

“WH” Frequency Application

Purpose:

The “WH” Frequency(WHQ) Application accepts a TRS file as input, and outputs a number of statistics about “WH” questions (WHQs) contained therein. Sample WHQs include “who,” “what,” “when,” “why,” “where,” and “how.”

Segment Filtering Process:

In order to filter the segments in the TRS file, the program looks for segments matching two criteria:

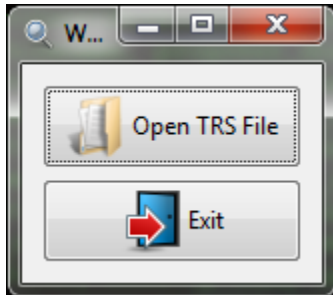
1. Transcriber code 4 (Utterance Type) must be set to “Q”, indicating that the segment is a question.
2. The transcribed phrase within the segment must contain one or more words beginning with the letters “wh”.

If both of these statements are true for a particular segment, then the segment is said to have “passed through” the filter, and it appears on the list in the segment window (see below).

If one or more of these statements is false for a particular segment, then the segment is said to have been “filtered out”, and it does not appear on the list in the segment window.

Main Window:





The main window allows you to either select a TRS file for processing as input, or exit the application.



Segment Window:

After you have selected a TRS file, the application will process it. During this time a progress bar will be displayed.

When processing is complete, the segment window appears:

WH-Frequency Counter										
Time	Phrase	Speakers	Target Listeners	WHQ Count	Who	What	Why	When	Where	How
15.26 - 16.08	this is a question with a wh	Female Adult Near	Target Child	1	0	0	0	0	0	0
15.26 - 16.08	this utterance has a what	Male Adult Near	Target Child	1	0	1	0	0	0	0
16.08 - 18.08	who what why when where	Male Adult Near	Target Child	1	1	1	1	1	1	0
18.08 - 39.90	bbbb how now	Overlapping Vocals Near	Target Child	1	0	0	0	0	0	1
Totals: WHQ Count: 4, Who: 1, What: 2, Why: 1, When: 1, Where: 1, How: 1										
Options										
<input type="checkbox"/> Group Linked Segments										
<input type="checkbox"/> Show Context										
<div>  Export Results  Add Count Column  Remove Count Column  Close </div>										

This window lists the results obtained from processing the TRS file. Initially, each row in the list corresponds to a segment (it can be altered to show utterances using the “Group Linked Segments” checkbox - see the section below for details.).

Segment List:

Most of the columns in the list should be fairly self-explanatory. However, some require additional explanation:

- The application attempts to identify the speaker(s) and target listener(s) in each segment. If the segment has multiple speakers or listeners, they will appear in the same cell, separated by a comma.
In some cases (eg. Overlapping speech transcribed using the angle bracket <> syntax), it may not be possible to determine a segment’s speaker and/or target listener. In these cases, the cell will display a dash “-”.
- The “WHQ Count” column represents the quantity that this row is counts toward the total WHQ count (the total is displayed in the status bar immediately below the list). Each row initially counts as one occurrence of a WHQ. The reason that this column exists is because during development, there was some uncertainty about how things should be counted. If the user wishes, they may adjust the value in the WHQ Count column. This can be done by clicking on the cell, then scrolling the mouse wheel to increment/decrement the number. Finally, the user may press enter or click away to save the new value. The total value in the status bar will be updated appropriately.

Target Listeners	WHQ Count	Who	What	Why
Target Child	1	0	0	0
Target Child	1	0	1	0
Target Child	1	1	1	1
Target Child	1	0	0	0

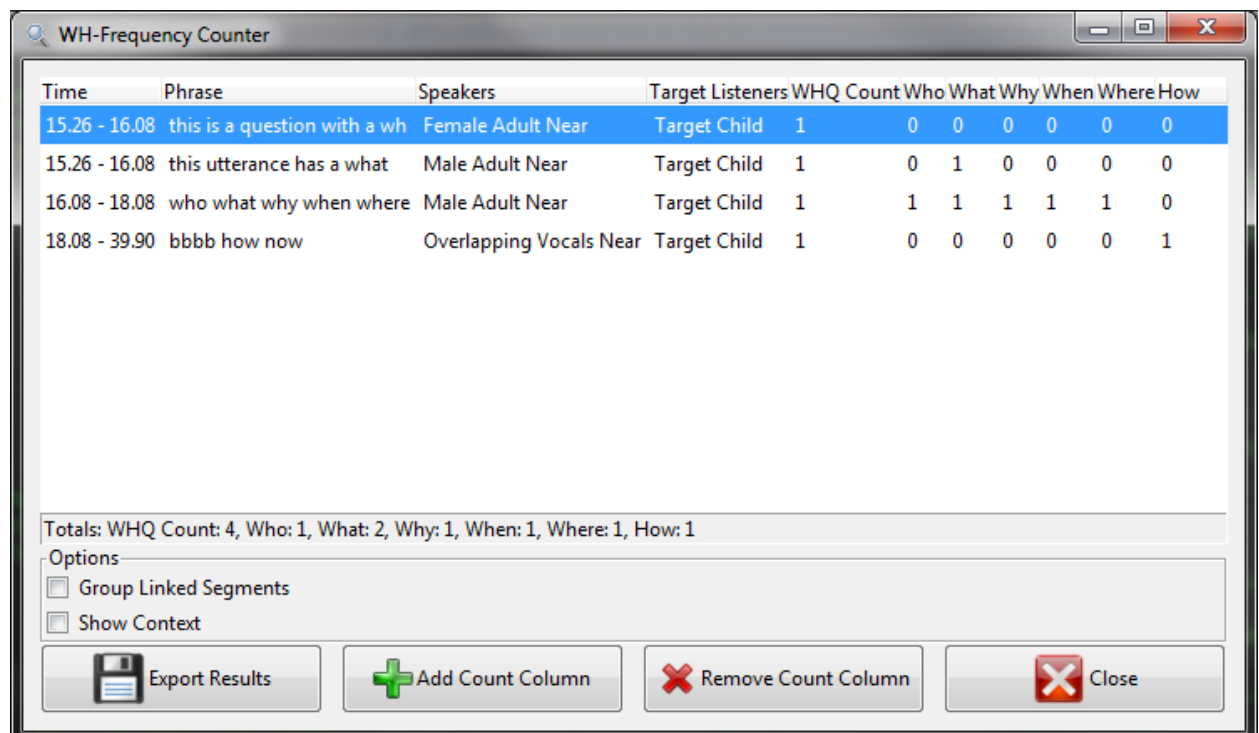
- The “Who,” “What,” “When,” “Why,” “Where,” and “How” columns display the number of occurrences of that particular word in the segment’s phrase (Note: this is a case-insensitive search). Their totals are also displayed in the status bar underneath of the list. These columns cannot be manually adjusted like the “WHQ Count” column.

Options:

The “Options” area below the list contains two checkboxes. Their functionality is described below:

- The “Group Linked Segments” checkbox allows you to consolidate segments that have been linked together by the transcribers using their “I”, and “C” codes (transcriber code 3). In other words, segments that have been linked together will be combined into a single row in the list, and will together count as only as single WHQ.

For example, consider this list as it appears before the “Group Linked Segments” checkbox is checked:



Time	Phrase	Speakers	Target Listeners	WHQ Count	Who	What	Why	When	Where	How
15.26 - 16.08	this is a question with a wh	Female Adult Near	Target Child	1	0	0	0	0	0	0
15.26 - 16.08	this utterance has a what	Male Adult Near	Target Child	1	0	1	0	0	0	0
16.08 - 18.08	who what why when where	Male Adult Near	Target Child	1	1	1	1	1	1	0
18.08 - 39.90	bbbb how now	Overlapping Vocals Near	Target Child	1	0	0	0	0	0	1

Totals: WHQ Count: 4, Who: 1, What: 2, Why: 1, When: 1, Where: 1, How: 1





Options

☐ Group Linked Segments

☐ Show Context

Export Results Add Count Column Remove Count Column Close

Now, consider it after the “Group Linked Segments” checkbox is checked:

WH-Frequency Counter										
Time	Phrase	Speakers	Target Listeners	WHQ Count	Who	What	Why	When	Where	How
15.26 - 16.08	this is a question with a wh	Female Adult Near	Target Child	1	0	0	0	0	0	0
15.26 - 18.08	this utterance has a what ->who what why when where	Male Adult Near, Male Adult Near	Target Child, Target Child	1	1	2	1	1	1	0
18.08 - 39.90	bbbb how now	Overlapping Vocals Near	Target Child	1	0	0	0	0	0	1
Totals: WHQ Count: 3, Who: 1, What: 2, Why: 1, When: 1, Where: 1, How: 1										
Options										
<input checked="" type="checkbox"/> Group Linked Segments										
<input type="checkbox"/> Show Context										
<div>  Export Results  Add Count Column  Remove Count Column  Close </div>										





Notice that the second and third segments in the original list have been combined into one (indicating that they are linked with “I” and “C” transcriber codes in the TRS file). The following changes have occurred:

- The “Phrase” column shows the “->” where the linkage took place.
- The “Speakers” and “Target Listeners” columns show the values from both original segments, separated by a comma.
- The values in the “Count” columns have changed to reflect the counts of values from the new, combined phrase.
- The “Totals” section in the status bar has been updated to reflect the new WHQ Count. Note the combined segments now count as a single WHQ.

Note that every time the “Group Linked Segments” checkbox is toggled, all of the entries in the “WHQ Count” column are reset to 1 (any manual changes will be lost). This is a programming limitation that would take a bit of “skull-gunnery” (that’s technical term...) to get around. If this becomes needed, let me know and I’ll implement it.

- The “Show Context” checkbox can be used to show the phrases of segments adjacent to the ones displayed in the list. Note that these segments may not necessarily contain WHQ questions – they are simply the adjacent segments in the TRS file. The intention is to give the user a sense of the conversational context in which the WHQ occurs.

Continuing with our previous example, the list would look like this when the checkbox is checked:

WH-Frequency Counter									
Time	Phrase	Speakers	Target Listeners	WHQ Count	Who	What	Why	When	Where How
15.26 - 16.08	(test>) this is a question with a wh (this utterance should not be present)	Female Adult Near	Target Child	1	0	0	0	0	0
15.26 - 16.08	(this utterance should not be present) this utterance has a what (who what why when where)	Male Adult Near	Target Child	1	0	1	0	0	0
16.08 - 18.08	(this utterance has a what) who what why when where (bbbb how now)	Male Adult Near	Target Child	1	1	1	1	1	0
18.08 - 39.90	(who what why when where) bbbb how now (-)	Overlapping Vocals Near	Target Child	1	0	0	0	0	1
Totals: WHQ Count: 4, Who: 1, What: 2, Why: 1, When: 1, Where: 1, How: 1									
Options									
<input type="checkbox"/> Group Linked Segments									
<input checked="" type="checkbox"/> Show Context									
 Export Results		 Add Count Column		 Remove Count Column		 Close			

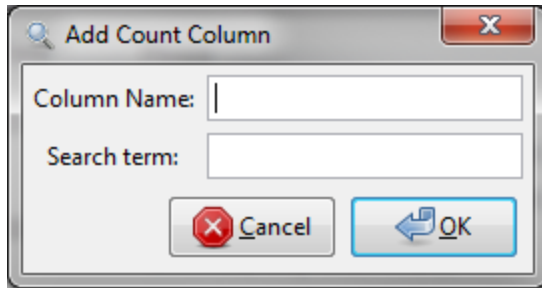
Only the “Phrase” column has changed. Each of its cells now contains three lines:

1. The phrase from the segment before the row’s WHQ segment is show in parenthesis. If there is no preceding segment (the row’s WHQ segment is the first in the TRS file), a dash “(-)” is displayed.
2. The phrase from the segment containing the WHQ is displayed (without parenthesis).
3. The phrase from the segment following the row’s WHQ segment is displayed. If there is no following segment (eg. End of file), a dash “(-)” is displayed.

Buttons:

The buttons at the bottom of the window perform the following functions:

- “Export Results” – This button exports the contents of the list to an Excel file (*.csv format). Clicking it will cause a save dialog to pop up, prompting the user for a location in which to save the resulting file.
Note: The list is exported as it appears in the window, so if either/both the “Group Linked Segments” or “Show Context” checkboxes are checked, the changes they cause to the list will also appear in the exported file. Similarly, any custom “count columns” (see below) that you have added will be exported as well.
- “Add Count Column” – This button allows you to add additional columns displaying the frequency of a particular word or character (or pattern) within the phrase. The column total is also appended to the status bar.
For example, suppose we wanted to add a column that counted all of the occurrences of the word “this”. Clicking the “Add Count Column” button brings up the following dialog box:



The “Column Name” entry is the name that will appear in the header row at the top of the list window.

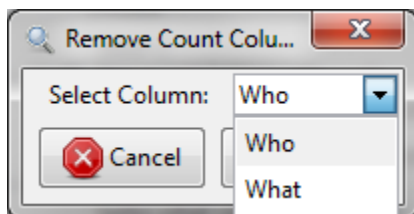
The “Search Term” entry is the text you wish to search for in the phrase of each segment (Programmers may be interested to note that this entry actually accepts a regular expression, allowing for more complex pattern matching. Others may not be so interested...)

Clicking OK adds the column the list. Its total is displayed in the status bar along with the other totals. This column will be exported along with other data when the “Export Results” button is clicked.

- “Remove Count Column” – This button allows you to remove a particular count column from the list.

Note: The first four columns (“Time”, “Phrase”, “Speakers”, “Target Listeners”, and “WHQ Count”) are permanent and cannot be removed.

Clicking the button causes a popup dialog to display. The dialog contains a drop-down list with the names of all removable count columns. Simply select the one you would like to remove, and click OK.



- “Close” – This button closes the window. It does not exit the application. New TRS files can be opened from the main window.