should be neglected is remembered a next session.

If the COMMAND EDITOR is open when scripts are auto recorded, you can watch how the scripts are recorded, but the COMMAND EDITOR is locked. You can only close and resize the window, but clicking on buttons or fields will not have any effect.

3.9.2 Running scripts

You can assign a saved script to a menu item of the scripts pop-up menu, which allows you to execute a script without having the 'Command Editor' open. To this end, you should first assign a script to a menu item.

ASSIGN A SCRIPT TO A MENU ITEM

- Click the Edit button on the toolbar. If there is no such button, you may select **Edit > Toolbar > Show edit script** first. The EDIT SCRIPT window appears as a drawer below the toolbar.
- Click Add. Select a script from the default scripts folder.
- In the EDIT SCRIPT window, enter a name for the script (after 'Script name').
- In addition you can enter a short description of the script.
- Click OK.

The script names are added to the 'Select script' pop-up menu, just above the Edit button. You can add as many scripts as you want. You cannot use the same script name for different scripts; neither can you add a script that is already assigned.

You can change the order of the script names (and hence the order of the menu items of the 'Select script' popup menu) with the  and  icons.

To delete a menu item from the script pop-up menu, select the menu item and click Delete.

RUNNING A SCRIPT

To run a script, select the script name from the 'Select script' pop-up menu. Than click the Run button. The Run button is disabled, if auto-recording is on, or the COMMAND EDITOR window is open. If an error is encountered, the COMMAND EDITOR window is opened with the script and an error message (see Command editor).

HOW IT WORKS

The menu items from the script pop-up menu, refer to the saved file that contains the script. For this reason the saved script should be in the default script folder, for the program to be able to find the file.

When you click Run, the script file is read, put into the script field of the COMMAND EDITOR, and executed, all invisible to the user.

If the script encounters an error (e.g. because a script line concerns a sequence variable that is not in the presently open SEQUENCE VIEWER file), the COMMAND EDITOR is opened with the script that contains the error.

The relevant information for each menu item in the script pop-up menu (script name, description and reference to a file) is stored into the SEQUENCE VIEWER preferences file. Thus, once created, the items from the script pop-up menu, remain available in next sessions.

If a script file cannot be found (e.g. because you trashed that script file, or changed the name of the script file in the finder), you obtain an error message.

3.9.3 Edit text

INSERT SYMBOLS

The format box allows you to easily insert symbols. A number of common symbols that are used for transcriptions, are available after 'Insert'. Click one of these symbols to insert it. Or you can select a wider range of symbols with the 'Other' pull down menu.

Symbols can only be inserted if symbols are allowed, e.g. the sequence text, or the Find box to find text in the sequence text. Names of variables for example should not contain symbols, and it is not possible to insert them in edit boxes for variable names.

Instead of using the format box, you can also use the option key together with some other key to insert a symbol. The up, down, left and right arrows can be inserted using the option key and one of the arrow keys.

TEXT STYLE

Text styles can only be applied to the sequence text. Select a piece of text in the sequence text, and choose a text style with the pull down menu. Available are plain, bold, italic, underline and strike out. More text styles can be applied to the same piece of text; e.g. both italic and bold is possible. If a piece of text has a particular text style, and you select that text style from the pull down menu, that text style is removed. 'Plain' removes all text styles.

TEXT KEYS

To assign a text key to one or more words of the sequence text, select these words, and choose one of the existing text keys with the 'Text keys' pull down menu. Text keys can only be assigned to whole words; if you select only part of a word, the text key is assigned to the whole word. If a word already has a text key, selecting that key will remove it.

3.9.4 Text layout

You can change the indentation of the first line of each paragraph in the sequence text. The main sequence window should be in codes and text mode. Move the slider to the left or to the right to change the indentation. The change is immediately visible in the main sequence window. Only the indentation of present sequence is changed; if the **all** checkbox at the right of the slider is checked, all sequences obtain the new indentation.

You can also change the spacing between paragraphs (and the linked codes, if codes and text are linked). Click a number of times at the up or the down triangles, until you're satisfied about the spacing. Here too, only the spacing of the present sequence is changed, but if the **all** checkbox is checked, all sequences obtain the new spacing.

3.9.5 Date/time conversion

With the 'data/time conversion' box in the toolbar you can easily change a human readable date or time stamp into a number (the number of days since January 1, 1970, respectively the number of seconds since January 1, 1970) and vice versa.

First select the type you want to convert with the 'Convert to/from' pop-up menu. Than either enter a human readable date or time stamp in the upper edit box, or enter a number in the lower edit box.

3.9.6 The help field

If you move the mouse over buttons and fields in a setup window, a description of this object appears in the help field of the toolbar.

The help field also provides you with information about possible errors in setups. Especially with more complex setup windows, the OK (or similar button) may be disabled because of errors in the setup, but it maybe unclear to you what the error is. In such a case, just move the mouse over the help button, and the help field will tell you what is wrong with the setup.

4 Sequences

Sequences are the primary units of analysis in SEQUENCE VIEWER. A file consists of a number of sequences, whereas each sequence consists of a number of events. See the Introduction for more information.

4.1 Define sequence

INTRODUCTION

With the DEFINE SEQUENCE setup window (**Sequences > Define sequence...**) you can perform a number of transformations on your sequences. All transformations are temporarily. For example, if a transformation concerns the removal of events, the removed events are by no means deleted from the data file. Only the analyses are performed on the transformed sequence.

You can observe the effects of transformations on your sequences. To this end:

- Define a sequence, that is, apply one or more transformations (explained in detail below) on the sequences in your file.
- Click Apply. The button Done also applies the sequence definition, but in addition closes the window.
- Set the layout mode of the main sequence window to codes only or to codes and event variables.
- Check  at the left side of the main sequence window to apply the sequence definition to the present sequence.

TRANSFORMATIONS

Mask events

With mask events you can neglect one or more code variables. First check the ☑ **Mask events** checkbox. A number of checkboxes appear, each corresponding to a code variable. Check the checkboxes of the code variables you wish to mask. Masked codes are shown as '•'.

If you only have one code variable, masking events does not make sense, and this option is not available. It is not possible to mask all code variables, as this neither makes sense.

Unite events

If two ore more succeeding events are similarly coded, you may view them as one event. Check ☑ **Unite events** to accomplish this.

The effect of 'uniting events' depends on 'masking events'. Two succeeding dissimilar events, may become 'similar', because of masking particular code variables.

Also the order of the masking and the uniting transformations may make a difference. To first apply 'uniting' and then 'masking', hold the mouse down at the button with the triangle () in front of ☑ **Mask events** and drag the whole line with mask events options to a position between ☑ **Unite events** and ☑ **Subsequence**. You may observe that the 'Mask event' line changes from grayed/transparent to normal, indicating that this is a legal position for the line. Leave the mouse and observe that the order of both transformations is now reversed. Use the  icon in the main sequence window to observe the effects (the layout mode of the main sequence window

should be 'codes only' or 'codes and event variables'; see Introduction above).

If events are united, you may wonder what happens with the event variables. If events have time information, the onset time of the united event will equal the onset time of the first event of the united events, whereas the offset time of the united event will equal the offset time of the last event of the united events. With respect to the other event variables you have a choice:

- You may assign the values of the first event of the united events to the united event.
- You may assign the values of the last event of the united events to the united event.
- You may assign the sum of the values of the united events to the united event.
- You may assign the mean of the united events to the united events.

Just select one of these options with the pop-up menu after ☑ **Unite events**. The choice is only important of course in case event variables are involved in the analyses on transformed sequences.

Subsequence

First check the ☑ **Subsequence** checkbox. There are three different ways to define a subsequence.

The first way is by position. Select 'position' with the pop-up menu after the Subsequence checkbox. To select event number 1 to 10 of a sequence, type these numbers in the fields after the pop-up menu. You can use the character 'L' to designate the last event in a sequence. For example, 'position L to L' just means the last event. You may also use an expression like 'L-1'. For example:

- 'position 1 to L-1'means all but the last event
- 'position L-3 to L' means the last four events.

A second way to define a subsequence is using particularly coded events. Select 'event' with the pop-up menu after the ☑ **Subsequence** checkbox. Type event codes in both fields. A question mark means any code. Hence, all question marks in the first edit box means the first event. If you type all question marks in the 'to:' box, this has the special meaning of the last event code. Codes that correspond to a masked event variable show up as a '•' and you cannot change such a character.

If events have time information, you can define a subsequence from a particular time to another particular time. Select 'time' with the pop-up menu after the ☑ **Subsequence** checkbox and type two time values.

Remove events

You can remove particular events from a sequence. To do this, first check the ☑ **Remove events** checkbox. Now you can type event codes in the edit box after this checkbox. Again, a question mark means any code. For example, an event code like 'I??' removes all events with the value 'I' on the first code variable from the sequence.

You can also specify conditions by using the pop-up menu in the 'Remove events' line. If you select 'after', a second field appears, where you can type an event codes. For example:

Remove I?? after R??

only removes event code I?? if it occurs after R??.

In a similar way:

Remove I?? before R??

removes all events coded with I??, if they are followed by R??.

To remove all events except I??, you should select 'except' with the pop-up menu.

Recode events

You can recode code variables in the following way. First check the ☑ **Recode events** checkbox. Now a number of radio buttons appear, corresponding to the code variables (if there is only one code variable, no radio buttons will appear, as the recode can only apply to the only one code

variable). Select the radio button corresponding to the code variable you want to recode. Type the codes to be recoded in the field after the radio buttons and type the new code in the field after =.
You can select another radio button, to recode a second code variable. You may note that the former radio button is surrounded by a square, informing you that the corresponding code variable is recoded.

Event checks

To keep either the checked or unchecked events of a sequence, check the **Event checks** checkbox, and select the appropriate item from the 'Remove' pop-up menu.

ORDER OF TRANSFORMATIONS

As discussed by 'Unite events', the order of the transformations definitely affects the eventual transformed sequence. You can easily change the order of the transformations by using the buttons with the triangles in front of each transformation line (, and). Click at a triangle, hold the mouse down, and move the whole transformation line to a lower or higher position. Use the data file window to watch the effects of different transformations and different orders.

ADDITIONAL INFORMATION

Applying a sequence definition from analyses windows

For many analyses the SEQUENCE DEFINITION window is also available from the analysis window itself. In that case you will find the icon DEF in the lower part of the analysis window. If you option-click on this icon, the SEQUENCE DEFINITION window appears. If you have defined a sequence, you can switch between using the defined sequence and not using the defined sequence by just clicking on the DEF button. It will turn into DEF if a sequence definition will be applied, whereas in the lower part of the window the name of the sequence definition will appear.

Saving sequence definitions

You can save nine different sequence definitions for later use. As soon as you click Apply or Done, the sequence definition is stored in the sequence file, and saved as soon as the sequence file is saved, together with the file. You can also change the name of a sequence definition with the CHANGE NAME icon at the top of the window.

Entering event codes

In setup windows where you can enter codes, usually a field is available that show you the allowed codes. You can click at such a code to enter it in the event code. In the SEQUENCE DEFINITION window, this field becomes available if you click at the 'show codes' triangle ( ) in the upper right corner.

4.2 Add sequences

To add a new sequence to your file, you can click at the button in the upper left corner of the data window. The new sequence will be added after the present sequence. To add a new sequence before the present sequence, hold the option key down when you click the button.

More options are available with the ADD SEQUENCES setup window.

- Select **Sequences > Add sequences...** to open the ADD SEQUENCES setup window.
- Decide upon whether or not you want to copy information from another sequence to the sequences that are added. If yes:
- Check **Copy information**.
- Type the number of a sequence in the 'from sequence' edit box, or go to a sequence in the main sequence window. Information from this sequence will be copied and pasted to the sequences that are added.
- After you have selected the sequence, be sure to check **lock**, to prevent that the sequence number changes if you go to a different sequence.
- Check all checkboxes of the information you want to copy to the new sequence(s).
- If you check **Copy sequence variables**, a field with variable names appears. Click at the names of the variables whose values you want to copy to the new sequence(s). To deselect a variable name, just click it once again. The variable SEQSIZE cannot be copied; its value for the new sequences will be 0, unless you also copied event codes. In that case the value on SEQSIZE equals the number of event codes.
- If you do not want to copy information from a sequence at all, uncheck **Copy information**.
- Type the number of sequences you wish to add in the 'Add' edit box.
- If you want to add the new sequences after the last sequence, select this option from the 'add where' pop-up menu.
- If you want to add the new sequences before or after a particular sequence:
- Be sure **lock** is checked
- Go to the sequence before or after which you want to add the new sequences.
- Select 'before sequence n', or 'after sequence n'. 'n' is the number of the sequence that is presently open. If you scroll through the sequences in the main sequence window, this number (n) will change accordingly. In this way you can select where you want to add the new sequences.

4.3 Delete sequences

To delete a sequence from your file, you can click at the button in the upper left corner of the sequence window. This will delete the presently open sequence. If your file has only one sequence, this button has the effect of clearing all information of this sequence (event texts, event codes, keys, etcetera), whereas sequence and event variables obtain 'M' for missing value (a SEQUENCE VIEWER file always contains at least one 'empty' sequence).

Some more options are available with the DELETE SEQUENCES setup window:

- Select **Sequences > Delete sequences** to open the DELETE SEQUENCES setup window.
- To delete a range of sequences, click at the **Sequence** radio button and type the number of the first and the last sequence in the range.
- Select **All marked sequences** to delete all marked sequences
- Select **All sequences** to delete all sequences (one empty sequence will be left, however). This is for example useful if you want to make a new file with exactly the same settings (like number and names of sequence variables), because names of sequence and event variables, their properties, names of code variables with allowed codes, file settings, et cetera are retained.

Also locked sequences will be deleted. If you want to delete all sequences, except the locked ones, unmark all sequences first (**Sequences > Mark sequences > Unmark all**). Then mark the locked sequences (**Sequences > Mark sequences > Mark locked**), and select the **All marked**

sequences radio button.

4.4 Sort sequences

To sort the sequences by a sequence variable, do this:

- Select **Sequences > Sort sequences...** to open the SORT SEQUENCES setup window.
- Select or enter a sequence variable.
- Select ascending or descending.
- Click Sort.

Sequences are sorted according to the values on the selected sequence variable. If this variable has missing values, these sequences are placed before the other sequences in case of an ascending sort, and after the other sequences in case of descending sort.

4.5 Split sequences

INTRODUCTION

You can split a sequence into two or more different sequences. The original sequence is kept and all parts of the original sequence are added to the end of the file as separate new sequences.

All information from the original sequence concerning the whole sequence, like sequence variables, information fields, or AV links, is copied to the new sequences. The marked property is not copied. Word links are only copied if both start and end fall within a new sequence. The sequence variables SEQSIZE and STARTTIME are adjusted.

SETUP

- Select **Sequences > Split sequences...** to open the SPLIT SEQUENCES setup window.
- Select if you want the split after or before a particular code, or between two particular codes.
- Type the event code (or event codes in case you want the split between two event codes) where the split should take place.
- You can select whether you want to split the sequence at the first instance the specified code is found in the sequence, the last instance, or all instances.
- You can mark which sequences are actually split, to enable you to inspect those sequences.
- Click Split.

ADDITIONAL INFORMATION

If you want to trace from which sequence the split sequences originated, you can add a sequence variable and assign the sequence number to this variable.

You can also choose to mark those sequences that are actually split. In that case you may also choose to unmark all sequences first. This will not affect the selection of marked sequences. If you use the marked sequences  option to split only sequences that are marked, marked sequences are first split; then both 'Mark options' are applied).

Also locked sequences will be split (because the original sequence remains untouched).

If you selected 'before' and the very first event in the sequence is found, the sequence will not be split at that event. The same is true if you select 'after': no split can be performed after the last event. No split is performed if codes and text are not linked, unless there is no text at all.

You can choose that the split is only performed if the event is checked with the  icon at the bottom of the window. This option is not available in case of a split between event codes.

See also:
Select sequences, Select marked sequences, Checking event codes.

4.6 Add events

INTRODUCTION

You can add an event code at a particular position in a sequence:
- before the first event of a sequence
- after the last event of a sequence
- before each event with a particular code
- after each event with a particular code

SETUP

- Select **Sequences > Add events...** to open the ADD EVENTS setup menu.
- Type an event code after the 'add' prompt. This is the event code that will be added.
- Select a position with the 'add where' pop-up menu. If you select 'Before' or 'After', an edit box appears where you can type the event code of the event before or after which the new events should be inserted. Please note that a new event is inserted before or after *each* instance of that event code.
- Type a text in the 'add event text' edit box. This text serves as the event text for the event. If ☑ **Use first character of event text as first code** in the FILE SETTINGS window is checked, the first character of the event code in the 'add' edit box is added to this text. Generally, you should type a text for the event, except in case you want to add events to sequences without event text at all. See below for a further explanation.
- Click OK.

ADDITIONAL INFORMATION

You may, or you may not type a text that should be added as event text. The following situations can be discerned with respect to a sequence.

(1) If you type a text in the 'add event text' edit box, the event will not be added if it becomes unclear to which event code this text belongs. For example, if the number of event codes does not equal the number of paragraphs of the sequence text, and you want to add an event code after for example the third event code, it is unclear where the new event text should be inserted in the sequence text.

(2) If you do not type a text in the 'add event text' edit box, insertion of an event code somewhere in the sequence, without inserting a corresponding event text, will break the links between the successive event codes and their event texts; hence the event will not be added. If there is no text at all in the text field, the event will be added however.

Generally, if a sequence has both event codes and event texts and these are correctly linked, the event will be added if you type a text in the 'add event text' edit box. In all other cases, the event will only be added if no ambiguity is created. For example:

If you have selected 'At end', the event will always be added if no text is typed in the 'add event text' edit box.
If you have selected 'At start', the event will be added if the sequence has no text at all, whether or not text is typed in the 'add event text' edit box.
If the sequence has no text at all, the event(s) will always be added if no text is typed after add text.

See also the explanation at Delete events.

See also:
Select sequences, Select marked sequences.

4.7 Delete events

To delete an event with a particular event code from your data file:

Select **Sequences > Delete events...** to open the DELETE EVENTS setup window. Type an event code after the 'delete event' prompt. A question mark means any code. If the event code consists of only question marks, all events of all (selected) sequences are removed, but the sequences themselves, including the sequence variables, remain.
Removing an event means that the event code, the event text and the event variables belonging to the event, are removed. Text links that start or end at a word of the removed event text, are deleted too.

If the number of event codes does not equal the number of lines (paragraphs) of the event texts, the event will not be removed, as it is unclear which paragraph of the text in the text field belongs to the event code. Hence, you may first link all codes and text (**Sequences > Link code and text...**), use **Sequences > Mark links...** to mark the sequences with bad links, and find the marked sequences to look why codes and text are not linked.
If the text field is empty, the event code with the event variables will be removed, because such ambiguity about which part of the text in the text field should be removed, does not exist.

See also:
Select sequences, Select marked sequences, Checking event codes.

4.8 Combine events

INTRODUCTION

You can combine two or more successive events if:
the successive events have *exactly* the same event code, these event codes equal a specified event code, and
codes and text are linked.
The program first searches an instance of the specified event code. The specified event code may include question marks for 'any code'. If such an event code is found, it is checked whether the next event code is exactly equal (including the codes at the position of question marks) to the next event; if that is the case it is checked whether the event code thereafter also equals the previous event codes; etcetera.

The events with equal event codes are combined. That is, the event texts are taken together. The (equal) event codes are replaced by one event code. The new (combined) event receives the values of the first of the equal successive event codes on the event variables; the values on the event variables of the other event(s) are removed.

To combine events, the sequence should be correctly linked (unless the sequence has no event texts at all).

SETUP

- Select **Sequences > Combine events...** to open the COMBINE EVENTS setup window.
- Enter an event code to specify which event codes should be combined. A question mark means 'any code'.
- You may check ☑ **Check combined events**, to put a check mark before the combined events, so you can identify events that are actually combined. In that case you may also want to check ☑ **Uncheck all events first**.
- Click Combine.

ADDITIONAL INFORMATION

You can specify which events should be combined, by typing an event code after 'combine equal event codes if code'. For example, if you have a coding scheme with four coding variables, an event code like 'ABCD' means that only events with this event code will be combined.

An event code like 'AB??' means that for example two successive events coded as 'ABXY' will be combined. But two successive event codes like 'ABXY' and 'ABPQ' will *not* be combined: the event codes should be exactly equal.

If the event code after 'combine equal event codes if code' consists of only question marks, all successive equally coded events will be combined.

If, for example, two events are combined, the event text of the second event is put after the event text of the first event; a space is added between both pieces of text.

Often the event text is preceded by a character to designate the speaker. This character may be used as the first code of the event code, if ☑ **Use first character of event text as first code** is checked in the **Coding** tab of the FILE SETTINGS.
If:
- this checkbox is checked, and
- the first character of the event text of an event that will be combined equals the first code, and
- the first character of the event text is followed by a colon and a space, and
- these conditions hold for all events that will be combined,

then:
the first two characters (the character for the speaker and the colon) of the second event text are *not* copied. For example:

event code	event text
IPX	I: Sentence one.
IPX	I: Sentence two.

will after the combination look like:

IPX	I: Sentence one. Sentence two.

In most cases this is a useful option. If you want to prevent this nevertheless, be sure to uncheck ☐ **Use first character of event text as first code** first in the **Coding** tab of the FILE SETTINGS. In that case the combined event will look like:

IPX	I: Sentence one. I: Sentence two.

See also:
Select sequences and Select marked sequences.

4.9 Lock sequences

With **Sequences > Lock sequences** you can lock or unlock all sequences at once. You can also lock or unlock the marked sequences only.

See also:
Marking sequences, Locking sequences.

4.10 Mark random

To check a random selection of sequences:

- Select **Sequences > Mark random...** to open the CHECK RANDOM setup window.
- You can enter either a percentage or the number of sequences. If you enter a percentage, the number of sequences is automatically entered and vice versa.
- Check ☑ **Unmark all sequences first** if appropriate (see below).
- Click OK.

The selection of sequences concerns all sequences, irrespective whether they are already marked or not. For example, suppose that about 50 percent of all sequences is already marked. If you mark again 50 percent of the events randomly (keeping ☐ **Unmark all sequences first** unchecked), you can expect that about 75 percent of the sequences will eventually be marked: about half of the randomly selected sequences was already marked, whereas about half of the unmarked events will become marked.

If you check ☑ **Unmark all sequences first**, all sequences will be first unmarked.

4.11 Mark links

The **Sequences > Mark links...** command is used to mark sequences with particular link characteristics between codes and text, respectively sequences and AV files. The command is useful to identify sequences without linked AV files and sequences with bad linked codes and text.

4.12 Copy marks

You can copy or add marks from one set to another set:

- Select **Sequences > Copy marks...** to open the COPY MARKS setup window.
- Select the set you wish to copy with the 'add marks from set' pop-up menu.
- The marks are always copied to the active mark set. You can select a different active mark set with the 'to mark set' pop-up menu. You can also change the active mark set in the main sequence window.
- To restrict the copy to sequences with a particular value, check ☑ **if sequence variable** and select or enter a sequence variable. Enter a value or select 'variable' if you want to compare the value of the selected sequence variable with the value of another sequence variable.
- Click Mark.

Only marks are copied; that is, if a sequence according the 'from' mark set is unmarked, whereas according to the 'to' mark set the sequence is marked, the sequence remains marked according to the 'to' mark set. To obtain an exact copy of the mark set, check ☑ **Unmark all sequences of set [set name] first**.

4.13 Check events

You can check each separate event. Checked events have a check mark in front of the event code. You can manually check or uncheck event codes by clicking before the event code.

The **Sequences** menu allows you to check all events, to uncheck all events, to toggle the checks, or to check a random selection of events.

4.14 Check random

To check a random selection of events:

- Select **Sequences > Check random...** to open the CHECK RANDOM setup window.
- You can enter either a percentage or the number of events. If you enter a percentage, the number of events is automatically entered and vice versa.
- Check ☑ **Uncheck all events first** if appropriate (see below).
- Click OK.

The selection of sequences concerns all events, irrespective whether they are already checked or not. For example, suppose that about 50 percent of all events is already checked. If you check

again 50 percent of the events randomly (keeping ☐ **Uncheck all events first** unchecked), you can expect that about 75 percent of the events will eventually be checked: about half of the randomly selected events was already checked, whereas about half of the unchecked events will become checked.

If you check ☑ **Uncheck all events first**, all events will be first unchecked; also the events from unmarked sequences in case you selected the marked sequences for the analysis.

See also:
Select marked sequences.

4.15 Check data

INTRODUCTION

Check data concerns the checking of codes, sequence variables, event variables, text keys and time keys. Select **Sequences > Check data...** to open the CHECK DATA setup window.

The different types of checks are grouped and can be obtained by the different tabs. Before doing any check, you are strongly recommended to perform a check on the integrity of the data, before performing any of the other checks.

In the upper part of the window you find two general checks, to be discussed in the section on 'General checks'.

To check variables, onset time, offset time and START TIME, text keys, or time keys, select the appropriate type, which will give you a number of different options.

In the lower part of the setup window you can tell whether the check should start with the first sequence, with the present sequence (the one shown in the main sequence window), or with the sequence thereafter. Select the appropriate radio button.

GENERAL CHECKS

Data integrity

Data can become corrupted because of a number of reasons, for example because of a computer crash, or because you copied data to an external device (e.g. a USB stick) without unmounting the device correctly.

You can perform a check on the integrity of the data. If the program encounters an error in the data, you are informed about the error and given the opportunity to correct it. Corrections are also possible for locked sequences.

The program may report to have made 'minor' corrections'. This does not mean that your data contains errors. Examples of such minor changes are removing carriage returns at the end of the sequence text or information fields, or changing a number from '0.2300' to '0.23', or from '0345' to '345'. Such changes usually decrease the size of the data a bit, and make calculations a bit faster.

AV file links

Links between a sequence and AV files maybe broken, in case you moved files to a different location on your hard disk, or in case an external device is not available. For analyses that use the linked AV files, you may first want to check whether the links are correct.

VARIABLES

Check sequence variables

Sequence variables are checked with respect to their (user defined) minimum and maximum values.

Check event variables

Event variables are checked with respect to their (user defined) minimum and maximum values.

Check codes

Codes are checked with respect to the allowed characters a code may consist of. For example, if you have coded your data file and then changed the allowed codes, codes may no longer be in accordance with the allowed codes.

First code ≠ character 1 of text

If ☑ **Use first character of event text as first code** in the **Coding** tab of the FILE SETTINGS window is checked, you can find event codes that do not fulfill this requirement.

TIME

If your events have time information (onset and offset times), you can check:
- if the onset time increases with each event;
- if the STARTTIME of a sequence does not exceed the onset time of the first event (provided the event times are absolute);
- if the difference between the offset time and onset time of an event is positive;
- if the offset time of an event does not exceed the onset time of the next event with the same first code.

Select one or more of these checkboxes, to perform these checks.

TEXT KEYS

Start of text key

The start of a text key should not be lower than 1 (the first word of the sequence text).

Last text key

The end of a text key should not exceed the number of words of the sequence text.

Overlap of same keys

If one or more words are assigned to different text keys, this is called overlap. For example, if text key A is assigned to word 2 to 5, and text key B to word 5 to 7, both keys overlap: both keys share word 5.

Such overlap is not allowed for keys with the same keyword. If a key with key word A is assigned to word 2 to 5, whereas a different key, but also with keyword A, this is not allowed (see also overlap of keys). If such overlap occurs, analyses with keys may give incorrect or unexpected results.

Overlap of different keys

Exclusive keywords

You may have defined your keywords in such a way that particular keywords should not overlap. Suppose for example that you distinguish between words or phrases as being hostile or friendly, and assign keys with these names to parts of the texts. In such a case a particular word quite probably should not be assigned to both the keyword 'hostile' and the keyword 'friendly' at the same time: you view the keywords are (mutually) exclusive.

To check if a set of keywords is exclusive, check **Overlap of different keys** and select 'Exclusive keywords' with the pop-up menu after this checkbox. Next select the keywords that should not overlap using the other pop-up menu. The keywords appear in the field below this latter pop-up menu. To remove a keyword, select the keyword once again with the pop-up menu.

Exhaustive keywords

A particular set of keywords may also be exhaustive. Suppose we have 'friendly', 'hostile' and 'neutral' in case of text keys. Probably you want to assign to each and every word in the sequence text one of these three keywords: these three keywords are exhaustive.

To check the exhaustiveness of a set of keywords, first select 'Exhaustive keywords'. Next select the exhaustive keywords with the pop-up menu in the same way as with exclusive keys.

Inclusive keywords

Finally, a particular keyword may only occur if another keyword occurs too. For example, you may want to assign different keys depending on whether a whether a person is talking, laughing or crying. In addition, if the person is talking, you may want to assign a key depending the particular topic of the talk. In such a case the keyword for topic 'talking about A' (included keyword) can only occur if the person is talking (including keyword).

To check if one keyword includes another one, first select 'Inclusive keywords'. Two pop-up menus with keywords appear. Select the included keyword from the first pop-up menu and the including keyword with the second one.

TIME KEYS

Start of time key

The start of the first time key should not be lower than the lowest onset time of the sequence.

Last text key

The end of a time key should not exceed the highest offset time.

Overlap of same keys

Time keys overlap if they share a particular period of time. Keys with the same keywords should not overlap. For example, if time key A is assigned to time 2 to 7, and time key B to time 5 to 12, both time keys between time 5 and 7 (see also overlap of keys).

Overlap of different keys

Exclusive keywords

Suppose you assign keys with respect to the looking behavior of a respondent in interviewer-respondent interaction. For example, the respondent may look at the interviewer, at showcards, or somewhere else. Because the respondent cannot look at these three objects at one and the same time, these keywords are exclusive.

To check if a set of keywords is exclusive, check **Overlap of different keys** and select 'Exclusive keywords' with the pop-up menu after this checkbox. Next select the keywords that should not overlap using the other pop-up menu. The keywords appear in the field below this latter pop-up menu. To remove a keyword, select the keyword once again with the pop-up menu.

Exhaustive keywords

In the example above about looking behavior, you may want to cover all looking behavior of the respondent with one of the three keys. In such a case these keywords are exhaustive.

Inclusive keywords

Finally, a particular keyword may only occur if another keyword occurs too. For example, if you want to assign a key if both the respondent and the interviewer look at each other, such a key (included keyword) can only occur if the respondent looks at the interviewer (including keyword).

To check if one keyword includes another one, first select 'Inclusive keywords'. Two pop-up menus with keywords will appear. Select the included keyword from the first pop-up menu and the including keyword with the second one.

See for overlap also Coincidence of keys.

5 Text

INTRODUCTION

SEQUENCE VIEWER offers a number of analyses on the sequence texts. Most setup windows for texts have a number of elements in common. These are described below. For some of these options it is important to know that SEQUENCE VIEWER views words as character strings delimited by a space, tab or return. A comma or a period is not viewed as a word delimiter. For example, in "if you,you are", "you,you" is viewed as one word. In "if you, you are", "you," and "you" are two different words: 'you,' and 'you'; the comma is added to the first 'you'.

WILDCARDS

In most setup windows where you can enter text that should be found, you can use the 'insert wildcard' pull down menu, to insert special characters in order to find any character(s). The special character should be preceded by a '^'.

- A '· ' will find any single character (please note that this is not a simple dot, but a raised dot). For example, 'a^· b' will find 'axb', ayb', etcetera.
- A '…' means any number of characters, including no characters. For example 'a^…b' will find 'ab', 'axb', axyb', etcetera.
- A '∂' will find any digit, whereas 'Δ' will find any sequence of digits (that is, any integer). For example, to find a floating point number, you can use '^Δ.^Δ'.

See Text analysis for some more examples.

You can only insert wildcards if the 'resemblance' (see below) equals 1.00.

MATCH OPTIONS

Case sensitive

If ☑ **Case sensitive** is checked, a match is only found if the case of each character in the find text is the same as in the match text. If not checked, differences with respect to case are neglected.

Text style

To find a string with a particular text style, select that text style with the 'Text style' pop-up menu. Text styles can only be applied to sequence texts.

Resemblance

To find a text that is not exactly the same as the text you typed in the find edit box, you can use the resemblance pop-up. The text will be found if the resemblance between the typed text and the found text, exceeds a particular value. The higher the value, the higher the resemblance; 1.00 means an exact match. If you select a value lower than 1, you cannot use the wildcard option to find any characters.

In case the resemblance is lower than 1, you can choose between resemblance based on words, or resemblance based on characters. Resemblance based on words is much faster. However, in case of typing errors (e.g. 'resemblance' and 'resemblence') both words will be viewed as different. But the resemblance of two different words (e.g. 'longer' and 'linger') is high if

the resemblance is based on characters.

Match position

Text is always searched paragraph by paragraph (or event text in case the sequence text is searched).

With the 'Match where' pop-up menu you can specify whether the text should be found anywhere in a paragraph, at the beginning of a paragraph, at the end of a paragraph, or should match a whole paragraph.

Whole words only

Check ☑ **Whole words only** if the first and the last character of the matched text, should be at a word boundary. Word boundaries are the beginning and end of a text, a space, a carriage return, or one of the additional word delimiters. For example, if you want to find 'male', you may check ☑ **Whole words only** to prevent that you also find 'female'.

Please note that if you use ☑ **Whole words only** in combination with ^..., the text found by ^..., will not contain spaces. For example, "w^...s" will not find "word delimiter", but "w^... ^...s" will (note the space between both instances of ^...).

TEXT ADJUSTMENTS

In some cases you may want to exclude particular parts or characteristics of the text from the search. Such text is simply temporarily removed from the search text, or the search text is temporarily adjusted, when searching the text. The following options are available in a number of text setup windows.

Neglect text

Often you may want to add comments to your sequence texts. Usually such comments should not be taken into account for text analyses. You can check one or more appropriate checkboxes to exclude such comments from the analyses. For example, if you want to exclude text between brackets, just check ☑ **[ ]**.

If you checked two or more of these checkboxes, please note that:
The program first looks to '(' and ')', than to '[' and ']', than to '{' and '}', and finally to '<' and '>' (as far as the corresponding checkboxes are checked of course).
Suppose both ☑ **( )** and ☑ **[ ]** are checked. The program first finds the first '(' in the text and than finds the first ')' after this character. All text in between is neglected, including brackets ('[' and ']').
Take a look at the following examples about how text is neglected (the red parts) and the effect on the number of words.

The (number of words of this) text is 4.
The (number of words (of this) text is 4.
The (number of words) of this) text is 6.
The (number) of words (of this) text is 6.
The (number [of) words (of] this) text is 5.
The [number (of] words [of) this] text is 4.
The [number (of] words [of) this) text is 6.
The nu(mb)er of words of this text is 9.
The (number) of words of this text is 8.
The(number)of words of this text is 7.

Diacritical marks

To find a match between the find text and the text that is searched, you may or may not want to take diacritical marks into account. For example, if you want to find 'Muller', you may also want to find 'Müller'. To this end, check ☑ **Remove diacritics** to temporarily remove diacritical marks from the search text (not from the find text).

Neglect characters

If you want to assign for example the length of an event text to an event variable, you may want to neglect punctuation marks or transcription symbols. You can type such characters in the 'Neglect characters' edit box. If you open a text setup window, this box is already filled with a number of common characters you may want to neglect, as far as these characters actually appear in the sequence texts of your file.

Neglecting characters is also useful to prevent that, punctuation marks or symbols are part of a word. For example, if you want to count the number of occurrences of the word 'sequence', the string 'sequence,' (note the comma) is viewed as a different word, and thus not counted. Adding a comma to the neglect characters prevents this.

Neglecting characters may also be important in counting the number of all words. For example, a question mark may be preceded and followed by a space. In such a case, the question mark is counted as a word. Adding such punctuation marks to the neglect characters, prevent that such 'widows' are counted as words.

Successive spaces

If you have entered characters that should be neglected, it my happen that such a character is preceded and followed by a space, for example "be ↑ sure". Neglecting "↑" causes that "be" and "sure" are now separated by two spaces; hence, if your find text is "be sure" (with only one space), "be ↑ sure" will not be found. Check ☑ **Treat successive spaces as one space** to prevent this. If checked, two or more successive spaces are replaced by one space before the search is performed.

Neglect first word

Not all 'words' of an event text are necessarily words in the usual sense. If the event texts consist of, for example interactions between interviewer and respondent, it is usage to precede each utterance by one or more characters to designate the speaker, e.g.:

Interviewer: What is your age?

or

I: What is your age?

In such a case, the first word should not be counted in, for example, calculating a speech rate. To this end, a checkbox is available called ☑ **Neglect first word** to temporarily remove the first word of each event text before the search is performed.

Additional word delimiters

You can add additional word delimiters (e.g. a hyphen, or a slash) in the setup of some text analyses. For example, 'twentieth-century' is viewed as two words, if a hyphen is added to the word delimiters. Additional word delimiters are simply temporarily replaced by a space before the search is performed.

NOTE

Characters that should be neglected should not occur in the 'Find text', because such text can never be found. For similar reasons, if you use for example brackets to neglect text, a '[' or ']' should neither occur in the find text, nor be entered as neglect character. The same is true for additional delimiters.

5.1 Finding

INTRODUCTION

You can find particular pieces of text from the sequence texts or the information fields of sequences. In addition to finding them, you also have options for replacement, (un)marking the sequences, or (un)checking the events wherein these texts occur. Select **Text > Find text...** to open the FIND TEXT setup window. You can only find texts within a paragraph. If you type a carriage return in the 'Find text' edit box, the Find button is activated.

THE SETUP WINDOW

Find text

Type a text in the 'Find text' field. A search always concerns the text within a paragraph. You cannot search for texts across paragraphs. If you leave the 'Find text' field empty, empty fields will be found. Such an 'empty' field however may actually contain text that is neglected (see Text adjustments).
Use the 'insert wildcard' pull down menu, to insert wildcards.
Specify where you want to find the text: in the sequence text, or in one of the five information fields. In case of finding text in the sequence text, you can restrict the search to event texts of checked events.

Search text adjustments

Here you have the options to neglect diacritical marks, neglect characters, and treat successive spaces as one space.

Match options

In addition to the options Case sensitive, Text style, resemblance and Whole words only, you can specify where the text should be found with the 'Match where' pop-up menu. You can find text *anywhere* in a paragraph, *at the start* of a paragraph or *at the end* of a paragraph. If you select the fourth option, 'whole text', the find text should match the whole text of the paragraph.

See also:
Select sequences, Select marked sequences and Checking event codes.

5.1.1 Find text

SETUP

You can find text in either the sequence text field or in one of the five information fields.

- Select **Text > Find text...** to open the FIND TEXT setup window.
- Select the **Find** tab.
- Specify the search options (see Finding).
- Select 'Search all' to search upwards from sequence number 1, 'Search up' to search from the present sequence upwards, or 'Search down' to search from the present sequence downwards.
- Click Find. If the text is found, it is highlighted and the button changes into Find next.

ADDITIONAL INFORMATION

The  icon is disabled if you selected one of the five information fields, because checked events have n meaning for these fields. The icon is also disabled if the 'Find text' field is empty; an empty 'find text' field will find sequences without sequence texts; thus restricting the search to checked events makes no sense.

See also:
Select sequences, Select marked sequences, and Checked events only.

5.1.2 Replace text

You can find text in either the sequence text field or in one of the five information fields.

- Select **Text > Find text...** to open the FIND TEXT setup window.
- Select the **Replace** tab.
- Specify the search options (see Finding).
- Type a text in the 'Replace with' field. If you leave the 'Replace with' field empty, the found text will be deleted.
- Select 'Search all' to search upwards from sequence number 1, 'Search up' to search from the present sequence upwards, or 'Search down' to search from the present sequence downwards.
- Click Find. If the text is found, it is highlighted and the button changes into Find next.
- Click Replace to replace the highlighted text with the 'Replace' text and to find the next instance of the find text. To prevent finding the next instance after the replacement, click Replace with the option key down.
- Or click Replace all to replace the text in all sequences at once.

NOTE

The whole highlighted found text is replaced, including characters that are neglected.

See also:
Select sequences, Select marked sequences, and Checked events only.

5.1.3 Mark/unmark

To mark or unmark sequences having a particular text in the sequence text field or one of the five information fields, do this:

- Select **Text > Find text...** to open the FIND TEXT setup window.
- Select the 'Mark' or the **Unmark** tab.
- Specify the search options (see Finding).
- Check ☑ **Unmark all sequences first**, respectively ☑ **Mark all sequences first** if appropriate.
- Click Mark or Unmark.

See also:
Select sequences, Select marked sequences, and Checked events only.

5.1.4 Check/uncheck

To check or uncheck events having a particular event text, do this:

- Select **Text > Find text...** to open the FIND TEXT setup window.
- Select the 'Check' or the **Uncheck** tab.
- Specify the search options (see Finding).
- Check ☑ **Uncheck all events first**, respectively ☑ **Check all events first** if appropriate.
- Click Check or Uncheck.

See also:
Select sequences, Select marked sequences, and Checked events only.

5.2 Show event texts

You can obtain all texts belonging to particular event codes or to particular keywords. The texts are shown in the output window. If your file is very large, it may take some time before these texts actually appear in the output window.

- Select **Text > Show Event texts...** to open the SHOW EVENT TEXTS setup window.
- Select the **by codes** or **by keys** tab.
- If you selected 'by codes', type the event code for which want to have the event texts. A question mark means any code. You will only get the texts from sequences with correctly linked codes and texts. See Linked codes and text.
- If you selected 'by keys', select a keyword with the pop-up menu. The keyword will be added to the keyword list. To remove a keyword from the list, select it once again.
- Click Show.

NOTES

Selecting Checked events only is not possible if 'by keys' is selected.
Text styles are not shown in the output window.

See also:
Select sequences, Select marked sequences, and Checked events only.

5.3 Word frequencies

With 'Word frequencies' you obtain a frequency table of all words of the sequence texts. The table is shown in the output window.

- Select **Text > Word frequencies...** to open the WORD FREQUENCIES setup window.
- Check or uncheck the checkboxes for Case sensitive and Remove diacritics, and enter appropriate Neglect characters and Additional delimiters.
- To prevent that the table becomes very long, you can set a minimum frequency for the occurrence of a word.
- Click Count.

See also:
Select sequences, Select marked sequences, and Checked events only.

5.4 Text count

You can assign the total number of words or characters, or the number of occurrences of particular words or characters of each event text to an event variable.

- Select **Text > Text count...** to open the TEXT COUNT setup window.
- Select the name of an event variable from the variables field. The count will be assigned to this variable (provided text and codes are linked, see Linked codes and text).
- Choose appropriate settings for Neglect text and Neglect first word, and enter Neglect characters, if necessary.
- Click the **Words** or the **Characters** tab. 'Words' counts the number of words, whereas 'Characters' counts the number of characters.

If you selected 'Words':

- You can enter other word delimiters in addition to space, tab and return.
- Select 'Count all' with the 'Count what' pop-up menu if you wish to count all words.
- To count only particular words, select 'Only count:', and to count all words except particular words, select 'Don't count:'. Enter these words, separated by spaces in the edit box below the pop-up menu. In addition you can check the Remove diacritics and Case sensitive checkboxes.
- Click Count.

If you selected 'Characters':

To count all characters, select 'Count all with the 'Count what' pop-up menu. Select 'Only count' to count only particular characters, and select 'Don't count to count all characters except particular characters. Type these characters in the field below the pop-up menu.
Click Count.

See also:
Text analysis, Select sequences, Select marked sequences, and Checked events only.

5.5 Text analysis

INTRODUCTION

You can assign the number of times particular words or strings occur in the event text to an event variable. The TEXT ANALYSIS setup window cannot be opened if no event variables are available.

In some cases it may be preferred to use Text count to count the number of occurrences of particular strings. 'Text count' is more straightforward than 'Text analysis'; moreover 'Text count' allows you to count the number of words, excluding particular words, and/or excluding the first word.

Generally, if the text to be counted consists of two or more words, you should use 'Text analysis'. If the text to be counted consists of only one word, you may use the TEXT COUNT setup window with the 'Count only' option instead.

For example, suppose you want to count the number of times 'behavior' occurs. May be also the UK spelling of this word ('behaviour') occurs in the event texts. This can be solved by entering 'behavior behaviour' in the count field of the TEXT COUNT setup window. The same is true if there may be typing errors in the event texts, e.g. 'resemblence' instead of 'resemblance'.

Such problems can also be solved by using wildcards in the TEXT ANALYSIS setup window. For example, you can enter 'behavi …r' (or 'behavio …r') as count string. Or you can use the 'synonyms' count type in 'Text analysis' (see below).

It is recommended to inspect a Word frequencies table before performing a 'Text count' or 'Text analysis'. If there are spelling errors in the words you want to count, or different spellings of the same word, you can correct them first, or decide upon the most appropriate analysis.

As another example to illustrate that most analyses in 'Text count' can also be performed in 'Text analysis', but in a less straightforward way, suppose we simply want to count the number of words of each event. This is very simple of course in 'Text count'. In 'Text analysis' we can obtain the same result, if we just enter the wildcard ' …' (any characters) in the 'count string' field, and check ☑ **Whole words only**.

SETUP

- Select **Text > Text analysis...** to open the TEXT ANALYSIS setup window.
- Type the string you wish to count in the text field below *Count string* and use wildcards if appropriate.
- Select a count type with the pop-up menu (see below).
- Specify the Match and Text adjustment options.
- If you check ☑ **Whole words only**, you may want to add additional word delimiters. For example, suppose you want to count the number of times the word 'male' appears in the event text. If you don't check ☐ **Whole words only** also 'female' is counted. But if you check ☑ **Whole words only** 'male/female' will not be counted. If you add '/' as additional delimiter, 'male/female' is counted as one instance of 'male'.
- Select the name of an event variable from the variables field. The number of words will be assigned to this variable (provided text and codes are linked, see Linked code and text).

COUNT TYPES

It is assumed that ☐ **Whole words only** and ☐ **Case sensitive** are unchecked in the examples below. Found instances are in red.

Literal

Literal counts the number of times the string occurs exactly in the event text.

- Example 1: the string "the fi" yields a count of 1 with the event text "In the field below".
- Example 2: the string "the" yields a count of 2 with "The meaning of these words".

Synonyms

Synonyms counts the number of times one of the words in the count string occurs in the event text.

- Example 1: the count string "more each" yields a count of 1 with the event text "Type one or more words".
- Example 2: the string "more each" yields a count of 2 with the event text "Type one or more words on each line".

This option is especially useful if, for example, you wish to count the number of times a word like 'behavior' occurs. If you type "behavior behaviour" in the 'Count string' field, you include both spellings.

Unordered

Unordered counts the number of times each and every word in the string occurs in the event text, irrespective of the order of the words in the count string. The difference with 'synonyms' is that 1 is added to the count, only if *each* word of the count string occurs in the event text.

- Example 1: The string "more each" yields a count of 0 with "Type one or more words".
- Example 2: The string "more each" yields a count of 1 with the event text "Type one or more words on each line".
- Example 3: The string "each more" yields a count of 1 with the event text "Type one or more words on each line".
- Example 4: The string "each more" yields a count of 2 with the event text "Type one or more words on each line or on each line more and more words".

Ordered

Ordered counts the number of times *all* words in the count string occur in the same order in the event text.

- Example 1: The string "more each" yields a count of 0 with "Type one or more words".
- Example 2: The string "more each" yields a count of 1 with the event text "Type one or more words on each line".
- Example 3: The string "each more" yields a count of 0 with the event text "Type one or more words on each line".

Characters

Characters counts the number of times each character in the count string occurs in the event text, irrespective of order.

- Example: The string "ao" yields a count of 5 with the event text "How are you today?"

NOTES

Especially if you use the ' ·' and ' …' symbols in combination with ☑ **Whole words only** checked, results may be sometimes difficult to predict.
Here are some hints, that may be useful for particular situations.

- ' ·' with ☐ **Whole words only** unchecked will find each and every single character in the event text and thus returns the number of characters of the event text (the length of the event text).
- ' ·' with ☑ **Whole words only** checked will find words of one character in the event text and returns the number of single character words in the event text.
- ' · · ·' with ☑ **Whole words only** checked will find words of three character in the event text and returns the number of three character words in the event text.
- ' …' with ☐ **Whole words only** unchecked will find the whole event text and will return 1.
- ' …' with ☑ **Whole words only** checked will find each and every single word in the event text and will return the number of words of the event text.
- ' … …' has the same meaning as ' …'.

You can use **Text > Find Text...** to see if what you get is what you want (to compare results, in the FIND TEXT setup window, the text style should be 'Any', the resemblance should be 1.00 and 'Match where' should be 'anywhere').

See also:
Select sequences, Select marked sequences, and Checked events only.

5.6 Speech rate

You can assign the speech rate of each event to an event variable. The speech rate is the number of words per second.

- Select **Text > Speech rate...** to open the 'Calculate SPEECH RATE setup window.
- Select the name of an event variable from the variables field. The speech rate will be assigned to this variable (provided text and codes are linked, see Linked codes and text).
- Choose appropriate settings for Neglect text and Neglect first word, and enter Neglect characters and Additional word delimiters, if necessary.
- Select the minimum number of words of the event text. For example, an event text that consists of only one word (e.g. 'Yes'), will not yield a reliable speech rate.
- Click Calculate.

See also:
Select sequences, Select marked sequences, and Checked events only.

5.7 Repeated words

INTRODUCTION

Repeated words concern the number of times the same string is repeated in an event. A common usage of this analysis is to discover 'restarts', that is if a speaker repeats an earlier part of the utterance. For example: "And how many hours uh how many hours did you work last week?" Such a restart can be viewed as a speech disturbance and should be distinguished from non-disturbing repeats. For example: "How concerned would you say you are about this issue?". The word "you" appears twice in this sentence but this should not be viewed as a restart.

The program cannot distinguish restarts from non-disturbing repeats, so you should decide yourself whether or not a repetition is a restart. The program can aid you in this process by finding and showing you the repetitions in an event, leaving it to you to decide upon what to do with a repetition.

Please note that the program does not count the repetitions themselves (that is, does not count the number of times a particular string is repeated), but the number of *sets* of repetitions. For example, the utterance 'No no, we have uh we have two cars.' has two sets of repetitions: 'No no' (assuming the match is not case sensitive) and 'we have uh we have' (assuming that repeated strings may be separated by one or more words - see below).

SETUP

- Select **Text > Repeated strings...** to open the FIND REPEATED STRING setup window.
- Optionally you can select the name of an event variable from the variables field. The number of sets of repetitions is assigned to this variable.
- Choose appropriate settings for Match options and Text adjustments.
- Checking ☑ **Whole words only** means that the whole word should match. For example, "I thi think" is considered a repetition, provided that this checkbox is *not* checked, because "thi" is repeated. If checked "I thi think" is not a repetition, because "thi" does not match "think".

In addition to these common options, a number of other settings are available.

- If ☑ **Repetition must be in same sentence** is checked, the event is split up in sentences, and the repetition should be within the sentence. A sentence is ends with '. ', '? ' or '! ' (note the space after these punctuation marks), a carriage return, or the end of the sequence text. A sentence starts at the beginning of an event text or after '. ', '? ' or '! ' (including the space after the punctuation mark).
- Repetitions may be separated 0 or more words. For example, the repetition of "no" in "no no" is separated by 0 words; the repetition of "no" in "no uhmm no" is separated by one word. The maximum number of words between repetitions can be set with the pull down menu 'Repeated strings may be separated by maximally [x] words'. If the number of words between a repetition exceeds this setting, it is not viewed as a repetition.
- A repetition may consist of one or more words. For example, the repetition in "no uhmm no" consists of one word ("no"). The repetition in "I have well I have" consists of two words ("I have"). The maximum number of words of a repetition can be set with the pull down menu 'Repetitions may consist of maximally [x] words'. If the number of words of a repetition exceeds this setting, it is not viewed as a repetition.
- You can (temporarily) delete particular words that may appear between a word and its repetition. For example, if you enter "uhmm' as a word that should be neglected, the number of words between the repetition of "no" in "no uhmm no" is 0.

- Click Find.
- Or click Assign if you selected an event variable. We recommend however to first use the Find button, and to observe if the results are as expected.

SHOW REPEATED STRINGS

If you click Find, and the program finds one or more repetitions according to the options described above, the event text will be shown in a separate window. The repetitions of a set are colored. If there are more sets of repetitions, each set will have a different color.

Up to nine different sets can be shown at a time. You can hide or show each set by clicking on the corresponding colored button. This is particularly useful if there are a number of sets, which partly overlap.

If you have selected an event variable, you can assign the number of sets of repetitions to the selected event variable (if you have selected one) by clicking Assign count to ['name'], where ['name'] is the name of the selected event variable. The number of *shown* sets is given after 'count =='. That means that you can hide inappropriate sets, and thus adjust the number of sets of repetitions that is assigned to the event variable. If there are more sets of repetitions, you can select an appropriate value with the pop-up menu after count =.

Click Next to find the next instance of repetitions in an event.

EXAMPLES

Example 1

Consider the sentence:

"Okay, well we wer- on your husband's question we were asking the average week."

with these settings:
- Nothing is neglected.
- ☐ **Whole words only** is unchecked.
- Repeated strings are separated by 2 words or less.
- Repetitions consist of 1 word.

(The other options are not relevant for the example).

This yields:

"Okay, well we wer- on your husband's question we were asking the average week."

There are two sets of repetitions, the red one and the blue one. You may click one of the shown buttons to observe the effect.
If we checked ☑ **Whole words only**, no repetitions would have been found.

We leave ☐ **Whole words only** unchecked, but specify that repeated strings may be separated by maximally 3 words instead of 2. We obtain:

"Okay, well we wer- on your husband's question we were asking the average week."

"Week" is added to the blue set, because "we were" and "week" are separated by 3 words.

Now we set the number of words between repetitions to 4. This is the result:

"Okay, well we wer- on your husband's question we were asking the average week."

There is only one set of repetitions, with "we" repeated 5 times.

Setting the number of words between repetitions to 5, yields the same result. But if we in addition

check ☑ **Whole words only** you obtain:

"Okay, well we wer- on your husband's question we were asking the average week."

Example 2

For our second example, we change the sentence a bit:

"Okay, well when we asked you- when we asked the average week."

with these settings:
- Nothing is neglected.
- ☐ **Whole words only** is unchecked.
- Repeated strings are separated by 3 words or less.
- Repetitions consist of maximally 2 words.
(The other options are not relevant for the example).

We obtain three sets of repetitions:

"Okay, well when we asked you- when we asked the average week."

You may also observe that both words "we" are flagged. If a word is flagged, it means that it belongs to two (or more) repetition sets. Because a word cannot have two colors at the same time, we can use the shown buttons, to hide a particular set. In this case we click at the green shown button to hide the green set. Now both words "we" become blue, but both words remain flagged. If we hide the blue set too, it appears that "we" also belongs to the red set.

Apparently we have a 'green' set, consisting of "we", "we", and "week", a blue set consisting of "when we" and "when we", and a red set consisting of "we asked" and "we asked". This may seem a bit superfluous, but the program cannot know what exactly you are after, and just shows all possibilities. In the present example, you probably would decide that there is just one set of repetitions, consisting of "when we asked" and "when we asked". This set is not shown, because we specified that repetitions should consist of no more than two words. If we change this setting to 3, this set is shown, together with the other three sets.

It will be clear from these examples that the options you select may greatly affect the results. For this reason it is not recommended in most cases to automatically assign the number of sets to an event variable. Rather you should use the program to find possible relevant sets, and use the shown buttons to keep the sets you want. This may be more tedious than automatically assigning the number of sets, but it is much less tedious than finding speech disturbances by reading all texts, and moreover much more reliable.

See also:
Select sequences, Select marked sequences, and Checked events only.

5.8 Auto assign keys

SETUP

Automatic assignment of text keys is intended to find a particular piece of text in the event text and assign a text keyword to this piece of text.

- Select **Text > Auto assign keys...** to open the AUTO ASSIGN KEYS setup window.
- Enter the find text and use wildcards if appropriate.
- Specify the Match and Text adjustment options.
- Select whether you want to assign the key to the found text, or to extend the text with the 'Assign key to' popup menu (see below).
- Type the keyword you want to assign after 'Assign keyword'. That can be a new keyword that does not exist yet, or an already existing keyword. You can also select an existing keyword with the pull down menu.
- Click Find to find the text. The button will change into Find next, and the found text becomes highlighted. The TEXT KEYS window is opened, if not open yet.
- Click Assign to assign the keyword to the selected text. The program will proceed to find the next instance of the find text.
- If you don't want the program to find the next instance of the find text after the key is assigned, hold the option key down when clicking Assign. You may observe that adding the key is reflected in the TEXT KEYS window.
- Click Assign all to assign the keyword to all pieces of texts that fulfill the find conditions.

ADDITIONAL INFORMATION

If a text is found, the selected text will always be at word boundaries, because keys are assigned to whole words, not to characters. For example if you search for 'our', this piece of text may be found in 'source'. Hence the whole word 'source' is selected, and the keyword is assigned to the whole word.

This behavior should not be confused with checking ☑ **Whole words only**. If this checkbox is checked, 'our' will not be found in 'source', because 'our' does not start and end with a word boundary (space, carriage return, and the very start and end of the sequence text).

You can assign the key exactly to the found words, or to an extended text. One option is to assign the key to the whole event text wherein the text is found. Another option is to use particular punctuation marks that serve as boundaries for the text to which the key should be assigned. Suppose you want to assign the key to the sentence wherein the find text is found. In that case, the 'left' boundary is a period *plus a space*, that is nearest to the found text, and you should type a period in the field after 'between'. The 'right' boundary is the first period *plus a space* after the found text. You can use period, question mark, exclamation mark, comma, colon and semi-colon as boundary characters. You can use only one, or more characters for a boundary. The characters for the 'left' boundary need not be the same as the characters for the 'right' boundary. If no boundary characters are found, the first, respectively the last word of the event text serve as boundaries. If a character is not followed by a space, it cannot be a boundary. This prevents for example that the period in a floating point number like '14.352' is viewed as boundary.

See also:
Select sequences, Select marked sequences, and Checked events only.

5.9 Auto code

Automatic coding is intended to find a particular event text and to assign a code to the linked event code.
You can restrict the search to events that already have a particular code, or, alternatively, do not have a particular code. There are also a number of options available with respect to the text to be found.

- Select **Text > Auto code...** to open the AUTO CODE setup window.
- Enter the find text and use wildcards if appropriate. In addition you can check ☑ **Does not match**. If checked, the event code is assigned to all event texts, that do not match the find text.
- Specify the Match and Text adjustment options.
- Type an event code after 'Code event as'. This is the code that will be assigned. A question mark means that the code on that position remains unchanged.
- Check ☑ **only if code** if you want to apply an additional condition. Click at '=' to change it to '≠' and vice versa, and enter a conditional event code.
  - '=' finds events with an event code that equals the conditional event code.
  - '≠' finds events with an event code that does not equal the conditional event code.
  - A question mark means any code. If the conditional event code consists of question marks only, a '=' means that every event code fulfills the condition (which makes the condition redundant), whereas a '≠' means that no event code fulfills the condition (which doesn't make sense).
- Click Find to find the text. The button will change into Find next, the found text becomes flagged, whereas the corresponding event code becomes highlighted.
- Click Replace to replace the highlighted event code with the new event code and to find the next instance of the find text. If you hold the option key down when clicking Replace, the event code is replaced, but the program does not continue to find the next instance of the find text.
- Click Replace all if you want to change all found events at once, without confirmation.

See also:
Select sequences, Select marked sequences, and Checked events only.

6 Keys

INTRODUCTION

You can assign a key to a piece of text of the sequence texts (text keys), or to parts of linked audio or video files (time keys).

Here we will describe a number of common features to both the text keys window and the time keys window.

The keys window

The key field

The large field in the keys window is the key field. In the key field, keys are represented by colored bars. The start, end and length of a colored bar, correspond to the start, end and length of the text key (piece of sequence text) or time key (e.g. an episode of a movie). A key is characterized by a keyword; in the keys window, keys with the same keyword all have the same color. Moreover, keys with the same keyword are horizontally aligned.

The key field slider

The key field slider at the top of the window, corresponds to a particular word of the sequence text in case of text keys, or to a particular time point in case of time keys. If you move this slider, you may observe that, in case of text keys, the corresponding word in the sequence text is highlighted.

In case of time keys, the corresponding frame of the linked AV file is shown (if selected). In addition, if text and codes are linked, and onset and offset times are available, the word that *approximately* (as estimated by the number of words of the event text, and the onset and offset times of the event) correspond to the position of the slider, is highlighted.

Zooming in on keys

The vertical slider to the left of the key field, and the horizontal slider below the key field (between the orange rectangles), can be used to zoom in on a particular part of the key field. With both scrollbars attached to the key field you can scroll through the key field.

Selecting a key

To select a key, just click on the corresponding colored bar. The bar becomes surrounded by a black rectangle and the color is changed (lighter) to show that it is selected. To deselect a key, just click at some white space between keys.

Create an image of the keys field

You can save the image with the keys as a PNG picture to your hard drive. If you click the icon  with the hard drive in the lower left corner, you are shown how the picture will look like before you save it.

Showing keys in the sequence text field of the main sequence window

You can make keys visible in the sequence text field of the main sequence window. Select 'text keys' or 'time keys' with the 'Show' pop-up menu in the lower right corner of the main sequence window, and select a particular keyword with the 'Select keyword' pop-up menu. The texts corresponding to this key become colored. In case of time keys, the texts that belong to a key are

inferred from the onset and offset times (if available); hence they are not exact.

Editing keys

You can perform a number of edit operations on keys. For some edit operations, it may be necessary to move the key field slider over the key to be edited. For other edit operations, it may be necessary to select a key.

Add a new key

- Move the key field slider to the position (a word, or a frame of a movie file) were you want the key to *start*.
- Click the New button in the upper left corner of the keys window. A drawer appears. Enter the name (the keyword) of the new key, or select an existing keyword. A keyword may not have spaces and cannot be longer than 12 characters.
- Move the key field slider to the position where the key should end.
- Click Add key in the drawer.

You can add as many keys to a sequence as you want. Neither is there any restriction to the number of different keywords.

Delete a key

Select the key you want to delete, and click Delete.

Split a key

Select the key you want to split, move the key field slider to the point where you want the split, and click Split.

Combine two keys

You can combine two keys with the same keyword, if there is no gap between both keys. Select either one of both keys, and click Combine.

Rename a key

Select a key and click the Rename button. A drawer appears. Type or select a new name for the key, and click the Rename button of the drawer.

Moving keys

Select the key you want to move.

- Click the  icon to move the key one unit to the left.
- Click the icon to move the key one unit to the right.

To move a key over a large distance:

- Select the key you want to move.
- Position the key field slider somewhere between the start point and the end point of the key.
- Check the  **Move** checkbox.
- Move the key field slider to the left or the right with the shift key down.

Resizing keys

Select the key you want to resize.

- Click the  icon to extend the key with one unit at the left.

- Click the icon to shrink the key with one unit at the left.
- Click the icon to shrink the key with one unit at the right.
- Click the icon to extend the key with one unit at the right.

To resize a key with a larger amount:

- Select the key you want to resize.
- Position the key field slider either at the left or the right border of the key.
- Check the 'Resize' checkbox.
- Move the key field slider to the left or the right with the shift key down.

The 'Resize' checkbox is only enabled if the slider is exactly at the right or left border of the key. It may be necessary to use the horizontal zoom slider, to zoom in at the key border.

Undo

The 'Undo' pull down menu in the lower right corner, gives you some options for undoing changes made to the keys. Hold the mouse down on this button, and select an appropriate 'undo' command.

The keys window drawer

You can click at the grey triangle **more** in the lower left corner of the keys window, to extend the window with a drawer. Here you find some information about which part of the sequence is displayed in the keys window.

The right part of the drawer provides you with an overview of which colors belong to the different keywords. Only those keywords that actually occur in the sequence are shown with their colors.

Notes

In the notes field you can enter any information about a particular key. To enter information, select a key first.

In the main sequence window, you can check **show notes** in the keys box (in the lower right corner). If **show notes** is checked, a yellow sticky appears. The sticky shows the notes for a key, if you move the mouse over the words of the sequence text that belong to the key. In case of time keys, the text that belong to a key are inferred from the onset and offset times (if available).

Duration and overlap of keys

If a text key is assigned to words 3 to 5, the key covers three words (3, 4 and 5). Hence, the 'length' of this key is three. If you assign another key to words 5 and 6, both keys overlap: they have word 5 in common.

If a time key is assigned to time 3 to 5, the duration of the key is 2 (5 minus 3). If another time key is assigned to time 5 to 6, both keys do *not* overlap: both keys don't have a period of time in common.

A text key that ranges from word 4 to word 4, has a length of 1 (word), which is perfectly possible. A time key from time 4 to 4 has a length of zero however, which doesn't make sense (use **Sequences > Check data** to check for such zero lengths.

Different keys with the same keyword should not overlap. For example, suppose you assign (text) keyword 'A' both to word 3 to 7, and to word 6 to 10. In the keys window you cannot distinguish between both keys: both bars have the same color. Hence, both keys look like one key, of length 8, whereas actually there are two keys, both with a length of 5. As a consequence, calculations may be unreliable or yield unexpected results because of such overlap (see also Coincidence of keys). To prevent such overlap you should check your data (see Check data).

See also:
Time unit.

6.1 Define keys

'Define keys' allows you to:
- Change the name of a keyword
- Change the color of keywords
- Add a label to a keyword
- Change the order of keywords

To change one of these properties, select **Keys > Define keys...** to open the DEFINE KEYS setup window. Than click either the **Text keys** or the **Time keys** tab.

You should first select a keyword: click at the name or the color of a keyword to select it. The color itself will be enlarged, to indicate which keyword is selected. You can deselect a keyword by clicking it once again.

Change the name of a keyword

Select a keyword, and type a new name. You should click Change name, to accept the change (SEQUENCE VIEWER first checks if the name is allowed, or already exists).

Change the color of a keyword

Select a keyword and change the color by moving the sliders to adjust hue, saturation and brightness. It is recommended that you don't use dark colors, because in same graphs, textual information (in black) appears in boxes that have the color of the keyword.

You can get the original color back, by clicking at the 'original color' field. Key colors are stored within your file if you save your file.

Add or change a keyword label

You can add a label to each keyword, for example a definition of the keyword. Just type the label in the keyword label field.

Change the order of keywords

The order of keywords (as they appear in the 'Text keys' or 'Time keys windows'), can be changed with the  and  buttons. Click these buttons, to move the selected keyword upwards or downwards.

6.2 Text keys

You can assign a key to a piece of text of the event texts. A piece of text consists of one or more successive words.

To assign a text key, you can use the key field slider in the upper part of the TEXT KEY window (see Keys). You can also assign a piece of text to a key by selecting a piece of text in the text field of the main sequence window:

- Click the New button and enter a name for the new key in the drawer that appears.
- Select the sequence text to which you want to assign the key.
- Click the Add key button.

You can use the and buttons to the left of the key field slider to move one word backwards of forwards.

For a very simple method to assign a piece of text to an existing keyword (that is, its name should already exist), see the Toolbar.

NOTES

If you delete or add one or more words from the sequence text, the values (start and end) of a key may change. For example, if a key points to word 17 to 20, and you delete word 18, the key will now point to word 17 to 19. Keys are automatically adjusted for such changes.

Also if you replace text using the FIND TEXT setup window, adjustments to the start and end points of text keys, are automatically made.

6.3 Time keys

You can assign a time key to a part of linked audio or video files. To use time keys, your file needs time information, that is onset and offset times of events (see Add event time)

To assign a time key, you can use the key field slider in the upper part of the TIME KEYS window (see Keys). You can also assign an audio fragment, or an episode from a movie to a key, using the precision slider in the Player window. Usually the best procedure is as follows:

- Move the key field slider of the TIME KEYS window roughly to the start of the audio fragment or video episode were you want the key to start.
- Use the precision slider of the Player window to move the playhead exactly to where the key should *start*.
- Click the New button and enter a name for the new key in the drawer that appears.
- Use the precision slider of the Player window again to move the playhead exactly to where the key should *end*. It may be necessary to use the precision slider a number of times in succession in case of lengthy keys.
- Click the Add key button.

In addition you can use the and buttons to the left of the key field slider of the TIME KEYS window to move one time unit backwards of forwards.

If **Play selected key** (at the bottom of the TIME KEYS window) is checked, then if you click at a colored key bar, the corresponding sound of the key will be played.

6.4 Find keys

SETUP

To find a particular keyword, do this:

- Select **Keys > Find keys...** to open the FIND KEYS setup window.
- Select the **Find text keys** or the **Find time keys** radio button. If no keys of a particular type are available in your file, that radio button is disabled.
- Click the **Find** tab.
- Select a keyword from the 'Find keyword' pop-up menu, or enter an existing keyword manually.
- Click Find. To find the next occurrence of the key, click Find next.

If a key is found, the TEXT KEYS or TIME KEYS window opens. The key that is found, is selected. If you zoomed in on a particular part of the keys field, the found key may be outside the visible part. In that case a message appears just above the blue separation line in the FIND KEYS setup window. You should zoom out, or scroll the keyfield to see the selected key.

In addition the corresponding text in the sequence text field of the main sequence window is flagged. In case of text keys, the flagged text matches the key exactly. In case of time keys, the match is only approximate, as it is inferred from the number of words of an event, its onset and offset times, and the start and end of the time key.

You can also use the **Mark** and **Unmark** tabs, to mark or unmark sequences having one or more instances of a particular keyword.

In all these cases (Find, Mark and Unmark), you can use the 'Checked events only'  icon. In that case, keys will only be found if the key, or *part* of the key is in a checked event. Suppose keyword 'KeyA' covers the last two words of event 3 and the first word of event 4. If event 3 is checked, and event 4 is not checked, event 3 will be found, whereas event 4 will not be found.

The **Check** and **Uncheck** tabs enable you to (un)check events that (partly) overlap with the key. For example, if a found key belongs to two successive events (e.g. the last two words of an event text and the first two words of the next event text, *both* events will be checked, if the **Check** option is used.

See also:
Select marked sequences, and Checked events only.

6.5 Delete keys

To delete keys of a particular keyword, do this:

- Select **Keys > Delete keys...** to open the DELETE KEYS setup window.
- Select either **Delete text keys** or **Delete time keys**. If no keys of a particular type are available, that radio button is disabled.
- Enter or select a keyword after 'Delete keys of keyword'.
- Click Delete.

If all keys of a particular keyword are deleted (e.g. all sequences are selected, no sequences are locked), also the keyword itself (including its label and color) is removed from the file.

See also:
Select sequences and Select marked sequences.

6.6 Split event text

If you have text keys assigned to your sequence texts, you can split the event texts in such a way that event texts coincide with the start or the end of a keyword.

- Select **Keys > Split event texts...** to open the SPLIT EVENT TEXT BY KEYS setup window.
- If you check ☑ **Split at start of keyword**, the event text begins with the first word of the key. If you check ☑ **Split at end of keyword**, the event text ends with the last word of the key. You can also check both checkboxes.
- You can either use every keyword to split the event texts, or select only one keyword.
- You can either apply the split procedure to sequences without any codes, or to all (selected) sequences.
- You have the option to mark sequences with events that are split if you want to inspect the sequences that are changed.
- Click Split.

NOTES

If the sequence has codes, and codes and text are (correctly) linked, an event text may be split into two or more new event texts. Each new event text receives the event code of the originally linked event text. The same is true for event variables, if available. Onset and offset times are assigned a 'M' for missing at the position of the split.

See also:
Select sequences and Select marked sequences.

6.7 Convert keys

You can assign properties of text or time keys, to sequence variables or event variables; or replace event codes depending on whether a particular event occurs during the event.

- Select **Keys > Convert keys...** to open the CONVERT KEYS setup window.
- Select ◉ **Convert text keys** or ◉ **Convert time keys**. If you don't have keys of a particular kind, that radio button is disabled.
- Enter or select a keyword.
- Select either 'sequence variable', 'event variable' or 'event code' with the 'Assign to' pop-up menu.
- In case of sequence and event variables, enter a variable name, and select one of the three radio buttons in the 'assign' box, depending on the key property you want to assign to the sequence or event variable; see below.
- In case of event variable and time keys, you have to decide to which event a pause belongs (see below).
- In case of event code, you can replace a particular event code by another event code, if the key occurs during the event. Type an event code that should be replaced after 'If event code is' and the new event code after 'Replace with'.
- Click Convert.

Sequence variable

If you selected to assign text keys properties to a sequence variable, you can assign either one of these properties:
The number of times the selected keyword occurs in the sequence text.
The total number of words of all instances of the selected keyword in the sequence text.
The average number of words per keyword is in the sequence text.
In case of time keys, you have the choice between:
The number of occurrences of the keyword during the sequence.
The total time the selected keyword takes in a sequence.
The average time of the keyword.

Event variable

The options for event variables are similar to those of sequence variables, but here of course it is not the whole sequence (text) that counts, but the event (text).

In case of time keys, you have to decide to which event a pause between two events belongs. You can view a pause following the event as belonging to the event (that is, between its offset time and the onset time of the next event: 'event plus pause), or the pause before an event ('pause plus event'), or exclude the pause.

Code variable

If you selected 'code variable', the program first determines whether or not the keyword occurs during the event. If yes, a particular event code can be assigned to the event, given that the event code is coded in a particular way (that is, a particular event code is replaced by another event code).

Hence, the event code is replaced if (1) the keyword occurs somewhere during the event, and (2) the event code matches the code entered after 'If event code is:' (a question mark means 'any code'). A question mark in the event code after 'Replace with' means that the original code at that position is retained.

Like with event variables, you have to decide to which event a pause between two events belongs in case of time keys.

See also:
Select sequences and Select marked sequences.

6.8 Keyword length

You can obtain some statistics about the length of a particular keyword with the KEYS LENGTH setup window. In case of time keys, the length is the duration of the key (the difference between the start and the end of the key); in case of text keys, the length is the number of words of the key.

The program can also produce the information about the length of a keyword for each value of a sequence variable.

- Select **Keys > Keyword length...** to open the KEYWORD LENGTH setup window.
- Choose either **Text keys** or **Time keys**.
- Select a keyword with the combo box, or enter the name of a keyword.
- To obtain statistics for the different values of a sequence variable, check **By sequence variable**, and select or type the name of the sequence variable in the usual way.

- You should check either ☑ **Average**, ☑ **Frequencies**, or both. If you check ☑ **Average**, you obtain a number of descriptive statistics concerning length of the keyword. You can also obtain a frequency distribution of all different lengths of a particular keyword with ☑ **Frequencies**.
- Click OK.

See also:
Select sequences, Select marked sequences and Table options.

6.9 Keywords table

To obtain a frequency distribution of all text keywords or time keywords, do this:

- Select **Keys > Keywords table...** to open the TABLE OF KEYWORDS setup window.
- Select ◉ **Text keys** or ◉ **Time keys**. If you don't have keys of a particular kind, that radio button is disabled.
- Click OK.

See also:
Select sequences and Select marked sequences.

6.10 Coincidence of keys

INTRODUCTION

Keys may (partly) overlap. For example, in case of text keys, you may have assigned keyword A to a particular sentence, and keyword B to some words of that sentence. To find out how much percent of the time key A coincides with key B, you can use this analysis.

SETUP

- Select **Keys > Coincidence of keys...** to open the COINCIDENCE OF KEYWORDS setup window.
- Select ◉ **Text keywords** or ◉ **Time keywords**. If you don't have keys of a particular kind, that radio button is disabled.
- Select one of the options with the 'Show' pop-up menu (see below).
- Click OK.

INTERPRETING THE RESULTS

The analysis yields a matrix, showing the amount of overlap in the cells. The cells in the matrix show either percentages or frequencies (in case of text keys), respectively durations (in case of time keys), depending on the settings of the table options. The last column of the matrix shows the total number of words (in case of text keys) or the total duration (in case of time keys). Thus these totals are not the totals of the cells in the corresponding rows. In addition you can choose to calculate phi for each cell, as a measure of association between two different keywords.

Percentages are calculated as a percentage of the totals in the total column. Percentages in a row do not necessarily add up to hundred. For example, key 'blue' may overlap 50% with key 'purple', and key 'orange' may overlap 70% with key 'purple', whereas keys 'blue' and 'orange'

overlap with 20 %. See the picture below.

Overlap of time keys is somewhat different from overlap of text keys. If text key A concerns word 2 to 5, and text key B concerns word 5 to 7, it is clear that both keys overlap (with one word: number 5). This is not the case with time keys however. If time key A ends at time 5, and time key B starts at time 5, they do not overlap. The analysis takes this into account, and calculates the overlap for text keys and time keys differently.

NOTE

Calculations may be incorrect if *different* keys with the *same* keyword themselves overlap. In the picture below, the smaller orange key (the selected one) completely overlaps the larger orange key.

In calculating the duration of the orange key *both* orange keys are taken into account, thus underestimating the percentage of overlap with the blue key. To prevent this, you should first check your data: select **Sequences > Check data...**, choose the **time keys** and/or the **text keys** tab, and check ☑ **Keys with the same keyword should not overlap**.

See also:
Select sequences and Select marked sequences.

6.11 Coincidence graph

SETUP

A coincidence graph shows the period of time two keys co-occur (time keys) or the number of words two keys have in common (text keys). To get a graph, do this:

- Select **Keys > Coincidence graph...** to open the COINCIDENCE OF KEYS GRAPH setup window.
- Select 'Text keys' or 'Time keys' with the pop-up menu.
- Click Calculate. The program (internally) calculates a matrix with the total joint duration of each pair of keys with different keywords (in case of time keys) or the total number of words that is assigned to both keywords (in case of text keys). The graph will show the highest joint durations or frequencies, thus showing which keywords tend to go together with which other keywords.
- You can now click the Graph button to show the graph.

You can adjust the graph to obtain a better picture. Just click on a field with a code, hold the mouse down and drag the field to another position.

OPTIONS

Only joint durations/frequencies exceeding a particular minimum will appear in the graph. This restriction may also cause that not all keywords appear in the graph. The maximum number of different keywords in a graph is 15.

You can adjust the minimum joint duration/frequency (and thus the number of different keywords: the higher the minimum, the lower the number of different keywords) with the Options button. If you click this button, a sheet appears. You can adjust the minimum with the slider, or you can type a value in the edit box after the slider. The number of relations between keywords given this minimum are shown after 'number of relations'. The number of different keywords that are involved in these relations are shown after 'number of keywords'.

In addition you can choose that relative durations/frequencies are shown instead of absolute durations/frequencies: the joint occurrence divided by the lowest occurrence of one of both keywords in the total column; see Coincidence of keys). You can also show the strength of the relations as phi-values, based on the joint occurrence (frequency of words, resp. duration) of two keys and the separate occurrence of each of both keys.

If you check **Indicate strength of relation**, the style of the lines represents the size of the absolute or relative co-occurrence, or the phi-value, depending on what you selected with the pop-up menu.

You can also use the options window to color the different fields. If **Show colored fields** is checked, the fields obtain the same colors as the bars of the corresponding keywords in the TEXT KEYS or TIME KEYS window (see Define keys how to change these colors).

SAVING OR COPYING THE GRAPH

You can save the graph as a PNG file using the disk icon at the bottom of the window. The output icon adds the graph to the output window, whereas the clipboard icon copies the graph to the clipboard (so you can paste it in for example a Word document). The maximum number of graphs (or other pictures) in the output window is ten.

Please note that the whole between title bar and blue separation line is saved as a picture. That is, if you have a lot of white space above the top of the graph, or below the bottom of the graph, this space will nevertheless be incorporated in the picture.

You may also want to adjust the size of your picture according to the width of the print area of the output field, i.e. the picture should fit into the output window. To this end, find the triangle at the blue separation bar (you may resize the window to make this triangle visible). The top of the triangle corresponds to the right margin of the output field. Your graph should not exceed this right margin if you wish to send the graph to the output window and print it.

7 Codes

Event codes (see also Event codes) are made up of one to nine codes, each representing a code variable. Codes (values on code variables) consist of one character, A..Z, a..z and 0..9; hence 62 different codes are possible for each code variable. Note that codes are case sensitive.

An event code is the succession of codes on the respective code variables. For example, the event code 'IQ2' means that the first code variable has the value (code) I, the second one is coded as Q and the third one as 2.

In addition to the allowed codes (A..Z, etc.), three other characters may occur in an event code:

- '–' has the general meaning of 'not coded'
- '?' means any code (e.g. useful for finding event codes)
- '•' means 'masked', that is 'neglect that code variable' (useful if you only want to take the other code variables into account for particular analyses).

SEQUENCE VIEWER offers options to define code variables (e.g. define which codes are allowed for a particular code variable), attach labels to codes, find and replace codes, and so on.

In addition SEQUENCE VIEWER provides you with a number of analyses on code variables, which in a sense is the heart of the SEQUENCE VIEWER program.

Manipulation of codes, and analyses on the codes of the sequences, are available from the Codes menu, but can also be found in some other menus. If you are looking for a particular analysis that involves codes, you may inspect these other menus too. In particular:

Edit > Assign empty codes
Text > Show event texts
Text > Auto code
Keys > Convert keys
Event variables > Convert to code variable
Event variables > Means

7.1 Define

INTRODUCTION

SEQUENCE VIEWER is especially designed for a so-called multivariate coding scheme. This means that each event is coded on a number of code variables. Each code variable covers a particular aspect or dimension of the event. Developing a coding scheme is a difficult task. If you are coding texts, a text should be viewed as a sequence of text parts, e.g. sentences. If you are coding social interactions, the events may be the utterances of the speakers, probably consisting of more than one sentence. It is strongly advised in that case to let the first code variable designate the source of a text part; for example the speaker if you are coding social interactions.

With 'Define code variables' you can change the name of a code variable, add a label to it, define the allowed codes (characters), define the type and define code labels.

Variable names can be maximally 12 characters long. Names are not case sensitive ('VAR1' is the same as 'var1'). Digits and spaces are allowed, but variable names should not consist of digits only. The first and the last character of a variable name cannot be a space.

SETUP

To change the properties of a code variable, do this:

- Select **Codes > Define event codes...** to open the DEFINE CODE VARIABLES setup window.
- Click at the name of a code variable in the 'Code variables' field.
- Type a new name in the field below the variable field and click the Change name button.
- Select a type for the variable. The type 'Category' means that the codes are categorical. If 'Number' is selected, only the numerical codes 0..9 are allowed. Numerical codes may be useful in case of agreement measures.
- Type the characters that are allowed for that code variable in the 'allowed codes' field. Only the characters 0..9, a..z and A..Z can serve as codes. Note that codes are case sensitive. Click Change allowed codes to accept the allowed codes, before you select a different code variable.
- To automatically fill the 'allowed codes' field with all possible codes, click at the 

button. To get the codes for the code variable as they actually appear in your file, click at the 

button
- To add a variable label, type a label in the 'variable label' field.
- To change a code label, click at a value (or its label) in the 'code labels field', type a new label and click Change label.
- Click Revert to set all properties of the selected variable back to its original values. This is not true however for code dependencies, as a dependency is not a property of one code variable, but of the combination of two code variables.
- Click Cancel to get rid of all changes (and close the window).
- Click Store to store all changes (and close the window).

NOTE

If you move the mouse over a code in the code field of the main sequence window with the shift key down, its label is shown as a tool tip.

7.1.1 Code dependencies

INTRODUCTION

Depending on your coding scheme, it is possible that allowed codes depend on preceding codes. For example, suppose the first code variable has the allowed codes 'A', 'B' and 'C' and the second 'X', 'Y' and 'Z'. It may be the case that after an 'A' on the first code variable, only a 'X' or a 'Z' may occur. Dependencies between codes only make sense of course if there are two or more code variables.

SETUP

To define such dependencies, do this:

- Select **Codes > Define event codes...** to open the DEFINE CODE VARIABLES setup window.
- Click the Dependencies... button.
- Select the 'preceding' or given code variable with the 'if' pop-up menu.
- Select the target code variable, that is the variable which codes depend on the codes of the given (preceding) variable with the 'then' pop-up menu.
- Type a 'preceding' code (e.g. 'A') on a line in the first (small) column and the codes (e.g. 'XZ') that are only allowed if preceded by 'A' after this code in the second column, on the same row.
- Instead of typing you can also click on a code in the field with allowed codes.

- Click Store before you define dependencies between other code variables. This button is disabled if nothing is entered yet.
- The Revert button restores the last stored dependencies.
- To undo all changes made to your dependencies, click the Reset button.
- After you have defined all dependencies, click OK.

OPTIONS

If you already have a file with codes, you can click Auto. The program now decides on the basis of which codes on a code variable actually occur after a particular code of the given code variable, that these codes are the only ones that are allowed after the given code.

In defining such dependencies, you can make logical errors. Suppose that after 'A' on code variable 1, only 'B' is allowed on code variable 2 and only 'C' on code variable 3. In such a case 'C' on code variable 3 should be allowed after 'B' on code variable 2. If 'C' is not allowed after 'B', this is inconsistent. You can check for such inconsistencies between the given ('if') code variable and the target ('then') code variable with the Check button. If code variables have many allowed codes, and the number of code variables between given and target code is large, such a check may take quite a long time. You may note that if you have nine event variables, and all codes are allowed for each code variable, the number of possible event codes is incredibly large (9^62).

If a code of a given code variable is not entered as 'if' code, all codes of the target ('then') variable, are allowed.

The Output button  sends an overview of all your dependencies to the output window.

With the Cancel button you can undo all changes since the last time you opened the DEPENDENCIES setup window. The Reset button does the same, but the DEPENDENCIES setup window remains open.

IMPORTANT

If you change the codes belonging to a code variable in the DEFINE CODE VARIABLES setup window, the code dependencies remain unaffected. For example, if you delete a particular code from the allowed codes, and this code is part of a dependency, this dependency is *not* deleted. So make sure to adjust your dependencies manually, in case you deleted codes.

7.2 Info

To get information about a code variable:

- Select **Codes > Show info...** to open the INFO OF CODE VARIABLE setup window..
- Click at one or more variable names to select then. Click a selected variable name once again to remove it from the list.
- Check ☑ **Show properties** to obtain labels for the variables and their codes.
- Check ☑ **Show frequencies** to obtain the frequencies of the different codes.
- Click OK.

You should either check ☑ **Show properties**, or ☑ **Show frequencies**, or both.

See also
Select marked sequences and Checked events only.

7.3 Add

To add a new code variable:

- Select **Code > Add code variable...** to open the ADD CODE VARIABLE setup window..
- Select the variable type.
- Type the position in the event code where the new code variable should be inserted.
- Type a character that will appear on this position of the event code.
- Click Add.

See also:
Event codes and Code variables

7.4 Delete

To delete a code variable:

- Select **Codes > Delete code variable...** to open the DELETE CODE VARIABLE setup window..
- Select the name of the variable to be deleted from the variables field.
- Click Delete.

7.5 View event codes

You can obtain a graphical representation of the course of a sequence (the succession of event codes) with this option. Each event is represented in the graph according to the 'length' of the event. This length can be based on time information (the onset and offset times, if available), or on the number of words of each event text (if codes and text are linked).

The graphical representation consists of a number of rows: one different row for each code variable. Each row consists of colored bars, each representing a particular code. The length of the bar corresponds to the length of the event as described above. If a particular code remains the same in successive events, these are combined into one bar.

Select **Codes > View event codes...** to open the VIEW EVENT CODES window. A pop-up menu in the lower right corner is available to choose between event times and event texts. In case of event times, you can apply a sequence definition. The graph is automatically created (and updated if you go to another sequence).

The graphical representation of the course of a sequence makes it more easy to see differences between codes on a particular code variable, or relations between codes of different code variables. For example, if the sequence concerns utterances in an interview, it is very easy to see which speaker talks the most; or if different speakers talk about different topics.

The picture can be saved as a PNG file with the 'hard drive' icon  in the lower left corner. You can also send the picture to the output window with  , or copy the picture to the clipboard using the clipboard icon  .

If you want to send the picture to the output window, the right side should not exceed the top of the triangle below the picture.

See also:
Define sequence.

7.6 Finding

With the FIND CODES setup window, you can find a particular event code in you file. In addition you can specify conditions for the position of the event code in the sequence or how the preceding or subsequent event code should look like.

The setup window also includes options for replacing event codes, (un)marking the sequences wherein these event codes occur, or (un)checking the found events.

See also:
Select marked sequences and Checked events only.

7.6.1 Find

SETUP

To find a particular event code, do this:

- Select **Codes > Find event codes...** to open the FIND CODES setup window..
- Select the **Find** tab.
- Type an event code after the 'Find' prompt. A question mark means any code on that position.
- You can restrict the search using the 'Where' pop-up menu. Type a conditional event code if appropriate.
- Select 'Search up' to search from the present sequence upwards, 'Search down' to search from the present sequence downwards or 'Search all' to search upwards from sequence number 1 with the 'Search direction' pop-up menu.
- Click Find.
- The button changes into Find next, allowing you to find the next instance of the event code.

ADDITIONAL INFORMATION

If you select 'If not followed by' with the 'Where' pop-up menu, the last code in a sequence will not be found, because it is assumed that (at least) one event code follows the target event. Similarly, if you select 'If not preceded by', the very first event will never be found. Use 'Last event' or 'First event' instead.

If you change one of the search specification, the Find next button changes to Find, and the search starts from the beginning (e.g. from the first sequence upwards. If during a search, you want to restart the search with the same specifications, just click the **Find** tab.

7.6.2 Replace

SETUP

To replace a particular event code by another event code, do this:

- Select **Codes > Find event codes...** to open the FIND CODES setup window..
- Select the **Replace** tab.
- Specify the Find options as described in the Find section.
- Type an event code after the 'Replace with' prompt. A question mark on a position means that the code on that position remains unchanged.
- Click Find. The button changes into Find next.
- Click Replace to change the event code and to find the next instance.
- Click Replace all to change all (checked) event codes of all (marked) sequences at once.

ADDITIONAL INFORMATION

If you select 'If (not) preceded by' with the 'Where' pop-up menu, there is a subtle difference between Replace and Replace all.

Suppose you want to replace 'A' with 'B', if 'A' is preceded by 'B'. Now consider the sequence 'B A A'. With Find you will find the first 'A'. If you now click Replace the sequence changes to 'B B A', and the program continues to find the next instance. This will be the remaining 'A' in the sequence, because it is *now* preceded by a 'B'. Clicking Replace will change this 'A' in a 'B' too.

However, if you use Replace all to change 'A' into 'B' if 'A' is preceded by 'B', only the first 'A' fulfills the condition, and only this 'A' will be changed. Hence, repeatedly clicking Replace may not have the same results as clicking Replace all.

7.6.3 Mark/unmark

To mark or unmark sequences having a particular event code, do this:

- Select **Codes > Find event codes...** to open the FIND CODES setup window..
- Select the **Mark** or **Unmark** tab.
- Specify the Find options as described in the Find section.
- Check ☑ **Unmark all sequences first**, respectively ☑ **Mark all sequences first** if appropriate.
- Click Mark or Unmark.

7.6.4 Check/uncheck

To check or uncheck events having a particular event code, do this:

- Select **Codes > Find event codes...** to open the FIND CODES setup window.
- Select the **Check** or **Uncheck** tab.
- Specify the Find options as described in the Find section.
- Check ☑ **Uncheck all sequences first**, respectively ☑ **Check all sequences first** if appropriate.
- Click Check or Uncheck.

7.7 Count event codes

INTRODUCTION

You can count the number of times a particular event code occurs in a sequence, and assign this number to a sequence variable. In addition you can enter a conditional event code.

SETUP

- Select **Codes > Count event codes...** to open the COUNT EVENT CODES setup window.
- Type an event code after the 'count' prompt.
- Select a condition with the 'when' pop-up menu, and type the conditional event code (unless you selected 'All occurrences' of course).
- Select or type the sequence variable to which the count should be assigned.
- Click Count.

ADDITIONAL INFORMATION

If you select 'not preceded by' with the 'when' pop-up menu, it is assumed that the to be found event code is preceded by another event code; hence, by definition, the first event code can never fulfill the condition 'is not preceded by'. Similarly, the last event code cannot fulfill the condition 'not followed by'. You can use the ☑ **include first/last event** checkbox, to nevertheless include the first or last event code in these cases.

If you apply a sequence definition, the count is performed on the transformed sequence.

See also:
Define sequence, Select sequences and Select marked sequences.

7.8 Copying

You can copy the codes of a code variable to another code variable, given a particular event code.

- Select **Codes > Copy code variable** to open the COPY CODES setup window.
- Select the code variable you wish to copy with the 'copy codes from' pop-up menu.
- Select the code variable to which the codes should be copied with the 'to code variable' pop-up menu.
- Check ☑ **copy only if event code is** and type an event code to restrict copying only if the event has that event code. A question mark means any code. If the event code consists of question marks only, copying is performed to all event codes in the file (and checking the checkbox has the same effect as not checking it).
- Click Copy.

See also:
Select sequences, Select marked sequences, and Checked events only.

7.9 Convert codes to event variable

You can change the values of an event variable, given a particular event code. If you choose the conversion based on a code variable, you can assign different values to an event variable for each different code on the code variable at once. If you choose the conversion based on an event code, only one value

- Select **Codes > Convert codes to event variable...** to open the CONVERT EVENT CODES TO EVENT VARIABLE setup window.
- Type the event code in the 'event code' edit box.
- Select or enter an event variable. This variable obtains the new value.
- Type a value for the event variable in the edit box after 'change [variable name] to:').
- Click Convert.

See also:
Select sequences, Select marked sequences, and Checked events only.

7.10 Convert codes to keys

INTRODUCTION

You can create keys that correspond to particular event codes. That is, you can either assign a text key to an event text with a particular event code, or assign a time key to the period between the onset time and the offset time of a particular event code.

SETUP

- Select **Codes > Convert codes to keys...** to open the CONVERT EVENT CODES TO KEYS setup window.
- Type an event code after 'Convert event code'.
- Select the **Assign text keys** or the **Assign time keys** radio button.
- Type a (new) name for the keyword below 'Assign keyword' or select an existing name with the combo box.
- In case of time keys, decide upon to which event code a pause belongs with the 'Event range pop-up menu. If there is a pause between the offset time of an event and the onset time of the next event, you should decide whether this pause belongs to the time key of the first event, to the time key of the next event, or to the time key of neither event.
- Click Convert.

ADDITIONAL INFORMATION

If you select text keys, applying a sequence definition is not possible, because it is not always clear which event texts belong to the event codes of the converted codes.

If you select time keys however, you can apply a sequence definition, because the event codes of the converted sequence have (adjusted) onset and offset times.

If there is a pause between the offset time of an event code and the onset time of the next event code, you can decide whether this pause belongs to the time key of the first event, to the time key of the next event, or to the time key of neither event.

See also:
Define sequence, Select sequences, and Select marked sequences.

7.11 Frequencies

To obtain a frequency distribution of event codes, do this:

- Select **Codes > Frequencies...** to open the FREQUENCIES OF EVENT CODES setup window.
- Type an event code in the 'enter' edit box. A question mark means any code. If, for example, one of the possible codes on the first code variable is 'R', and you type "R???" (suppose you have four code variables), you will obtain a frequency distribution of all event codes starting with 'R'. If the event code consists of only question marks, you will obtain a frequency distribution of all (selected) sequences.
- You can also obtain a two dimensional table of event codes by one code variable. Select 'code variable' with the 'by variable' pop-up menu. Enter or select a code variable. You can only select a code variable that corresponds with a question mark in the 'enter' edit box.
- To obtain a two dimensional table of event codes by an event variable, select 'event variable' with the 'by variable' pop-up menu. Select or enter an event variable.
- Finally you can obtain a two dimensional table of event codes by a sequence variable. Select 'sequence variable' with the 'by variable' pop-up menu and enter or select a sequence variable.
- You choose between frequencies and percentages in your table.
- Click OK.

See also:
Define sequence, Select sequences, Select marked sequences and Table options.

7.12 Frequency graph

INTRODUCTION

A frequency graph shows the frequencies of co-occurring codes of different code variables. For example, if an event has the event code 'IQ1A' then the codes 'I', 'Q', '1', and 'A' are occurring together. The program internally calculates a matrix with the joint frequencies of each code of each code variable with each code of all other code variables. The graph only shows the highest joint frequencies, thus showing which codes on a particular code variable tend to go together with which codes on other code variables.

SETUP

To get a graph, do this:

- Select **Codes > Frequency graph...** to open the CODE FREQUENCY GRAPH setup window.
- Click Calculate.
- The graph shows a number of rectangles, each representing a code on a code variable. The joint frequency between two codes is the number attached to the line between two codes.
- To adjust the graph, hold the mouse down on a rectangle and move it to a different place.
- Only joint frequencies exceeding a particular minimum will appear in the graph. This restriction will also limit the number of codes that appear in the graph. The maximum number of different codes is 15.
- With the Options button a number of options become available (see below).
- If you click the Graph button, the graph is redrawn.

OPTIONS

If you click the Options button, a sheet appears, with some options.

You can adjust the minimum joint frequency (and thus the number of different codes in the graph) with the slider. You can also type a value in the box after the slider. The number of relations between codes having a joint frequency of the minimum frequency or more are shown after 'number of relations'. The number of different codes that are involved in these relations are shown after 'number of codes'.

Because the dependency between codes may be defined by yourself (with Dependencies... in the DEFINE CODE VARIABLES window), you can give these relations a red color, to distinguish them from 'real' relations. A relation may also be colored red, if a code variable only has one allowed code.

In addition you can choose that relative frequencies are shown instead of absolute frequencies (the absolute frequency divided by the number of times the code occurs).

If you check **Indicate strength of relation** the style of the arrows represents the size of the absolute or relative frequency.

You can give the rectangles different colors, depending on the variable involved.

Finally you can choose to accept or neglect the code '–' for not coded yet with the **Accept dash (for not coded yet)**.

KEEPING THE GRAPH

You can save the graph as a PNG file using the disk icon at the bottom of the window.

The output icon adds the graph to the output window. The maximum number of graphs (or other pictures) in the output window is ten. Be sure the picture does not exceed the margins of the output window. The triangle on the blue separation bar corresponds with the right margin of the output window.

Use the icon to copy the graph to the clipboard, to paste it for example in a Word document.

See also:
Define sequence, Select sequences, and Select marked sequences.

7.13 Next event

INTRODUCTION

The 'Next event' analysis shows the observed and expected frequencies of all event codes following a particular event code. In the 'Next event' (and 'Previous' event) analyses the expected frequencies are calculated using the correction suggested by Sackett. The expected frequencies calculated with or without this correction will differ in case of structural zeros, but will lead to the same result if there are no structural zeros. For more information, see for example R. Bakeman and J. M. Gottman, Observing Interaction, second edition, Cambridge University Press, Cambridge (UK),1997.

SETUP

To obtain a frequency distribution of event codes (or target event codes), following a particular event code (the given event code) do this:

- Select **Codes > Next event...** to open the NEXT EVENT setup window.
- Type the given event code after the 'enter' prompt.
- Select whether you want all occurrences of the given code in a sequence, only the first occurrence of the given code, or only the last occurrence of the given code in the sequence.
- Select the number of lags.
- Select either 'expected frequencies' or 'by sequence variable' with the pop-up menu.
- If 'by sequence variable' is selected, you will obtain a frequency distribution of the target codes of the sequences having particular values on a sequence variable. Select or enter a sequence variable and choose if you want frequencies or percentages in the output table.
- If 'expected frequencies' is selected, you should decide on whether or not the target event and the given event can have exactly the same event code; see below. This option shows the expected frequency of each event code. Check ☑ **Don't show events with observed frequencies of 0**, if you want to prevent that these event codes are incorporated in the table.
- Click OK.

EXPECTED FREQUENCIES

Expected frequencies depend on whether or not zero transitions are structural. As a trivial example, if you throw a dice a number of times, an event (the number of pips) can very well be followed by the same event. Or two successive notes in music are very well possible. If you are studying turn taking in a group discussion (who takes the turn after who) however, the same person cannot have the turn after him or herself. Thus, the expected frequency that person A takes the turn after person A is 0. This is called a structural zero. Thus, if you analyze turn taking, musical compositions, or chess games in SEQUENCE VIEWER (which is very well possible!), or whatever kind of sequence, you should decide upon whether or not an event can be followed by a similarly coded event.

In SEQUENCE VIEWER you should check ☑ **Given event cannot be followed by target event with the same event code**, if that is the case. In some cases, it is already decided by SEQUENCE VIEWER whether or not this checkbox should be checked. For example, if a sequence definition is applied that unites equal event codes, a particular event code cannot immediately be followed (or preceded in case of a 'Previous event' analysis) by the same event code.

You can also check whether repeated consecutive event codes actually occur in your file with the Check button. If that is the case, the checkbox is checked and becomes disabled. Also if the event code of the given event contains one or more question marks (for 'any code'), SEQUENCE VIEWER assumes that there will be identical consecutive event codes.

Please note that if two similar event codes occur in succession, also if these event codes equal the given event (the event specified after 'enter'), it is assumed that the given event can be followed by a target event with the same code, whether or not this is actually the case.

See also:
Define sequence, Select sequences, and Select marked sequences and Table options.

7.14 Previous event

The 'Previous event' analysis shows the observed and expected frequencies of the event codes preceding a particular event code. Except this difference, all options are essentially similar to the Next event analysis. See Next event for more information.

7.15 Matrix

INTRODUCTION

A matrix analysis concerns the transitions between successive event codes, separated by zero or more other event codes, depending on the number of lags.

A lag of 1 means that the event codes follow each other immediately; a lag of 2 means that one other event code between the two event codes occur; and so on. The maximum is 9 lags.

The output is shown in the form of a transition matrix. A number of different matrices are available.

Frequencies

Shows the number of times each event code (the given event) is followed by which other event code (the target events). You can save the transition matrix, if you wish, for example to analyze the matrix with other statistical programs. The format is a simple tab/return-delimited ASCII file.

Proportions

Shows the relative frequency the given event is followed by the target event (relative to the total number of given events). Please note that the total number of given events equals the row sums of the frequency transition matrix. This means that the count of the number of given events ends at position *lag-1* of the sequence; if the number of lags = 1, the last event cannot be followed by an other event; if the number of lags = 2, neither the last event, nor the one but last event can be followed by another event; etcetera. This matrix can be saved too.

Expected frequencies

Expected frequencies concern the frequency that is expected if the occurrence of a target event would be completely independent of the given event, and is based on the overall occurrence of the target event.

The exception to this is when identical events cannot occur in succession. For example because you checked ☑ **unite events** in your sequence definition, or because the structure of your data prevents that identical events occur in succession. In this case, a target event should not be counted if it is equal to the given event (because the expectation is zero). The result is however that the row and column totals of the expected frequencies don't fit the row and column totals of the observed frequencies. To correct for this, expected frequencies are calculated using iterative proportional fitting, which ensures that the row and column totals of the expected frequencies are equal to the row and column totals of the observed frequencies.

z-values and probabilities

These values express whether a particular target event occurs below or above chance, given the 'given' event. Two somewhat different algorithms are available, Allison/Liker and Sackett.

Yule's Q and Phi

These statistics express the strength of the relation between the given event and the target event.

In addition to these matrices, also a general measure for the dependency between the given events and the target events can be calculated.

Additional information

For more information, see for example R. Bakeman and J. M. Gottman, Observing Interaction, second edition, Cambridge University Press, Cambridge (UK),1997.

SETUP

To perform a matrix analysis, do this:

- Select **Codes > Matrix...** to open the TRANSITION MATRIX setup window.
- Select the number of lags.
- Decide upon whether or not there are structural zero's (see expected frequencies above). If there are structural zero's, you should check ☑ **Given event cannot be followed by target event with the same event code**. The checkbox may be disabled (and set to the correct value) if the program infers from your sequence definition (e.g. because of 'unite events') that consecutive identical events cannot occur.
- You can also click Check, to check the file on the occurrence of consecutive identical event codes; after all, if consecutive identical event codes actually occur in your data, it's apparently allowed.
- Check the types of analyses you want to perform.
- Click OK.

Matrices can become very large. Although you can easily send a 500 x 500 matrix to your output file, this hardly ever makes sense. Moreover, it will take some time for the program to display the matrix in a neat format for the output window. Hence, in case of large matrices, you obtain a warning if you really want do display the matrix. If you click No, the analysis continues, the file is saved if you checked that option, but the matrix is not send to the output window. If you click Yes, the analysis may take some time if the matrix size is larger than 1000. Very large matrices may not be analyzed at all.

See also:
Define sequence, Select sequences, and Select marked sequences and Table options.

7.16 Transition graph

SETUP

A transition graph shows the transitions between event codes with a lag of 1 in the form of a graph. To get a graph, do this:

- Select **Codes > Transition graph...** to open the TRANSITION GRAPH setup window.
- Click Calculate. The program calculates a transition matrix, on which the graph will be based.
- You will see a number of circles, each representing an event code, and connected by arrows. For an explanation, see below.
- To adjust the graph, hold the mouse down on a circle and move the circle to a different place.

- Only transitions exceeding a particular minimum frequency will appear in the graph. This restriction will also cause that not all event codes appear in the graph of course. The maximum number of different event codes is 15.
- With the Options button you can adjust the minimum transition frequency, color the circles, and let the arrows themselves reflect the transition frequency (see below).
- If you click the Graph button, the graph is redrawn.

THE GRAPH

The circles represent the events. Above the event codes you can find the total frequency of that event code, corresponding to the column total of the event code in the transition matrix, i.e. the frequency of the 'target events'. Below the event code you find the frequency of that event code as 'given event' (the row totals). See also Matrix for an explanation of these concepts.

The circles are connected by arrows. The number attached to an arrow is the number of times an event code is followed by the other event code. If both the given event and the target event have exactly the same event code, this is indicated by a curve. To adjust the position of a curve, hold the mouse down on the attached frequency and drag it to another position.

OPTIONS

If you click the Options button, a sheet appears, with some options.

You can adjust the minimum transition frequency, and thus the number of different event codes, as well as the number of relations (the arrows) with the slider. You can also type a value in the box after the slider. The number of transitions with a transition frequency equal to or exceeding the minimum frequency are shown after 'number of relations'. The number of different codes that are involved in these relations are shown after 'number of codes'.

In addition you can choose that transition proportions (relative frequencies) are shown instead of transition frequencies (see Matrix: proportions).

If you check **Indicate strength of relation** the style of the arrows represents the size of the transition frequency or transition proportion.

Check **Color by code variable** to color the circles to better distinguish between different types of event codes. The color depends on the code of the code variable you select with the pop-up menu after the checkbox.

KEEPING THE GRAPH

You can save the graph as a PNG file using the disk icon at the bottom of the window.

The output icon adds the graph to the output window. The maximum number of graphs (or other pictures) in the output window is ten. Be sure the picture does not exceed the margins of the output window. If you resize the window to its minimum width, the picture exactly fits between the margins of the output window. The triangle on the blue separation bar corresponds with the right margin of the output window.

Use the icon to copy the graph to the clipboard, to paste it for example in a Word document.

See also:
Define sequence, Select sequences, and Select marked sequences.

7.17 Tree

SETUP

A tree shows the number of times a particular event code is followed by other event codes in the form of a tree. To perform a tree analysis, do this:

- Select **Codes > Tree...** to open the TREE setup window.
- Type a start event code after the start prompt. This event code serves as the root of the tree. A question mark means any code. If the start event code consists of only question marks, the first event code in a sequence will always serve as a root, else the first event code in the sequence that equals the start event is the root for that sequence.
- Select a tree depth. The tree depth limits the length of a tree: the maximum number of successive event codes a tree may consist of (the root is included in the tree depth).
- Type a minimum frequency. The minimum frequency also limits the size of the tree: only branches with the minimum frequency or more, are shown.
- If One branch per sequence is checked, only the first event code in the sequence that equals the start event code, is added to the tree with its successive event codes.
- If ☑ One branch per sequence is checked, after having added a sequence to the tree, the program continues with the next sequence. If not checked, the program looks into the same sequence if there is another instance of the start event code, after the last event code that is added to the tree, and adds it to the tree if it is found.
- Click OK.

ADDITIONAL INFORMATION

Results are shown as a tree, starting with the start event, preceded by the frequency of this event. After the start event one or more branches are shown: events immediately following the start event with their frequencies (provided that frequency equals or exceeds the minimum frequency). And so on.

If a branch ends with '[]', preceded by a frequency, this frequency is the number of times the preceding event code is also the last one in a sequence. For example ' 8 RA2•• 6 []' means that in 6 cases RA2•• is the last event in a sequence, whereas in two cases (8 - 6) it is followed by other events. If the last event code is not followed by '[]', there may be (or may not be) other event codes after this event, but with a frequency less than the minimum frequency.

See also:
Define sequence, Select sequences, and Select marked sequences.

7.18 Path

A path is quite similar to tree, but in addition an end event is specified. The output then shows all paths between the start event and the end event, provided the length of the path does not exceed the path depth. All options in the PATH setup window are the same as in the TREE setup window.

7.19 Patterns

INTRODUCTION

A sequence pattern is a particular succession of event codes that may occur in a sequence. SEQUENCE VIEWER evaluates each sequence whether or not the sequence has such a pattern. Here are some examples:

- event code 'C' is somewhere in the sequence
- event code 'C' is somewhere in the sequence and is immediately followed by event code 'X'
- event code 'C' is the first event, followed, but not necessarily immediately, by event code 'X', after which no event code 'W' occurs.

As already may be clear from these examples, applying a sequence definition highly affects whether or not a sequence has a particular pattern. Especially in combination with a sequence definition, sequence patterns is a powerful analysis, making it possible to define nearly every kind of pattern you can imagine. The drawback is that sometimes it can be difficult to understand why a particular sequence has the pattern, or does not have that pattern.

SETUP

- Select **Codes > Patterns...** to open the SEQUENCE PATTERNS setup window.
- Activate one or more pattern lines, by checking the checkbox before a pattern line. All lines below the activated (checked) line are deactivated (unchecked), whereas all line above the activated line, become activated (checked) too.
- Enter specifications for each activated pattern line (explained below in more detail).
- You can delete a pattern line with the  icon.
- You can insert a pattern line, with the  icon. Because the maximum number of pattern lines is six, the sixth pattern line is lost.
- Click Options to assign particular values to sequence, code or event variables of sequences having (or not having) the defined pattern. You can also (un)mark sequences with(out) the pattern. See pattern options for a discussion of these options.
- Click Apply to apply these options.

The output depends on the applied options. If you click Apply without any options, the output only gives you the number of sequences with and without the patterns.

CHECKING THE SEQUENCE PATTERN

Before applying a pattern definition, it is good practice to first inspect whether the pattern definition gives results according to your expectations:

- In the main sequence window, switch to the codes only mode.
- If you use a sequence definition, be sure the SEQUENCE DEFINITION window is open (click Apply instead of Done in this window).
- Click Find to find the first instance of the pattern in your file.
- If you use a sequence definition, check the  icon at the left side of the main sequence window. Now the converted sequence (because of the sequence definition) is shown.

The event codes that should be in the sequence, that is, with a $\in$ instead of a $\notin$ in the pattern line, are flagged. This makes it easy to check whether the indicated event codes indeed have the pattern you are looking for. If necessary, you have to adjust your pattern definition.

You can save up to ten different patterns (excluding the applied options). These patterns are

stored into your sequence file and can be used for later analyses. A pattern is automatically saved with the selected pattern name when you save your sequence file.

EXAMPLES

A pattern consists of maximally 6 event codes. In its most simple form, a pattern concerns a number of coded events that should appear in a sequence, let us say event codes A and D.

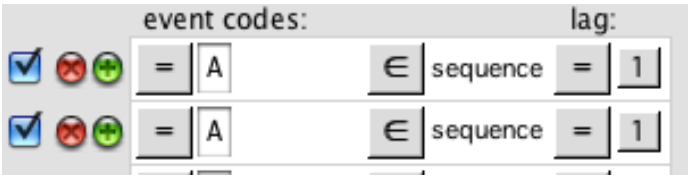
Example 1

Suppose that event A should occur at the first position in the sequence and event D at the second position. The setup should look like this:

The checkboxes of the first two lines are checked, because the pattern concerns two events; al other checkboxes should be unchecked.
Event code 'A' should appear in the sequence, hence the toggle button after the 'A' should have the symbol ∈ ('is in sequence').
Event code 'A' should be the first code of the sequence, hence, the lags = 1.
Event code 'D' should appear in the sequence, hence the toggle button after the 'D' should have the symbol ∈.
Event code D should follow code A immediately, hence, the lags = 1.
For each sequence, the program first checks if the first event is 'A'. If that is the case, then it is checked if the event code immediately after 'A' equals 'D'. If yes, the sequence has the pattern.

Example 2

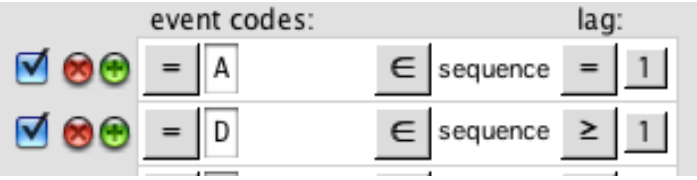
Now we change the 'D' on the second pattern line to 'A'.



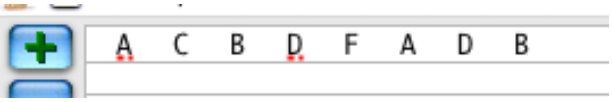
The program now finds sequences where both the first and the second event code equal 'A'.
Suppose in addition that we use a sequence definition, with **Unite events** checked. Then the pattern cannot occur: in the converted sequence two successive event codes cannot be the same.

Example 3

Now suppose you want to identify sequences with A at the first position whereas D should occur in the sequence, but not necessarily at the second position. Click at the toggle button with the equals sign below 'lag:' in pattern line 2 (with the D); it will change to ≥.



If we click the Find button, the program may find a sequence like this (in codes only mode):

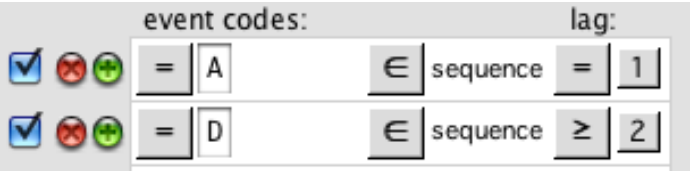


As you can see, the event codes from both pattern lines are flagged, making it easy to identify the pattern in the sequence. This is of course particularly useful in case of long sequences, complex patterns, and a larger number of code variables.

You may also observe that the pattern occurs twice in the sequence; events 6 and 7 also fulfill the pattern. SEQUENCE VIEWER identifies only the first occurrence of a pattern.

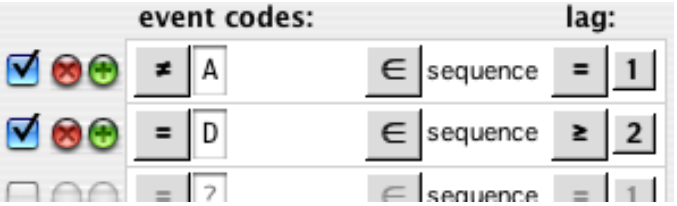
Example 4

Suppose you want to find sequences with A at the first position whereas D should occur in the sequence, but not immediately after 'A'. To this end, select 2 as the number of lags on pattern line 2.



Example 5

Now click at the toggle button before 'A', to change it to '≠'.



The program first checks if the very first event code does not equal 'A'. If that is the case, the program looks if this event code is followed (excluding the event code immediately after this event code, because lags ≥ 2) somewhere by a 'D'.

Example 6

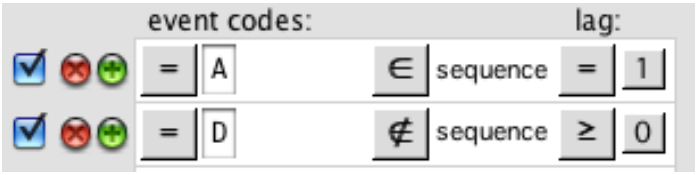
Thus far the basics. Let's make things a bit more complicated. Suppose you want to have sequences with 'A' somewhere in the sequence, whereas after the 'A' nowhere in the sequence a 'D' should occur. Take a look at this setup:

Note the ∉, ('is not in sequence'). You can click this button to toggle between ∈ and ∉.

A sequence like 'B X A C F' certainly satisfies the pattern. There is an 'A', and this 'A' is nowhere followed by a 'D'. Thus the program will find such a sequence.

But what about 'B X A'? Here too, there is an 'A', and this 'A' is not followed by a 'D'; nevertheless, the sequence will *not* be identified as having the pattern.

The reason is that 'lag ≥ 1' on the second pattern line *implies that 'A' should be followed by at least 1 event*. If you nevertheless also want a sequence like 'B X A' to be identified as "somewhere an 'A', which is not followed by a 'D'", the solution is changing the number of lags from 1 to 0.



You should note the difference between '≠' before the event code, and '∉' ('is not in sequence'). That is, "event code 'A' is not in the sequence", is usually not the same as "an event code different from 'A' should be in the sequence". However, the pattern line:

(1) = A ∉ sequence = 3

and the pattern line

(2) ≠ A ∈ sequence = 3

will find the same sequences.

Consider for example the sequence 'B D F G A'. Both pattern lines evaluate the third event code, a 'F'. Event code 'A is not in this part of the sequence, hence the sequence fulfills pattern line (1). The 'F' is also different from 'A', that is, it is a 'not A'. Thus, the sequence also fulfills pattern line (2).

If you have the choice between different pattern lines to obtain the same result, it is recommended to use a pattern line with '≠' instead of '∉'. First, the search algorithm is usually much faster for pattern lines with '≠' than for pattern lines with '∉'. Second, only event codes from pattern lines with '∈' are identified. For example, if you use the Find button, the 'F' in sequence 'B D F G A' will only be flagged if you use pattern line (2).

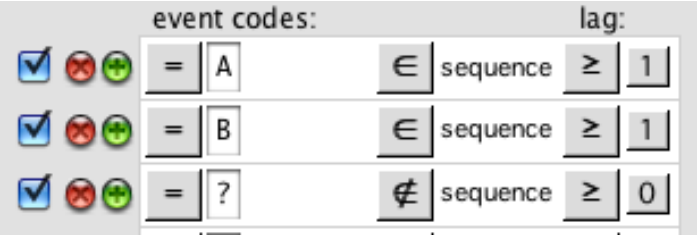
In addition:

A number of lags of 0, can only be selected for pattern lines with ∉, not with ∈.
If the number of lags is 0, only the ≥ operator for lags is possible; = is not allowed, and the toggle function does not work in that case.
Combination of '≠ event code' and '∉ sequence' is not possible.
Note that '∉ sequence' does not mean 'is not in *whole* sequence'; instead 'is not in that part of the sequence, bounded by the positions of event codes from the previous and next pattern lines' is somewhat more accurate.

Example 7

Suppose you want to find sequences with an 'A' somewhere in the sequence, whereas the very last event code in the sequence should be a 'B'. This can be accomplished by adding the pattern line:

= ? ∉ sequence ≥ 0



Remember that '?' means any code. If such pattern line is the last pattern line, it has the special meaning that the event code of the previous line should be the last event code in the sequence. If that previous event line contains a ∉, the position of that event is undetermined, and nothing will be found.

If there is more than one code variable, all codes of the event code in the pattern line should be a question mark.

Example 8

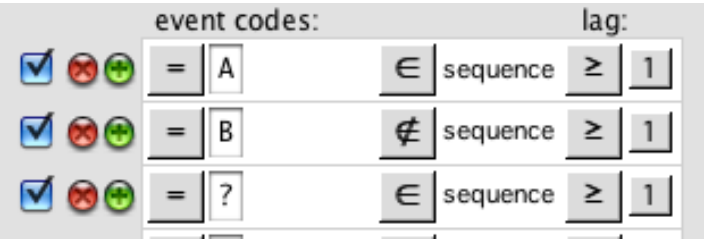
As another example take this setup:

Between A and C, event D should not occur. The 'lags ≥1' on the second pattern line implies that between A and C at least one other event should occur. Thus:

'A X C' fulfills the pattern. You may note that only event codes that can be positively identified, that is, having '∈sequence' in their pattern line, are flagged.
'A C' does not fulfill the pattern, because there should be at least one event code between 'A' end C'.
'A C C' fulfill the pattern; there is a (non-D) event code between A and the second 'C'.
'A D C A B C' fulfills the pattern.
'A B D C' does not have the pattern, because D appears between A and C.
But now suppose that instead of
=D ∉sequence lag ≥1'
pattern line 2 reads:
≠D ∈sequence lag ≥1
'A B D C' does fulfill the pattern: there is a 'not D' between A and C, i.e. B.

Example 9

Finally, take a look at this setup:



What about sequence 'A C B D F'?

The result may be a bit surprising ('A C B D F'), because you may think that 'B' should not occur in the sequence. Earlier however, we saw that 'B ∉sequence' rather means that B should not occur in that part of the sequence bounded by the previous and the next pattern line.

The 'A' in the sequence of course corresponds to the 'A' in the first pattern line. The 'B' in the sequence however corresponds to '?' in the third pattern line: 'B' is an 'any code'. Between the event code found according to the first pattern line and the third pattern line, should be at least one event code: the number of lags ≥ 1. And between both event codes in the sequence a 'B' should not occur. This appears to be the case: no 'B', just a 'C'.

If the number of lags in pattern line 2 is ≥0, instead of ≥1, the result would have been: 'A C B D F'; if the number of lags in pattern line 2 is ≥2, the pattern does not occur in the sequence.

See also:
Pattern options, Define sequence, Select sequences, and Select marked sequences.

7.19.1 Pattern options

A number of options can be applied regarding the sequences with and without the specified pattern.

- To change the value of a sequence variable, for those (selected) sequences that have the pattern, check ☑ **Change sequence variable**, and enter or select a sequence variable and a value.
- To compare the means of a sequence variable for the sequences with and without the pattern, check ☑ **Means of sequence variable** and enter or select a sequence variable. If you changed the value of the sequence variable (see previous option), the mean is calculated after the change.
- You can change the code of a code variable, for those event codes that are positively identified (that is, with '∈sequence' in their pattern line). That are also the event codes that are flagged when you use the Find button. To this end, check ☑ **Change code variable**, enter or select a code variable, and type a code.
- You can assign the position in the pattern of the positively identified event codes in the sequence to an event variable. For example, suppose the pattern has two pattern lines. The event code from the first pattern line is identified as event code number 3 in the sequence, whereas the event code from pattern line 2 is identified as event code number 5 in the sequence. The event variable than obtains value '1' on the third event and value '2' on the fifth event. If you use the Find button in the SEQUENCE PATTERN setup window, you can find at the bottom of this window at which positions particular pattern lines are found.
- Finally you can mark or unmark the sequences with or without the pattern.

7.20 Count sequences

Your sequence file may contain a large number of similar sequences. To obtain a frequency distribution of similar sequences:

- Select **Codes > Count sequences...** to open the COUNT SEQUENCES setup window.
- Type a minimum frequency. Sequences occurring less than the minimum frequency will not be displayed, to prevent a large table with a lot of infrequently occurring sequences.
- Click Count.

See also:
Define sequence, Select sequences, and Select marked sequences.

7.21 Define agreement

INTRODUCTION

Two sequences, considered as strings of codes, can be more or less alike. Apparently, the sequences:
A B G K
and
A B G K
agree perfectly.

On the other hand, the sequences:
A B G K
and
D M X
seem to be completely dissimilar.

Generally, the agreement between two sequences can be expressed as a number between 0 (completely dissimilar) and 1 (exactly the same). SEQUENCE VIEWER offers you a number of algorithms to calculate the similarity between two sequences.

- The DT algorithm is the fastest one. Moreover you can use an approximation, which will usually give the same result as the exact solution (especially if sequences are relatively short and/or contain few similar codes) but may be ten times as fast, or even faster than the exact algorithm.
- The OM algorithm (Optimal Matching) is the most widely used approach, but yields sometimes quite implausible results.
- The tuples approach is the slowest one, but probably also the best one.

DT

DT is based on the number and order of the elements both sequences have in common. DT offers you 4 versions, alfa, beta, gamma and delta. Generally, the lower the proportion of common elements, the lower the agreement. Alfa is the most sensitive to the proportion of common elements; that is, alfa decreases fastest with a decreasing proportion of common elements. Gamma is the least sensitive, whereas beta takes an intermediate position. Conversely, gamma is more sensitive to the similarity of the order of the common elements in both sequences, whereas alfa is the least sensitive in this respect.

Delta is somewhere between beta and gamma with respect to sensitivity to the proportion of common elements, but this measure has some other properties. When sequences have different lengths (different number of elements), they will always have at least as much elements not in common as they differ with respect to their length. Delta is less sensitive for differences in lengths than the other measures.

For a description of DT, see W. Dijkstra & A. W. Taris (1995) 'Measuring the Agreement between Sequences', Sociological Methods & Research, 24 (2), 214-231.

If event codes consist of more than one code variable, you can calculate the average agreement (in case of DT). Suppose you have the sequences:
ABX CQY VWZ RTU

AQX DQY VWZ RSU
Both sequences are pretty dissimilar; they only have one event code in common. However if we mask for example the second code variable (see Define sequence), we obtain
A•X C•Y V•Z R•U
A•X D•Y V•Z R•U
and both sequences have three event codes in common. We can calculate the agreement for each separate code variable, and use the average of these agreements as our measure. In the example, we would calculate the agreements between:
A C V R and A D V R
B Q W T and Q Q W S
X Y Z U and X Y Z U

Optimal matching

Optimal matching is based on the number of necessary operations, to make two sequences similar. Operations are deletions, insertions and substitutions. Insertions and deletions are also called indels. For example, to make the sequences:
1: A B C D
2: A B E D F
equal, one might delete 'C' from sequence 1 and 'E' and 'F' from sequence 2. Instead, one can also substitute 'E' in sequence 1 for 'C' (and delete 'F' from sequence 2). In the latter case, only two operations are necessary. This is called the 'Levenshtein' distance, or simple alignment. To obtain a distance measure between 0 and 1, one should divide by the maximum number of operations that is necessary to make both sequences equal. If substitutions are allowed, this will be the number of elements of the longest sequence. An agreement measure is obtained by subtracting the distance from 1.
The process of making both sequences equal is called alignment. The aligned sequences may look like:
1: A B C D
= = s = d
2: A B E D F
'=' signifies that both elements are equal, 's' signifies a substitution, and 'd' a deletion ('i' would signify an insertion).
Aligning sequences can be fairly complex, especially if long sequences are involved. Generally, more solutions are possible; optimal alignment or optimal matching looks for the solution with the least number of indels and substitutions. The number of indels and substitutions is called the cost of the alignment. It is possible to weight the cost of an indel operation, relative to a substitution. A '1' means that a substitution cost as much as an indel operation. A 2 means that a substitution cost as much as a deletion plus an insertion (because the same effect of a substitution can be accomplished both an insertion and a deletion). Values lower than 1 for the substitution/indel ratio usually do not make much sense. A value of 0 means that substitution costs nothing (and hence any two sequences of equal length will have an agreement of 1). For more information, see for example A. Abbott & A. Hrycak (1990), 'Measuring resemblance in social sequences', American Journal of Sociology, 96, 245-274.

If event codes consist of more than one code variable, you can weight the cost of substitutions. For example, the event codes 'IQ0' and RA1 are completely dissimilar. The event codes IQ1 and RA1 however, have the third code variable in common, and hence it can be argued that substitution should cost less. To weight substitutions in this way, you can check ☑ **Weight by event length** in the setup window.
If one or more code variables are numerical, you can also weight substitutions by the distance between corresponding codes. For example, to substitute 'RA1' by 'RA2' should cost less than substitution of 'RA1' by 'RA5' (provided of course that the third code variable represents an interval scale). You can apply this weight procedure by selecting 'OM weighted' from the

'Agreement measure' pop-up menu. In that case, substitutions are also weighted by event length (if there are more code variables), as described above. If you don't have numerical code variables, the item 'OM weighted' is disabled.

Tuples

The tuples approach is based on the number of k-tuples that both sequences have in common. What is a k-tuple? Consider the sequences:

1: A B C D
2: A B E D F

A 2-tuple is 'A B', 'B C', or 'A C' in sequence 1. Or 'A B', 'B E' or 'B F' in sequence 2. A 3 tuple in sequence 1 is 'A B D'. There are also 1-tuples of course: 'A', 'B', etc.

The more k-tuples both sequence have in common, the more they are alike. In the example, both sequences have the 1-tuples 'A', 'B' and 'D'; the 2-tuples 'A B', 'A D' and 'B D'; and the 3-tuple 'A B D' in common. It can easily be seen that if sequence 1 was 'A B E D', they would have more k-tuples in common. On the other hand, if sequence 2 was 'A X E D F', they would have less k-tuples in common. Given the number of common k-tuples and the maximum number of k-tuples of both sequences (fully determined by the lengths of both sequences), this yields an elegant measure of agreement, without the need to delete or insert elements.

A constraint can be placed on the maximum size of tuples. In case of long sequences, you may decide to look only to tuples no longer than, for example, five. Especially in case of long sequences, this drastically reduces calculation time. Generally, the agreement will be higher if the maximum size of tuples is reduced, because sequences usually do not have long tuples in common, whereas long tuples do add to the maximum number of tuples in a sequence. In the program, you can fix the maximum tuple size to a number between 1 and 9, or to the length of the shortest sequence ('min') or to the length of the longest sequence ('max'). Selecting either 'min' or 'max' does not affect the agreement (because tuples with a size larger than the shortest sequence cannot occur in that sequence).

More information about the tuples measure can be found in: C. H. Elzinga (2003), 'Sequence Similarity - A Non-Alignment Technique', Sociological Methods & Research, 32 (4), 3-29.

THE SETUP WINDOW

The DEFINE AGREEMENT window serves two purposes.

First, it shows you how the options of the different measures affect the agreement. In case of optimal alignment it also shows you how sequences are aligned.

Secondly, the window is used to define which measure and options are used in calculating the agreement between sequences in the analyses that use agreement measures. To change the agreement measure in one of the setup windows that use agreement measures, click the  icon at the bottom of that window. This brings the DEFINE AGREEMENT window to the front.

SETUP

- Select **Codes > Define agreement...** to open the DEFINE AGREEMENT setup window.
- Select an agreement measure.
- In case of a DT measure, either select **Exact** or **Approximate**.
- Check **Average** to calculate the average agreement of the separate code variables.
- In case of an OM measure, enter a ratio for the relative cost of a substitution versus an indel operation.
- Check **Weight by event length**, if you want to weight the cost of a substitution by the number of dissimilar codes of an event code.
- In case of tuples, select an appropriate maximum tuple size.
- Enter a number for both sequences that should be compared.

• In case of a OM measure you can click the Align button, to obtain an optimal alignment solution.
• Click Define to use the selected agreement measure and its options for use by other analyses.

ADDITIONAL INFORMATION

The agreement between two empty sequences is assigned a 'M', but this may be different in some other analyses. The agreement between an empty sequence and a non-empty sequence is always 0 (because they have nothing in common).

Optimal matching and especially the tuples approach may take a long time in case of very long sequences (some hundreds of events). You can abort a tedious calculation with command-. (key-period).

The agreement measure as calculated by the tuples approach is usually much lower, compared to the DT and OM algorithms.

Very long sequences cannot be displayed in the DEFINE AGREEMENT window, due to limits of the width of text in a field.

You can apply a sequence definition, to observe the effect of, for example, masking on the agreement. The sequence definition is not part of the agreement definition however.

See also:
Define sequence.

7.22 Assign agreement

INTRODUCTION

You can assign the agreement between the sequences in your file and a sequence of event codes to the sequences in your file.
To this end, you should first enter a sequence (the target sequence) with which you want to compare the sequences in the file. The target sequence can be an existing sequence from your file, an edited existing sequence, or a completely new sequence. You should also select a sequence variable, to which the agreements will be assigned.

SETUP

• Select **Codes > Assign agreement...** to open the ASSIGN AGREEMENT setup window.
• Enter the target sequence (see below).
• Select or enter a sequence variable. The agreement with the target sequence is assigned to this variable.
• Click Assign.

THE TARGET SEQUENCE

You can use an existing sequence from your file as the target sequence. Just type the number of this sequence in the 'Show sequence' edit box. The sequence is shown in the target sequence field. You can use the  and  buttons to go to the next or the previous sequence.

To edit an event code of the target sequence, click at an event code in the target sequence field. Edit the event code in the 'Edit selected event code' edit box'. Changes are immediately reflected in the target sequence field.

With the Add button you can add an empty event code (dashes only) after the last event code of the sequence (irrespective of which event code is selected). With Insert you insert an empty

event code before the selected event code, and with Delete you delete the selected event code.

If you don't want to use an existing sequence, but enter a completely new sequence, first click the Add button and than proceed in the way described above.

NOTE

If you apply a sequence definition, this sequence definition is *not* applied to the target sequence, *except* the masking transformation (if appropriate). Masked codes are not shown in the target sequence field however.

See also:
Define sequence, Define agreement, Select sequences, and Select marked sequences.

7.23 Agreement matrix

INTRODUCTION

An agreement matrix calculates the agreement between all pairs of (a selection of) sequences in your file. Alternatively you can also calculate the distance, defined as 1 minus the agreement.

You can save the matrix as a file to disk for analysis by different programs, e.g. to perform a cluster analysis. You can also write the matrix to the output window.

You should be aware of the fact, that if your file contains many sequences, the calculation of the matrix can take a very long time, and the matrix can be very large.

If a large matrix is sent to the output window, you are informed about how many kB or mB the matrix may add to your output window and you are allowed to cancel the analysis. If you have thousands of sequences, you can easily add many megabytes to the output window. The maximum size of a matrix that can be analyzed by SEQUENCE VIEWER is 5000 * 5000. If empty sequences (sequences without codes) occur in your data file (or empty sequences occur because of a sequence definition), you have to decide how to handle them: you can neglect empty sequences; assign a 'M' in case an empty sequence is involved; or assign a '1' in case both sequences are empty (because they are 'alike') and a '0' if only one of both sequences is empty.

The time the calculation of the matrix takes depends of course on the number of sequences; in case of a 5000 x 5000 matrix, nearly twelve and a half million agreements should be calculated. Calculation of one of the DT measures takes much less time than the other measures, especially if the approximate measure is used, whereas the length of the sequences themselves also considerably affects the duration of the calculations.

As usual, during the analysis, you are constantly informed about how long the remaining analysis will approximately take and you are (also as usual) allowed to abort the analysis (with command-period).

SETUP

To calculate an agreement matrix, do this:

- Select **Codes > Agreement matrix...** to open the AGREEMENT MATRIX setup window.
- Decide whether to calculate the **Agreement** or the **Distance**.
- Decide whether to calculate a **Square matrix** or a **Triangular matrix**.
- In the latter case, decide whether or not to **Include diagonal 1's** (or diagonal 0's, in case of distances).
- Decide upon what to do with empty sequences.

- Check ☑ **Send to output window** only in case of a small matrix. Large matrices can hardly be interpreted because of the massive amount of data.
- Checking ☑ **Write to file**, is especially useful if you want your matrix to be analyzed by some statistical program.
- Click OK.

See also:
Define sequence, Define agreement, Select sequences, Select marked sequences and Table options.

7.24 Cluster sequences

INTRODUCTION

A cluster analysis takes three steps.
- First, all agreements between all pairs of sequences are calculated.
- Second, clusters of sequences with a high inter or within agreement (relative to the intra or between agreement) are formed.
- Third, the number of the cluster to which a sequence belongs can be assigned to a sequence variable.

The first step involves the calculation of an agreement matrix. Because the calculation of the original agreement matrix can be very tedious, the program keeps this matrix in memory. You are informed whether or not the matrix is available. As soon as you change something that could affect the matrix (e.g. changing the selected sequences), the matrix will be removed from memory and will be calculated again. Because the calculation of the matrix can take a lot of time, we recommend to use one of the DT measures with the approximate calculation, at least for the exploratory phase of a cluster analysis. After all, if there are clear clusters in your data, you will find them, irrespective of the agreement measure. If there are no clear clusters, neither agreement measure will be of much help.

The maximum size of an agreement matrix is 5000 x 5000 (see also Agreement matrix).

Clusters

Clustering sequences is based on a hierarchical cluster analysis. The procedure is as follows.

Starting with the agreement matrix as calculated in step 1, the pair with the highest agreement is selected as the first cluster in the first run. In the next *run*, either another sequence pair is selected, or the combination of the cluster from the first run with some sequence, depending on the size of the between agreements of the constituting sequences. In the next runs, a cluster may consist of two other sequences not belonging to a cluster yet, the combination of a not-clustered sequence and a cluster, or the combination of two existing clusters.

Which cluster is formed depends on the mean size of all agreements between the constituting sequences.

The forming of clusters stops when all sequences are clustered, or when the number of runs has reached the limit as set by the user.

Example

Suppose, we have cluster A, consisting of sequence 2, 5 and 6; cluster B, consisting of sequence 1 and 7; whereas sequence 3 and 4 are not clustered yet (suppose there are only 7 sequences).

The program now compares the mean agreements between:

- sequence 3, 2, 5 and 6 (that is, 3 and the sequences from cluster A)
- sequence 3, 1 and 7 (that is, 3 and the sequences from cluster B)
- sequence 3 and 4
- sequence 4, 2, 5 and 6 (that is, 4 and the sequences from cluster A)
- sequence 4, 1 and 7 (that is 4 and the sequences from cluster B)

Whichever combination has the highest mean agreement will become the next cluster. Suppose this is 3 and B, then sequences 3, 1 and 7 will form a new cluster, C, whereas cluster B is thrown out.

Evaluation of output

To evaluate the results, for each run you will get information about:
- The number of clusters
- The mean within agreements
- The mean between agreements

Both within and between mean agreements are calculated in two different ways, unweighted and weighted (or adjusted).

The unweighted within mean is calculated as follows. First, for each cluster the mean of all agreements between sequences belonging to that cluster is calculated. Next the mean of all these mean agreements is calculated. The size of the cluster is not taken into account; the contribution of a small cluster is equal to the contribution of a large cluster. The weighted version on the other hand takes these differences into account. You should further note that the agreement of a sequence with itself is not taken into account.

The between agreements are calculated in a similar way, and concern the agreement between all sequences in a cluster on the one hand, and the sequences belonging to different clusters.

In all these calculations, sequences that are not clustered yet, are not taken into account.

Generally, a cluster solution is better, the higher the within agreements and the lower the between agreements. For this reason, the ratio within/between agreements is also calculated. This does not necessarily means that the solution with the highest ratio is the best one. For example, you may end with two clusters, one consisting of two sequences with an agreement of 1, and a cluster consisting of all other sequences, that happen to have very low agreements with the two-sequence cluster. You should inspect the run information carefully, in order tot decide upon the best solution.

The output also gives you the sequence numbers belonging to each cluster. Finally, for each cluster the event codes of the sequence with the highest mean agreement with the other sequences in the cluster is shown. You may view this sequence as the most typical sequence of the cluster, which may help you to interpret the meaning of a particular cluster.

Assign clusters

Finally, you can assign the number of the clusters to a sequence variable. To this end you should decide upon the run which gives you the best solution.

Empty sequences

You should be aware of empty sequences. If there are a number of empty sequences, they are quite probably put together into one cluster (because the agreement between two empty sequences is 1 by definition). You are advised to exclude empty sequences from a cluster analysis. Moreover, because an empty sequence has an agreement of 0 with a non-empty sequence, you may obtain a cluster solution that looks better than justified. Especially if you use a sequence definition, a non-empty sequence can become empty, thus it may not be obvious from the original sequence if a sequence becomes empty.

SETUP

To perform a cluster analysis, do this:

- Select **Codes > Cluster sequences...** to open the CLUSTER SEQUENCES setup window.
- Select the **Get clusters** tab.
- Type the range of the runs about which you want to be informed. The maximum number of runs is the number of sequences that should be clustered minus 1. The minimum number is 1. Because the lower runs are usually not very informative, you better type a higher number as minimum number, for example at about 90 % of the number of sequences.
- Check ☑ **Exclude empty sequences**, unless you want a separate cluster for the empty sequences.
- Click Cluster.
- Decide upon one or more run numbers with probably good solutions.
- Enter these run numbers in successive analyses as the last run in the range (in the 'to run' edit box). This provides you with additional information about that solution (e.g. the cluster matrix and the 'most typical sequence').
- Decide upon the run number with the best solution.
- Select the **Assign clusters** tab.
- Enter the run number of the best solution.
- Enter or select a sequence variable. The cluster number to which a sequence belongs is assigned to this variable. A sequence that does not belong tot a cluster, is assigned 'M'.
- Click Assign.

See also:
Define sequence, Define agreement, Select sequences, and Select marked sequences.

7.25 Reliability

INTRODUCTION

To calculate the reliability of the codes, it is assumed that you have two files of the same data available, coded by different coders. The first one is the presently open file. The second file should be comparable with the first one, that is, the number and order of code variables should be the same, the number and order of sequences should be the same and the number of events per sequence should be the same. If the names of the code variables are different, you will get a warning, but you can continue the analysis, if you want. The reliability measure used is the common Cohen's kappa.

SETUP

- Select **Codes > Reliability...** to open the RELIABILITY setup window.
- Click Get file to select the file that should be compared with the open file.
- You can specify a particular event code. For example, if this event code reads 'R???' you will obtain the reliability of all events that will have an 'R' at the first position (the first code variable) in both files.
- Check ☑ **matrix** to obtain a matrix of all different codes from the one file by all different codes of the other, showing how many times both coders agree or disagree.
- Check ☑ **kappa per event** to obtain the reliability of each event code.
- Click OK.

CALCULATION OF KAPPA

Kappa is calculated according to the formula:
(obs freq - exp freq) / (total - exp freq).

- *obs freq* is the number of events that are coded the same by both coders.
- *exp freq* is the number of events that are expected to be coded the same by both coders, if the coders assign codes randomly. The expected frequency is calculated as the product of both marginal frequencies divided by the total number of events; and summed over all different event codes.
- *total* is the total number of events that are coded by both coders.

In case of kappa per event, the meaning of the elements of the formula above, is a bit different. For a particular event code the meaning is:

obs freq is the number of times that both coders assign that event code to an event (if summed over all different event codes, it yields the overall observed frequency, as is used in calculating overall kappa).

- *exp freq* is the product of both marginal frequencies divided by the total number of events (if summed over all different event codes, it yields the overall expected frequency).
- *total* is the number of times coder 1 assigns a particular event code, plus the number of times coder 2 assigns that event code and divided by 2 (if summed over all different event codes, it yields the total number of events).

For example, suppose the matrix looks like this ('A', 'B' and 'C' are event codes):

	coder 1			
coder 2	A	B	C	total
A	23	5	1	29
B	3	62	8	73
C	2	6	44	52
total	28	73	53	154

This yields the following results:

	coder		frequency				
code	1	2	obs	exp	total	kappa	perc
A	28	29	23	5.27	28.50	0.763	80.7
B	73	73	62	34.60	73.00	0.714	84.9
C	53	52	44	17.90	52.50	0.754	83.8
total	154	154	129	57.77	154.00	**0.740**	**83.8**

For example, the results for event code 'A' are calculated as follows:
- expected frequency: (28*29)/154 = 5.27
- total: (28+29)/2 = 28.50
- kappa: (23-5.27)/(28.50-5.27) = 0.763
The percentage agreement ('perc') is calculated as:
(23/28.50) * 100 = 80.7

Weighted kappa

The weighted kappa takes into account that two coders may have assigned not completely different codes to an event. For example, if one observer assigns code 'ABCD' and the other observer 'ABFD', they have assigned the same code to three of the four code variables. The weighted kappa weights the disagreement between two codes, according to the number of code variables that are coded differently. For example, if there are four code variables, and coder 1

codes an event as 'RA0X', whereas the second coder codes that same event as 'RA1Y', the corresponding cell in the matrix is weighted by 0.5.

IMPORTANT

You should be careful to apply a sequence definition. You can safely use the mask option (e.g. '?•••' gives you the reliability of the first code variable), or define a subsequence using event positions (e.g. from event 2 to L neglects the very first event in each sequence). However, all transformations that may change the number of events in a sequence may either cause the length of two sequences to become unequal (which will abort the analysis) or cause discorrespondence between the events of a sequence in both files.

You should further note that if you select sequences (e.g. the marked sequences only), this will apply to the open file. That means if for example sequence number 1 is selected because it is marked in the open file, sequence number 1 from the other file will be selected too, irrespective whether it is marked in the other file or not.

See also:
Compare sequences, Define sequence, Select sequences, and Select marked sequences.

7.26 Compare sequences

INTRODUCTION

You can calculate the average agreement between corresponding pairs of sequences of two different files. The number and order of code variables should be the same, the number and order of sequences should be the same in both files. If the names of the code variables are different, you will get a warning, but you can continue the analysis, if you want.

This analysis can be an alternative for the reliability analysis (kappa), in case two observers coded the same sequence, but came up with a different number of events. The agreement between two corresponding sequences is an indication of the reliability of the coding procedure, including dividing the sequence into events (see also V. Quera, R. Bakeman and A. Gnisci, "Observer agreement for event sequences: Methods and software for sequence alignment and reliability estimates." Behavior Research Methods, 2007, 39 (1), 39-49).

Moreover you can assign, if you wish, the agreement between both sequences to a sequence variable in the open file. In this way it is easy to locate sequences that are quite differently coded by both coders (they will have a low agreement).

In addition to the average agreement between the sequence pairs, the program also calculates the weighted average: the agreement between the sequences of each pair is weighted by the length of both sequences in calculating the average agreement. Thus, longer sequences contribute more to the weighted average agreement.

SETUP

- Select **Codes > Compare sequences...** to open the COMPARE SEQUENCES setup window.
- Click Get file to select the file that should be compared with the open file.
- To assign the agreement of each pair to a sequence variable, check ☑ **Assign agreement to**, and enter or select a sequence variable.
- Select an appropriate agreement measure.
- Click OK.

If you use the select (marked) sequences option, the criteria for the selection are checked against the open file. For example, if according to the selection criteria sequence number 3 is skipped, this sequence is skipped in both files, irrespective of the values for the selection criteria in the other file.

See also:
Reliability, Define sequence, Define agreement, Select sequences, and Select marked sequences.

8 Event variables

Event variables are variables, assigned to events. For example, if sequences concern verbal interactions, and events concern the utterances of the speakers, you can add the onset time and offset time of each utterance to the event as event variables.
The onset and offset time of events are the most common types of event variables and always appear together. They are also called event times.
From the onset and offset times you can calculate additional event variables, like the duration of the utterance. Other examples of event variables may be the number of words of the utterance or the speech rate.

Event variables have numerical values, consisting of maximally 12 characters (including minus sign and decimal point). The number of event variables is unlimited (unless you exceed the maximum size of a data file, which is about 4Gb).

To select an event variable for analyses, you usually have a list available with event variables. You can just click at the name of an event variable to select it. The name will appear in an appropriate box. You can also type the name of an event variable in this box. As soon as you type one or more characters, a name is suggested. This is a handy feature if you have a lot of event variables.

In addition to options for editing and manipulating event variables, also some statistical analyses are available on event variables.

See also:
General introduction.

8.1 Define

INTRODUCTION

You can change the name of an event variable, add a label to it, and add an unlimited number of value labels. In addition you can define the variable type, change the number of decimal digits, and define minimum and maximum values (see also Numerical data).
Variable names can be maximally 12 characters long. Names are not case sensitive ('VAR1' is the same as 'var1'). Digits and spaces are allowed, but variable names should not consist of digits only (to distinguish them from values in computations). The first and the last character of a variable name cannot be a space.
You cannot change the name, the number of decimal digits and the label of the event time variables Onset time and Offset time. These event times are always whole numbers, with a minimum value of 0. But you can change their minimum (zero or more) and maximum values, and add value labels. The variable type of the event times can only be changed with the **Time** tab of the File settings (together with the variable type of the sequence variable STARTTIME).

See also:
Event variables.

SETUP

To change the properties of an event variable, do this:

- Select **Event variables > Define variables...** to open the DEFINE EVENT VARIABLES setup window.
- Click at the name of an event variable in the variables field to select it.
- To change the name of the selected event variable, type a new name in the edit box below the variable field and click the Change name button.
- To change the variable type, just select a different type with the pop-up menu. If you select 'date' or 'time stamp', the number of decimal digits can only be 0.
- If the variable type is 'number' you can change the number of decimal digits. This pop-up menu is disabled if the variable type is 'date' or 'time stamp'.
- Type minimum and maximum values for the variable. You can click at the  icon to obtain the minimum and maximum values that are allowed for event variables, given the variable type and the number of decimal digits. Or you can click at the  icon to obtain the actual minimum and maximum values (as appear in your file).
- To add or change a variable label, type the label in the 'variable label' edit box.
- To lock a variable, check the  **Lock variable** checkbox. A locked variable cannot be edited, changed (e.g. by replace or compute) or deleted, until it is unlocked.
- To change a value label, click at a value (or its label), type a new label and click Change.
- To delete a value label, click at a value or value label and click Delete.
- To add a value label, click below the last value, type a value and label and click Add.
- Click Revert to set all properties of the selected variable back to its original values.
- Click Cancel to get rid of all changes.
- Click Store to store all changes.

EFFECTS OF CHANGING PROPERTIES

Changing the variable type or the number of decimal digits may affect the minimum and maximum values allowed by SEQUENCE VIEWER, or may change the value because of rounding, as explained below. After you click the Store button, such values are automatically adjusted. If one or more values are actually adjusted, you are informed in the output window. In such a case, you can undo the 'Define event variables' command (select the appropriate menu item of the **Edit** menu), and inspect the impact of the changes first.

Changing only the user defined minimum and maximum (with the 'Min' and 'Max' edit boxes) will not affect the data; if a value exceeds the maximum as defined by the user, but not the allowed maximum, it remains unchanged. See also minimum and maximum values.

Variable type

If you change the variable type of a variable, this affects the allowed minimum and maximum values. For example if you change a variable from 'number' to 'date', it is very well possible that a particular value now exceeds the maximum allowed for a date variable. Values exceeding the allowed minimum and maximum are assigned 'M' for 'missing value'.

Decimal digits

If you change the number of decimal digits, all values will be rounded according to the new number of decimal digits. This may cause the value to be changed. For example, if 2.537 is rounded to two decimal digits, the value becomes 2.64; a value like 2.300 does not change of course if rounded to two decimal digits.

If the number of decimal digits is increased, this will not affect the values. However, now it may occur that a value exceeds the allowed minimum or maximum, because the allowed maximum

and minimum depend on the number of decimal digits. In that case the value obtain 'M' for missing value.

8.2 Info

To get information about an event variable:

- Select **Event variables > Show info...** to open the INFO OF EVENT VARIABLES setup window.
- Click at a variable name to select it. The name is added to the list below 'Show info of'. To delete a variable from the list, click at the variable once again.
- Check ☑ **Show properties** to obtain the properties (see Define event variables) for each variable.
- Check ☑ **Show statistics** to obtain statistics (like mean and variance) for each variable.
- Click OK.

See also
Select marked sequences and Checked events only.

8.3 Add

To add a new event variable:

- Select **Event variables > Add variable...** to open the ADD EVENT VARIABLE setup window.
- Select the variable type
- Select the number of decimal digits.
- Type a minimum value for the new variable.
- Type a maximum value for the new variable.
- Type a value that will be assigned to all events.
- Click Add.

The number of decimal digits and the minimum and maximum values are used by the program to prevent the input of values that are out of range. Moreover, these values are used to determine the maximum possible length of the values of the event variable. This in turn is used for the layout, when values of event variables are shown in the main sequence window. For these reasons, don't set these values higher than necessary.

See also:
Define event variables.

8.4 Delete

To delete an event variable:

- Select **Event variables > Delete variable...** to open the DELETE EVENT VARIABLE setup window.
- Select or enter the name of an event variable.
- Click Delete.

8.5 Rearrange

INTRODUCTION

Rearranging event variables can be useful, if you are doing a lot of analyses using a particular variable. Moving it upwards, prevents that you have to scroll to select this variable. Variables that are used seldom, can best be put at the end of the variables list. Rearranging may also be useful if you want to compare values on two event variables. Positioning them after each other, makes the comparison easier.

The positions of Onset time and Offset time cannot be changed.

SETUP

To rearrange the event variables:

- Select **Event variables > Rearrange...** to open the REARRANGE EVENT VARIABLES setup window.
- Click at the name of a variable in the variables field below *New order* to select it.
- Click at the  and  icons to move the selected variable upwards or downwards. You can also use the up and down arrow keys to this end.
- Select another variable to move it; etc.
- Click `Rearrange`.

8.6 Add event time

The event time actually consists of two event variables: the Onset time of the event and the Offset time of the event. You can only add (or delete) both variables at once.

To add the event time variables, select **Event variables > Add event time...**. Both variables are always added to the data file as the first two event variables, also if there are already other event variables. If you add the event time variables, all values are set to missing ('M').

See also:
Event times.

8.7 Delete event time

Use **Event variables > Delete event time...** to delete both event time variables. You obtain a warning first.

See also:
Event times and Add event time.

8.8 Finding

In addition to finding values on event variables, you also have options for replacing values, (un)marking sequences and (unchecking events. All these options make use of the same setup window, with tabs for **Find**, **Replace**, **Mark**, **Unmark**, **Check** and **Uncheck**.

To find a particular value, you can use one of the operators '=', '≠', '>', '<', '≥' or '≤' to indicate whether the value of the selected event variable should equal a particular value, or should be larger than a particular value, etcetera. You can compare the value either with a constant, or with the value on another event variable. For example, 'EVTVAR1 = EVTVAR2' means that an event is found if the value of EVTVAR1 equals the value of EVTVAR2 for that event. Or 'EVTVAR3 ≠ 4' finds events where the value for EVTVAR3 does not equal 4, including 'M'.

You can add a second event variable with similar options to the first event variable. To this end, select 'and', 'or', or 'xor' with the pop-up menu.
'and' means that both conditions should be fulfilled
'or' means that one or both conditions should be fulfilled
'xor' means that either one of both conditions should be fulfilled.
If you select 'none', only the condition for the first sequence variable is taken into account.

8.8.1 Find

To find a particular value of an event variable, do this:

- Select **Event variables > Find value of event variable...** to open the FIND VALUE OF EVENT VARIABLE setup window.
- Select the **Find** tab.
- Click at the 'Find if' edit box to select it and enter or select an event variable.
- Select '=', '≠', '>', '<', '≥' or '≤' with the pop-up menu.
- Either select 'value' and type a value, or 'variable' and enter or select an event variable.
- To select a second event variable, first select 'and', 'or' or 'xor' from the pop-up menu.
- Select 'Search up' to search from the present sequence upwards, 'Search down' to search from the present sequence downwards or 'Search all' to search upwards from sequence number 1 with the 'Search direction' pop-up menu.
- Click Find.

If an event is found, the value of the event variable is shown in the main sequence window. In the codes and text and codes only modes, the 'Show event variable' pop-up menu is set to the selected event variable, and the value with the corresponding event code is highlighted. In the codes and event variables mode, the found value is highlighted.

The Find button changes into Find next, allowing you to find the next instance.

See also
Finding, Select marked sequences and Checked events only.

8.8.2 Replace

To replace values of an event variable, do this:

- Select **Event variables > Find value of event variable...** to open the FIND VALUE OF EVENT VARIABLE setup window.
- Select the **Replace** tab.
- Specify the Find options as described in the Find section.
- Click at the 'Replace' edit box to select it and enter or select an event variable. The values of this variable will be replaced. This variable may or may not be the same as the 'find' variable.
- Either select 'value' and type a value, or 'variable' and enter or select an event variable in the 'with' edit box.
- Click Find. The button changes into Find next.
- Click Replace to change the value and to find the next instance.
- Click Replace all to change all values of all (marked) sequences at once.

If an event is found, the value that should be replaced is surrounded by a rectangle in the codes and event variables mode.

See also
Finding, Select marked sequences and Checked events only.

8.8.3 Mark/unmark

To mark or unmark sequences having a particular value on an event variable of one of the events of a sequence, do this:

- Select **Event variables > Find value of event variable...** to open the FIND VALUE OF EVENT VARIABLE setup window.
- Select the **Mark** or the **Unmark** tab.
- Specify the Find options as described in the Find section.
- Check ☑ **Unmark all sequences first**, or ☑ **Mark all sequences first**, if appropriate.
- Click Mark or Unmark.

See also:
Finding and Checked events only.

8.8.4 Check/uncheck

To check or uncheck events having a particular value on an event variable, do this:

- Select **Event variables > Find value of event variable...** to open the FIND VALUE OF EVENT VARIABLE setup window.
- Select the **Check** or the **Uncheck** tab.
- Specify the Find options as described in the Find section.
- Check ☑ **Uncheck all events first**, or ☑ **Check all events first**, if appropriate.
- Click Check or Uncheck.

See also:
Finding, Select marked sequences.

8.9 Compute

SETUP

To compute the values of an event variable:

- Select **Event variables > Compute event variable...** to open the COMPUTEF EVENT VARIABLE setup window.
- Select **Show event variable**.
- Click in the 'Compute variable' edit box and enter or select a sequence variable.
- Now click in the large equation field to select this field. Here you can enter the right part of the equation.
- You can click at a variable name or at one of the operators below the equation field to enter the text in the equation field.
- You can also enter the name of a sequence variable. In that case the value of the sequence variable will be used for all events in the sequence. Select Show sequence variables to obtain the names of the sequence variables. To enter the names of event variables, you can select **Show event variables**.
- Values are rounded according to the number of decimal digits of the event variable to be calculated.
- If a value cannot be calculated, a missing value ('M') is assigned. Dividing by zero also yields a missing value.
- If there are errors in the equation, the Compute button remains disabled. Move the mouse over the Help icon and the help field of the toolbar tells you what's wrong. Usually the part of the equation where the error is detected, will become red.
- Click Compute.

If there are no errors in the equation, SEQUENCE VIEWER displays the value for the first event of the present sequence at the bottom of the window, using six decimal digits.

If the variable to be computed is a date or a time stamp variable, you cannot use a floating point number, a floating point variable or a divisor ('/') in the equation.

INDEXING EVENT VARIABLES

You can index an event variable. For example, suppose you have an event variable called EVTVAR1. Then EVTVAR1[2] means the value of EVTVAR1 on the second event in the sequence.

Example 1:
EVTVAR2=EVTVAR3+EVTVAR1[2]
adds the value of event variable EVTVAR1 on the second event to the value of the event variable EVTVAR3 for each event and assigns this sum to event variable EVTVAR2.

Event variables can also be indexed with the keywords 'LAST' and 'EVTNO'. 'LAST' refers to the last event in a sequence. EVTNO+1 refers to the next event (EVTVAR1[EVTNO] is the same as just EVTVAR1, and thus doesn't make sense).

Example 2:
pause=onset time-offset time[EVTNO-1]
assigns the duration between the end of the previous event and the start of the present one to the event variable 'pause'. The first event will obtain a missing value.

Example 3:
remaining=offset time[LAST]-offset time
is the remaining time after a particular event until the end of the sequence.

APPLYING A SEQUENCE DEFINITION

You can apply a sequence definition. Values will only be assigned to events that are retained in the converted sequence. If the sequence definition unites events, the value will be assigned to the first of the original (not united) events. Or if events are (temporarily) removed, because of a sequence definition, the values of the event variable that is computed, will not be changed for the removed events. In this way, it is for example possible to compute the value of an event variable, only for events with a particular event code; or only for checked events, etcetera.

If you apply a sequence definition, indexed event variables refer to the converted sequence. For example EVTVAR[2] is the second event of the converted sequence, which is not necessarily the second event of the original sequence.

In case the sequence variable SEQSIZE is used in the equation, you can check ☑ Use converted SEQSIZE to use the number of events of the converted sequence as value for SEQSIZE. If this checkbox is not checked, the number of events of the original sequence will be used.

See also:
Compute sequence variables, Define sequence, Select sequences and Select marked sequences.

8.10 Transform

SETUP

You can transform an event variable according to some elementary functions and assign the result to another (or the same) event variable.

- Select **Event variables > Transform...** to open the TRANSFORM EVENT VARIABLE setup window.
- Select a transformation from the pop-up menu.
- Click at the 'Transform variable' edit box and enter or select the variable to which the function will be applied. If the 'random' function is selected, there will be no transform variable.
- Click at the 'Assign variable' edit box and enter or select the variable to which the result will be assigned. This may be the same variable as the transform variable.
- Depending on the selected transformation, other fields may appear. Enter appropriate values.
- Click Transform.

TRANSFORMATIONS

The following functions are available.

- ln
Returns the natural logarithm of the value.

- log10
Returns the logarithm with base 10.

- *exp*
Returns e to the power of the value.

- *pwr*
Returns the value to the power of a specified number. Enter this number in the field below 'raise to power'.

- *sqrt*
Returns the square root of the value.

- *cosine*
Returns the cosine of the value. Values are in radians. To convert radians to degrees, use the degree function.

- *sine*
Returns the sine of the value. Values are in radians. To convert radians to degrees, use the degree function.

- *tangent*
Returns the tangent of the value. Values are in radians. To convert radians to degrees, use the degree function.

- *degree*
Converts radians to degrees.

- *hour*
Returns the hour in 24-hour time from the value of a time stamp variable; for example 1329984720 ('Feb 23, 2012 20:12:00') returns '20'.

- *weekday*
Returns the number of a day (Sunday is 1, Monday is 2, etcetera) from the value of a date or time stamp variable.

- *month*
Returns the number of the month (January is 1, February is 2, etcetera) from the value of a date or time stamp variable.

- *year*
Returns the year from the value of a date or time stamp variable. For example 13578 (March 6, 2007 in case of a date variable) returns 2007.

- *date*
Returns the number of days from a time stamp variable. Be sure to change the variable type of the variable to which the result is assigned, to 'date'. See also Numerical data.

- *time stamp*
Returns the number of seconds from a date variable. Be sure to change the variable type of the variable to which the result is assigned, to 'time stamp'. See also Numerical data.

- *abs*
Returns the absolute value.

- random
Returns a random value between the minimum and the maximum. Enter minimum and maximum in their respective fields.

- maximum, minimum
Returns the maximum, resp. minimum value of two or more event variables of a sequence.

Transformed values will be rounded according to the number of decimal digits of the variable to which the transformation is assigned. Values exceeding the allowed minimum and maximum values, obtain 'M' for missing value. For example, the maximum of a floating point variable with 3 decimal digits = 99999999.999; if a transformation exceeds this value, a 'M' is assigned.
A number of transformations may not yield meaningful values. For example, transforming a number variable (in particular a floating point number) to the number of the month, cannot be performed, because the program should know whether the values to be transformed represent seconds (in case of a time stamp variable), or days (in case of a date variable). Or assigning a cosine transformation to a date variable, will usually not result in a meaningful date (that is, the number of days since Jan 1, 1970). In such cases the 'transform' and/or the 'assign to' variables are disabled in the variables list (and the Transform button will be disabled). If you type such a variable nevertheless, it will become red. You can move the mouse over the Help icon; in the help field you can read why the transformation cannot be performed.

See also:
Select sequences, Select marked sequences, and Checked events only.

8.11 Convert to code variable

You can change the code of a code variable of an event, given the value(s) of an event variable. For example, you can change the code of code variable 'CodeVar1' to code 'B', for all events having a value between 3 and 6 on event variable 'EvtVar3'.

- Select **Event variables > Convert to code variable...** to open the CONVERT EVENT VARIABLE TO CODE VARIABLE setup window.
- Enter or select an event variable after 'If'. The actual minimum and maximum values as appear in your file will be shown.
- Select a comparison operator with the pull down menu, and enter a value.
- To specify a range of values (like from 3 to 8), check **and [variable name]**.
- Select a code variable with the 'change code variable' pop-up menu. Select a code with the pull down menu.
- Click Convert.

See also:
Select sequences, Select marked sequences, and Checked events only.

8.12 Assign event number

You can assign the number of each event in a sequence to an event variable.

- Select **Event variables > Assign event number...** to open the ASSIGN EVENT NUMBER setup window.

- Select an event variable from the list.
- Check ☑ Include locked sequences, if the event number should also be assigned in case of locked sequences.
- Click Assign.

8.13 Frequencies

To obtain a frequency distribution of an event variable, do this:

- Select **Event variables > Frequencies...** to open the FREQUENCIES OF EVENT VARIABLE setup window.
- Click at the 'Show frequencies of' edit box and enter or select an event variable.
- Click OK.

See also:
Define sequence, Select sequences and Select marked sequences.

8.14 Means

SETUP

You can calculate and compare the means of an event variable for different combinations of event codes.

- Select **Event variables > Means...** to open the MEAN OF EVENT VARIABLE setup window.
- Enter or select an event variable.
- Enter an event code in the 'by event code' edit box. If you type a question mark for a particular code variable, the mean is calculated for each code on that code variable.
- Click OK.

ADDITIONAL INFORMATION

The program performs an analysis of variance on the means for the different groups of event codes.

Usually you should use this analysis in combination with a sequence definition. Suppose you want to compare the means for all different codes of only one code variable. In that case, you use a sequence definition, masking all other code variables, and typing a question mark for the one unmasked code variable. For example, suppose you have five code variables, and the first code variable is either 'I' or 'R'. You want to compare the means of an event variable for event codes with 'I' at the first position and 'R' at the first position. The event code in the 'by event code' edit box should look line '?•••' after applying the sequence definition.

If you just want the mean of an event variable for all event codes, use **Event variables > Show info...**.

See also:
Define sequence, Select sequences and Select marked sequences, Show info.

8.15 ANOVA

INTRODUCTION

You can perform an analysis of variance with an event variable as dependent variable, and a sequence variable as independent variable. Because a sequence usually has a number of different events, and hence a number of different values on the selected dependent variable, the different sequences are nested under the (different values of the) independent variable. Hence you essentially have a nested design, with 'sequence number' (a temporary variable, called SEQNO) as an additional variable. This variable is automatically incorporated in the analysis and the design will be a nested one, with SEQNO (designating different sequences) nested under the independent variable. If you happen to have a sequence variable 'SEQNO' in your file, SEQUENCE VIEWER will not confuse both variables.

The most appropriate error term for the effect of the independent variable is the sum of squares of the nesting variable (because SEQNO usually can be viewed as a random factor). If the different sequences can be viewed as a fixed factor, select 'within cells' as error term. If the means of the dependent variable do not differ significantly from each other, you can use the pooled sum of squares as error term, to increase the power of the test.

SETUP

- Select **Event variables > ANOVA...** to open the ANOVA setup window.
- Click at the 'dependent (event) variable' edit box to select it, and enter or select an event variable.
- Click at the 'by (sequence) variable' edit box. The variables list now shows the sequence variables. Enter or select a sequence variable as independent variable.
- Select an appropriate error term (see above).
- Click OK.

You may use a sequence definition to unite or remove events, define a subsequence, or apply a time interval. The analysis will be performed on the transformed sequences. For example, if you unite events, the united events may have adjusted values on the event variables (depending on which option is selected in the SEQUENCE DEFINITION setup window after 'Unite events').

See also:
ANOVA (for sequence variables), Define sequence, Select sequences and Select marked sequences.

9 Sequence variables

Sequence variables are variables that belong to a sequence, that is, they describe characteristics of a sequence. Sequence variables have numerical values, consisting of maximally 12 characters (including minus sign and decimal point). The number of sequence variables is unlimited (unless you exceed the maximum size of a data file, which is about 4Gb).

To show the values of the sequence variables of the presently open sequence, you can click at the  icon before 'vars' in the lower left corner of the main sequence window. A drawer appears with the sequence variables and their values.

The first two sequence variables are always SEQSIZE and STARTTIME. SEQSIZE is simply the number of events in a sequence. The value of SEQSIZE can only be changed by adding to or deleting events from a sequence. STARTTIME is the time the sequence starts, usually referring to the linked audio or video file (see also Audio and video).

SEQUENCE VIEWER offers a number of options to define and manipulate sequence variables. In addition a number of common statistical analyses on sequence variables are available. For more sophisticated analyses however, you should export the sequence variables.

To select a sequence variable for analyses, you usually have a list available with sequence variables. You can just click at the name of a sequence variable to select it. The name will appear in an appropriate edit box. You can also type the name of a sequence variable in this edit box. As soon as you type one or more characters, a name is suggested. This is a handy feature if you have a lot of sequence variables.

See also:
General introduction.

9.1 Define

INTRODUCTION

You can change the name of a sequence variable, add a label to it, and add an unlimited number of value labels. In addition you can define the variable type, change the number of decimal digits, and define minimum and maximum values (see also Numerical data).

Variable names can be maximally 12 characters long. Names are not case sensitive ('VAR1' is the same as 'var1'). Digits and spaces are allowed, but variable names should not consist of digits only (to distinguish them from values in computations). The first and the last character of a variable name cannot be a space.

You cannot change the name, the number of decimal digits and the label of the variables SEQSIZE and STARTTIME. SEQSIZE is always a whole number with a minimum of 0. The variable type of STARTTIME can only be changed with the 'Variable type' pop-up menu in the **Time** tab of the File settings (together with the variable type of the event time variables Onset time and Offset time).

See also:

Sequence variables.

SETUP

To change the properties of a sequence variable, do this:

- Select **Sequence variables > Define variables...** to open the DEFINE SEQUENCE VARIABLES setup window.
- Click at the name of a sequence variable in the variables field to select it.
- To change the name of the selected sequence variable, type a new name in the edit box below the variable field and click the Change name button.
- To change the variable type, just select a different type with the pop-up menu. If you select 'date' or 'time stamp', the number of decimal digits can only be 0.
- If the variable type is 'number' you can change the number of decimal digits. This pop-up menu is disabled if the variable type is 'date' or 'time stamp'.
- Type minimum and maximum values for the variable. You can click at the  icon to obtain the minimum and maximum values that are allowed for sequence variables, given the variable type and the number of decimal digits. Or you can click at the  icon to obtain the actual minimum and maximum values (as appear in your file).
- To add or change a variable label, type the label in the 'variable label' edit box.
- To lock a variable, check the  **Lock variable** checkbox. A locked variable cannot be edited, changed (e.g. by replace or compute) or deleted, until it is unlocked.
- To change a value label, click at a value (or its label), type a new label and click Change.
- To delete a value label, click at a value or value label and click Delete.
- To add a value label, click below the last value, type a value and label and click Add.
- Click Revert to set all properties of the selected variable back to its original values.
- Click Cancel to get rid of all changes.
- Click Store to store all changes.

EFFECTS OF CHANGING PROPERTIES

Changing the variable type or the number of decimal digits may affect the minimum and maximum values allowed by SEQUENCE VIEWER, or may change the value because of rounding, as explained below. After you click the Store button, such values are automatically adjusted. If one or more values are actually adjusted, you are informed in the output window. In such a case, you can undo the 'Define sequence variables' command (select the appropriate menu item of the **Edit** menu), and inspect the impact of the changes first.

Changing only the user defined minimum and maximum (with the 'Min' and 'Max' edit boxes) will not affect the data; if a value exceeds the maximum as defined by the user, but not the allowed maximum, it remains unchanged. See also minimum and maximum values.

Variable type

If you change the variable type of a variable, this affects the allowed minimum and maximum values. For example if you change a variable from 'number' to 'date', it is very well possible that a particular value now exceeds the maximum allowed for a date variable. Values exceeding the allowed minimum and maximum are assigned 'M' for 'missing value'.

Decimal digits

If you change the number of decimal digits, all values will be rounded according to the new number of decimal digits. This may cause the value to be changed. For example, if 2.537 is rounded to two decimal digits, the value becomes 2.64; a value like 2.300 does not change of

course if rounded to two decimal digits.

If the number of decimal digits is increased, this will not affect the values. However, now it may occur that a value exceeds the allowed minimum or maximum, because the allowed maximum and minimum depend on the number of decimal digits. In that case the value obtain 'M' for missing value.

9.2 Info

To get information about a sequence variable:

- Select **Sequence variables > Show info...** to open the INFO OF SEQUENCE VARIABLES setup window.
- Click at a variable name to select it. The name is added to the list below 'Show info of'. To delete a variable from the list, click at the variable once again.
- Check ☑ **Show properties** to obtain the properties (see Define sequence variables) for each variable.
- Check ☑ **Show statistics** to obtain statistics (like mean and variance) for each variable.
- Click OK.

See also:
Select marked sequences.

9.3 Add

To add a new sequence variable:

- Select **Sequence variables > Add variable...** to open the DELETE SEQUENCE VARIABLE setup window.
- Select the variable type
- Select the number of decimal digits.
- Type a minimum value for the new variable.
- Type a maximum value for the new variable.
- Type a value that will be assigned to all sequences.
- Click Add.

The number of decimal digits and the minimum and maximum values are used by the program to prevent the input of values that are out of range. For this reason, don't set these values higher than necessary.

See also:
Define sequence variables

9.4 Delete

To delete a sequence variable:

- Select **Sequence variables > Delete variable...** to open the DELETE SEQUENCE VARIABLE setup window.
- Select the name of the variable to be deleted from the variables field.

• Click Delete.

9.5 Rearrange

INTRODUCTION

Rearranging sequence variables can be useful, if you are doing a lot of analyses using a particular variable. Moving it upwards, prevents that you have to scroll to select this variable. Variables that are used seldom, can best be put at the end of the variables list.

The positions of SEQSIZE and STARTTIME cannot be changed.

SETUP

To rearrange the sequence variables:

• Select **Sequence variables > Rearrange...** to open the REARRANGE SEQUENCE VARIABLES setup window.
• Click at the name of a variable in the variables field below *New order* to select it.
• Click at the  and  icons to move the selected variable upwards or downwards. You can also use the up and down arrow keys to this end.
• Select another variable to move it; etc.
• Click Rearrange.

9.6 Finding

In addition to finding values on sequence variables, you also have options for replacing values and (un)marking sequences having a particular value on a sequence variable. All these options make use of the same setup window, with tabs for **Find**, **Replace**, **Mark** and **Unmark**.

To find a particular value, you can use one of the operators '=', '≠', '>', '<', '≥' or '≤' to indicate whether the value of the selected sequence variable should equal a particular value, or should be larger than a particular value, etcetera. You can compare the value either with a constant, or with the value on another sequence variable. For example, 'SEQVAR1 = SEQVAR2' means that a sequence is found if the value of SEQVAR1 equals the value of SEQVAR2 for that sequence. Or 'SEQVAR3 ≠ 4' finds sequences where the value for SEQVAR3 does not equal 4, including 'M'.

If you selected 'value' and the selected sequence variable has value labels, you can, instead of typing a numerical value, also select a value label with the pull down menu.

You can add a second sequence variable with similar options to the first sequence variable. To this end, select 'and', 'or', or 'xor' with the pop-up menu.
'and' means that both conditions should be fulfilled
'or' means that one or both conditions should be fulfilled
'xor' means that either one of both conditions should be fulfilled.
If you select 'none', only the condition for the first sequence variable is taken into account.

9.6.1 Find

To find a particular value of a sequence variable, do this:

- Select **Sequence variables > Find value of sequence variable...** to open the FIND VALUE OF SEQUENCE VARIABLE setup window.
- Select the **Find** tab.
- Click at the 'Find if' edit box to select it and enter or select a sequence variable.
- Select '=', '≠', '>', '<', '≥' or '≤' with the pop-up menu.
- Either select 'value' and type a value, or 'variable' and enter or select a sequence variable.
- To select a second sequence variable, first select 'and', 'or' or 'xor' from the pop-up menu.
- Select 'Search up' to search from the present sequence upwards, 'Search down' to search from the present sequence downwards or 'Search all' to search upwards from sequence number 1 with the 'Search direction' pop-up menu.
- Click Find.

If a sequence is found, the name(s) of the sequence variable(s) are highlighted in the sequence variables drawer. The Find button changes into Find next, allowing you to find the next instance.

See also
Finding and Select marked sequences.

9.6.2 Replace

To replace values of a sequence variable, do this:

- Select **Sequence variables > Find value of sequence variable...** to open the FIND VALUE OF SEQUENCE VARIABLE setup window.
- Select the **Replace** tab.
- Specify the Find options as described in the Find section.
- Click at the 'Replace' edit box to select it and enter or select a sequence variable. The values of this variable will be replaced. This variable may or may not be the same as the 'find' variable.
- Either select 'value' and type a value, or 'variable' and enter or select a sequence variable in the 'with' edit box.
- Click Find. The button changes into Find next.
- Click Replace to change the value and to find the next instance.
- Click Replace all to change all values of all (marked) sequences at once.

If a sequence is found, the name(s) of the sequence variable(s) are highlighted in the sequence variables drawer. The value that should be replaced is surrounded by a rectangle.

See also:
Finding and Select marked sequences.

9.6.3 Mark/unmark

To mark or unmark sequences having a particular value on a sequence variable, do this:

- Select **Sequence variables > Find value of sequence variable...** to open the FIND VALUE OF SEQUENCE VARIABLE setup window.
- Select the **Mark** or the **Unmark** tab.
- Specify the Find options as described in the Find section.
- Check **Unmark all sequences first**, or **Mark all sequences first**, if appropriate.
- Click Mark or Unmark.

See also:
Finding.

9.7 Compute

SETUP

To compute the values of a sequence variable:

- Select **Sequence variables > Compute sequence variable...** to open the COMPUTE SEQUENCE VARIABLE setup window.
- Select **Show sequence variable**.
- Click in the 'Compute variable' edit box and enter or select a sequence variable.
- Now click in the large equation field to select this field. Here you can enter the right part of the equation.
- You can click at a variable name or at one of the operators below the equation field to enter the text in the equation field.
- You can also enter indexed event variables (see below for an explanation). To enter the names of event variables, you can select **Show event variables**.
- Values are rounded according to the number of decimal digits of the sequence variable to be calculated.
- If a value cannot be calculated, a missing value ('M') is assigned. Dividing by zero also yields a missing value.
- If there are errors in the equation, the Compute button remains disabled. Move the mouse over the Help icon and the help field of the toolbar tells you what's wrong. Usually the part of the equation where the error is detected, will become red.
- Click Compute.

If there are no errors in the equation, SEQUENCE VIEWER displays the value for the present sequence after the equals sign.
If the variable to be computed is a date or a time stamp variable, you cannot use a floating point number, a floating point variable or a divisor ('/') in the equation.

INDEXING EVENT VARIABLES

You can enter the indexed event variables. The index refers to the number of the event in the sequence. The value of the event variable on this event will be used in the equation.

Example 1:
SEQVAR1=SEQVAR2+EVTVAR1[2]
adds the value of the event variable EVTVAR1 on the second event of the sequence, to the value of the sequence variable SEQVAR2 and assigns this sum to sequence variable SEQVAR1.

You can also index an event variable using the keyword LAST. This refers to the last event in the sequence. An index like [LAST-1] refers to the last but one event.

Example 2:
SEQVAR2=SEQVAR3+EVTVAR1[LAST-1]
adds the value of event variable EVTVAR1 on the one but last event to the value of the sequence variable SEQVAR3 and assigns this sum to sequence variable SEQVAR2.

Example 3:
SEQVAR2=EVTVAR1[LAST-5]
cannot be calculated if the sequence has 5 events or less. In that case a missing value is assigned.
You may also note that [LAST+1] does not make sense.

Example 4:
duration=offset time[LAST]-onset time[1]
assigns the duration of the sequence to sequence variable 'duration.

ADDING THE SUM OR MEAN OF EVENT VARIABLES TO THE EQUATION

You can use the name of an event variable followed by ':SUM' or ':MEAN', to use the sum, respectively mean of an event variable of all events in the sequence in the computation.

Example 5:
SEQVAR1=EVTVAR2:MEAN
assigns the mean of all values of event variable EVTVAR2 in a sequence to sequence variable SEQVAR1.

INDEXING SEQUENCE VARIABLES

You can index a sequence variable with the sequence number. The index refers to the number of a sequence.

Example 6:
SEQVAR1=STARTIME[1]+SEQVAR1
adds the STARTTIME of the first sequence to the value of SEQVAR1.

Example 7:
SEQVAR1=STARTTIME[SEQNO+1]-STARTTIME
The expression 'STARTTIME[SEQNO+1]' takes the value of STARTTIME of the next sequence. If sequences themselves are ordered with respect to the time, 'STARTTIME[SEQNO+1]-STARTTIME' is the duration of the sequence (including any pause between sequences).

APPLYING A SEQUENCE DEFINITION

You can apply a sequence definition. In that case, indexed event variables refer to the converted sequence. For example EVTVAR[2] is the second event of the converted sequence, which is not necessarily the second event of the original sequence.
For example, you can use a sequence definition to assign the sum of the values of an event variable to a sequence variable, for the checked events only.

If you apply a sequence definition and use the sequence variable SEQSIZE in the equation, you can check ☑ **Use converted SEQSIZE** to use the number of events of the converted sequence (which

may be less than the original sequence) as value for SEQSIZE. If this checkbox is not checked, the number of events of the original sequence will be used. The checkbox is disabled, if the variable SEQSIZE (without index) is not used in the equation.

The expression SEQSIZE[SEQNO+1] is the length of the next sequence. In case you use a sequence definition, this next sequence is not converted, and thus equals the length of the original sequence.

NOTE

Using STARTTIME and the event times in an equation, will yield different results for absolute and relative event times. For example, suppose you want to know the duration between the last offset time of a sequence, and the start of the next sequences. In case of absolute event times:

duration = STARTTIME[SEQNO+1]-offset time[LAST]

gives you the answer. This is not the case if event times are relative however. Usually this can be solved by adding STARTTIME to the equation:

STARTIME[SEQNO+1]-offset time[LAST]+STARTTIME.

See also:
Compute event variables, Define sequence, Select sequences and Select marked sequences.

9.8 Transform

SETUP

You can transform a sequence variable according to some elementary functions and assign the result to another (or the same) sequence variable.

- Select **Sequence variables > Transform...** to open the TRANSFORM SEQUENCE VARIABLE setup window.
- Select a transformation from the pop-up menu.
- Click at the 'Transform variable' edit box and enter or select the variable to which the function will be applied. If the 'random' function is selected, there will be no transform variable.
- Click at the 'Assign variable' edit box and enter or select the variable to which the result will be assigned. This may be the same variable as the transform variable.
- Depending on the selected transformation, other fields may appear. Enter appropriate values.
- Click Transform.

TRANSFORMATIONS

The following functions are available.

- ln
Returns the natural logarithm of the value.

- log10
Returns the logarithm with base 10.

- *exp*
Returns e to the power of the value.

- *pwr*
Returns the value to the power of a specified number. Enter this number in the field below 'raise to power'.

- *sqrt*
Returns the square root of the value.

- *cosine*
Returns the cosine of the value. Values are in radians. To convert radians to degrees, use the degree function.

- *sine*
Returns the sine of the value. Values are in radians. To convert radians to degrees, use the degree function.

- *tangent*
Returns the tangent of the value. Values are in radians. To convert radians to degrees, use the degree function.

- *degree*
Converts radians to degrees.

- *hour*
Returns the hour in 24-hour time from the value of a time stamp variable; for example 1329984720 ('Feb 23, 2012 20:12:00') returns '20'.

- *weekday*
Returns the number of a day (Sunday is 1, Monday is 2, etcetera) from the value of a date or time stamp variable.

- *month*
Returns the number of the month (January is 1, February is 2, etcetera) from the value of a date or time stamp variable.

- *year*
Returns the year from the value of a date or time stamp variable. For example 13578 (March 6, 2007 in case of a date variable) returns 2007.

- *date*
Returns the number of days from a time stamp variable. Be sure to change the variable type of the variable to which the result is assigned, to 'date'. See also Numerical data.

- *time stamp*
Returns the number of seconds from a date variable. Be sure to change the variable type of the variable to which the result is assigned, to 'time stamp'. See also Numerical data.

- *abs*
Returns the absolute value.

- random
Returns a random value between the minimum and the maximum. Enter minimum and maximum in their respective fields.

- maximum, minimum
Returns the maximum, resp. minimum value of two or more sequence variables of a sequence.

Transformed values will be rounded according to the number of decimal digits of the variable to which the transformation is assigned. Values exceeding the allowed minimum and maximum values, obtain 'M' for missing value. For example, the maximum of a floating point variable with 3 decimal digits = 99999999.999; if a transformation exceeds this value, a 'M' is assigned.
A number of transformations may not yield meaningful values. For example, transforming a number variable (in particular a floating point number) to the number of the month, cannot be performed, because the program should know whether the values to be transformed represent seconds (in case of a time stamp variable), or days (in case of a date variable). Or assigning a cosine transformation to a date variable, will usually not result in a meaningful date (that is, the number of days since Jan 1, 1970). In such cases the 'transform' and/or the 'assign to' variables are disabled in the variables list (and the Transform button will be disabled). If you type such a variable nevertheless, it will become red. You can move the mouse over the Help icon; in the help field you can read why the transformation cannot be performed.

See also:
Select sequences, Select marked sequences.

9.9 Assign sequence number

You can assign the (order) number of the sequence in your file to a sequence variable.

- Select **Sequence variables > Assign sequence number...** to open the ASSIGN SEQUENCE NUMBER setup window.
- Click at the 'Assign to variable' edit box and enter or select a sequence variable.
- Click Assign.

If you want to assign the sequence number also to locked sequences,check **include locked sequences**.

9.10 Assign number of checks

You can assign the number of checked events of a sequence to a sequence variable.

- Select **Sequence variables > Assign number of checks...** to open the ASSIGN NUMBER OF CHECKS setup window.
- Click at the 'Assign to variable' edit box and enter or select a sequence variable.
- Click Assign.

9.11 Frequencies

SETUP

To obtain a frequency distribution of a sequence variable, do this:

- Select **Sequence variables > Frequencies...** to open the FREQUENCIES OF SEQUENCE VARIABLE setup window.
- Click at the 'Show frequencies of' edit box and enter or select a sequence variable.
- Check ☑ **Perform analysis for each value of:** and enter a variable name after this checkbox, to obtain a frequency distribution for each value on this sequence variable.
- Click OK.

HISTOGRAM

You can also obtain a histogram of the selected sequence variable. Just select a variable, apply any selection criteria and click the Show histogram button. A sheet appears with the histogram.

  - You can vary the number of intervals between 2 and 9 (provided that there are two or more values on the sequence variable) with the 'number of intervals' pull down menu.
  - You can also specify a lower and an upper bound. Enter these values, or select a value with the pull down menus. Click the Show button. If the entered value exceeds the values in the data file, the Show button will be disabled.
  - Select a different color for the bars with the 'Bar colors' button.
  - If the number of different values exceeds 9, you can use ☑ **Don't group** checkbox, to either group the values with equal intervals, or don't group the values. If the values are not grouped, the lowest value is the one that is entered in the 'lower bound' edit box.
  - Check ☑ **Show grid**, to show gray horizontal lines in the graph.

See also:
Select sequences and Select marked sequences.

9.12 Crosstable

To obtain a crosstable of two sequence variables, do this:

- Select **Sequence variables > Crosstable...** to open the CROSSTABLE OF SEQUENCE VARIABLES setup window.
- Click at the 'Column variable' edit box and enter or select the column variable.
- Click at the 'Row variable' edit box and enter or select the row variable.
- Select 'frequencies', 'percentages' or 'both' as table output with the 'show' pop-up menu.
- You may check ☑ **Include missing values**, to include the sequences having a missing value on one or both sequence variables into the table.
- Check ☑ **Perform analysis for each value of:** and enter a variable name after this checkbox, if you want a crosstable for each value on this sequence variable.
- Click OK.

See also:
Select sequences and Select marked sequences.

9.13 t-test

To perform a t-test on sequence variables, do this:

- Select **Sequence variables > T-test...** to open the T-TEST setup window.
- Click at the edit box below 'Dependent variable' and select or enter the dependent variable.
- Click at the edit box below 'Group variable' and enter or select the group variable.
- In both value fields the minimum and the maximum values of the group variable appear. Of course you can type different values if you want. If the values have labels, you can also select appropriate labels with the pull down menu 'label'. The corresponding values are automatically inserted.
- if you want a t-test for each value on a sequence variable, (e.g. the effect of respondent gender on SEQSIZE for each interviewer) check ☑ **Perform analysis for each value of:** and enter a variable name after this checkbox.
- Click OK.

See also:
Select sequences and Select marked sequences.

9.14 Oneway

To perform a one-way analysis of variance on sequence variables, do this:

- Select **Sequence variables > Oneway...** to open the ONEWAY setup window.
- Click at the edit box below 'Dependent variable' and select or enter the dependent variable.
- Click at the field below 'Group variable' and select or enter the group variable.
- Click OK.

See also:
Select sequences and Select marked sequences.

9.15 ANOVA

INTRODUCTION

You can perform a two-way analysis of variance, with two independent variables. In addition you can add a covariate.

You have a number of design options, depending on the relation between both independent variables or factors. Moreover you have the choice between a number of different error terms to test the effects. For a more detailed description of all these options, we refer to handbooks about analysis of variance, e.g. B. J. Winer, D. R. Brown & K. M. Michels (1991), Statistical principles in experimental design, New York: MacGraw-Hill. Here we will provide only some elementary information.

An orthogonal design assumes that both factors are unrelated and that there are no missing values. This is usually only the case in a real experimental design, with equal numbers of cases in each cell, without missing values. If both factors are unrelated, but there are missing values, this may cause an artificial dependency between both independent variables. To correct for this

situation, an unweighted means analysis is the best solution. In an unweighted means analysis the cell frequencies are treated as if they are all equal; the cell means are not weighted by cell frequencies.

This situation should be distinguished from the situation that both factors are really related (that is, this relation reflects the relation in the population; for example years of education and gender or race). This means that part of the variance in the dependent variable that is explained by one factor, is also explained by the other factor. According to the least squares solution, first the variance that is explained by the first factor is calculated (which equals the variance explained by this factor in a one-way analysis of variance). Next the variance that is explained by the second factor, excluding the part that is already explained by the first one, is calculated. Hence the order of both factors is relevant. This solution is in SPSS known as 'Type I Sum of Squares' ('Type II Sum of Squares' in SPSS calculates the variance explained by each factor, excluding the contribution of the other factor; by switching the order of both factors with the switch button, it is easy to obtain this solution in SEQUENCE VIEWER). In the weighted means analysis (equivalent to 'Type III Sum of Squares' in SPSS) the means are weighted according to the cell frequencies. This is the most common design for non-experimental data. If you select a 'weighted' or 'least squares' solution, the number of levels for each factor should not exceed 30. The OK button will be disabled of one of the factors has a higher number of levels in your datafile.

Factors may be related in the sense that particular levels of one factor only occur within one particular level of the other factor. Such a design is called a nested design. An example is interviewer number and respondent number. A particular respondent number does only occur for one interviewer number (an interviewer may interrogate two or more respondents, but respondents are usually not interviewed by two interviewers). The nested variable should be the second factor; e.g. interviewer number should be the first factor and respondent number the second one. If the design is actually not nested, but the factors are (partly) crossed, or the first factor is nested under the second one, the OK button will be disabled. This is determined by SEQUENCE VIEWER given the actual data in your file.

Error terms depend on whether factors to be tested are random or fixed. A fixed factor is a factor that contains all potential levels of a factor; for example gender. A random factor contains a random sample of all potential levels; for example interviewer number, because there are potentially more interviewers than available in your data. As a general rule, fixed factors are tested using the Mean of Squares of the error in the denominator of the F-value, whereas random factors use the mean of squares of the interaction. If an effect itself is not significant, you can pool it with the error term. The appropriate denominator in a nested design is the 'factor 2' within 'factor 1' (or: 'factor 2 [factor 1]') mean of squares. Here too, if the effect of factor 1 is itself not significant, you can pool the mean of squares of 'factor 1' and 'factor 2 [factor 1]' to obtain an appropriate error term, usually yielding much more power.

If you use only one factor in your design, you can also perform a Oneway analysis of variance, which gives similar results as the ANOVA procedure, if no covariate is specified; the Oneway procedure is a bit faster however.

SETUP

To perform a two-way analysis of variance on sequence variables, do this:

- Select **Sequence variables > ANOVA...** to open the ANOVA setup window.
- Click at the edit box after 'dependent variable' to select it, and enter or select the dependent variable.
- Check ☑ **covariate** to add a covariate. Than click at the field after 'covariate' and select or enter a covariate.
- Click at the edit box after 'factor 1' and select or enter the first factor.

- Check ☑ **factor 2** to add a second factor to the design. Than click at the field after 'factor 2' and select the second factor. If you don't add a second factor a one-way analysis is performed. In that case the other options are not relevant.
- Select an appropriate design with the 'design' pop-up menu.
- If you have selected 'nested', you may check ☑ **Suppress means of [variable]**, if you don't want these means in the output; in the example above of respondent number and interviewer number, you are probably only interested in the means of different interviewers, not in the means of each respondent.
- If you have selected another design, except 'orthogonal' (because you assume that both independent variables are related), you can check ☑ **Estimated marginal means**, to obtain marginal means, adjusted for differences in cell frequencies.
- Select appropriate error terms.
- Click OK.

See also:
Select sequences and Select marked sequences.

9.16 Correlation

INTRODUCTION

To calculate a correlation matrix on sequence variables, do this:

- Select **Sequence Variables > Correlation...** to open the CORRELATION MATRIX setup window.
- Click at the 'Row variables' edit box to select it.
- Click at the name of a sequence variable in the variable list to put it into the list of row variables.
- Click at another name of a sequence variable in the variable list to add it to the list of row variables. In the variable list, names of variables that are selected, are highlighted.
- To remove a variable from the list of row variables, just click it once again in the variable list.
- If you want a square matrix, you can click OK. The column variables will be the same as the row variables in the matrix.
- If you want different column variables, click at the 'Column variables' field to select it.
- Create a list of column variables in the same way as you created the list of row variables.
- To select all sequence variables, use the Select all button. All variables will appear in the selected list. The Clear all button removes all variables from the selected list.
- Click OK.

ADDITIONAL INFORMATION

You can select between listwise and pairwise deletion. Listwise deletion removes the whole case if the case has a missing value on at least one of the variables in the row or column lists. Pairwise deletion only removes a case if one (or both) of the variables between which the correlation is calculated, has a missing value.

If ◉ **pairwise deletion** is selected, you have the option to add the number of valid cases to each correlation in the output.

If ◉ **show significance** is checked, also the exact one-sided p-value is shown.

See also:
Select sequences and Select marked sequences.

9.17 Multiple regression

To perform a multiple regression analysis on sequence variables, do this:

- Select **Sequence Variables > Multiple regression...** to open the MULTIPLE REGRESSION setup window.
- Click in the 'Dependent variable (criterion)' edit box and enter a dependent variable.
- Click in the 'Independent variables' field and click at a name in the variable list to put it into this field.
- Click at another name of a sequence variable in the variable list to add it to the list of independent variables (or predictors). In the variable list, names of variables that are selected as independent variables, are highlighted.
- To remove a variable from the list of independent variables, just click it once again in the variable list.
- Check ☑ **Show variance** and/or ☑ **Show means** to calculate these statistics for the independent variables.
- Click OK

See also:
Select sequences and Select marked sequences.

Contents