Reliability Application

This program's purpose is to measure any potential differences between human and LENA judgments of transcription segment properties. A basic usage scenario looks something like this:

First, the user selects corresponding TRS/CSV and WAV files. Next, the program chooses segments from the files (based on criteria given by the user) and presents them, one at a time. For each segment, the user can listen to the audio clip and input their opinions on several properties of the segment. Each segment is presented to the user twice – once "without context" (i.e. the audio clip played begins at the segment start time and ends at the segment end time), and once "with context" (i.e. the audio clip played begins several seconds before the segment start time and ends several seconds after the segment end time). Finally, the user can export the results to an Excel spreadsheet.

Running the program:

The reliability application is accessible on the LENA machine via a desktop shortcut of the same name.

Terminology:

The following terminology is useful for understanding how the program works:

- "Check" – A check is a combination of settings used to run tests (see below). These settings include the TRS/CSV file, corresponding WAV file, the number of segments to choose, the type(s) of segments to choose, and the number of seconds worth of context to provide for "with context" tests.

- "Running a Check" – This action involves performing one or more tests (see below). The total number of tests performed is equal to the number segments specified in the check being run. This number can be specified as a percentage of the total number of segments in the file, or as a specific number. Additionally, a limit can be placed on the total number of selected segments (Since, for example, 1% of a file with 50,000 segments is still 500 segments => 1000 tests…). If the specified percentage would yield more segments than the limit, the number of segments picked will be equal to the limit.
- "Test" – Performing a test involves randomly selecting one segment from the TRS/CSV file, playing it for the user, and collecting feedback. The user provides feedback using dropdown boxes (indicating the type of segment that they believe they heard). The dropdown options are described in detail in a later section.
There are two types of tests: tests "with context" and those "without context". Tests with context play an extra couple of seconds (the exact number is specified in the check) on either side of the audio clip in an attempt to allow the user to get a better sense of what's going on in the environment. Tests without context do not provide these extra seconds.

- "Exporting Results" – A check can be run multiple times. Each time a check is run, an entry is recorded in the application's database. Exporting results involves selecting one of these entries (i.e. selecting one particular "run of the check") and receiving output in the form of an Excel spreadsheet. More on this in later sections.

Main Window:

The main application window looks like this. The buttons are described below the figure.
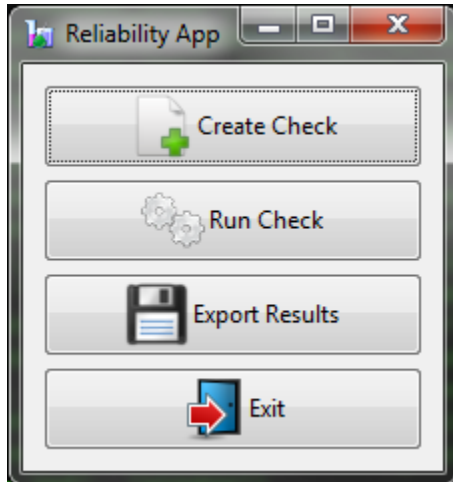
1. "Create Check" – This button allows the user to select and save the settings that make up a check. When it is clicked, the "Create Check" form displays:

The inputs are described below:

- "Name" – enter a name for the check – this is the name you will select when running the check (or exporting results for a particular run) later on.
- "TRS/CSV File" – specifies the TRS or CSV file to grab segments from. A CSV file must contain the following columns: "Elapsed_Time, Segment_Duration, Speaker_ID"
- "WAV File" – specifies the WAV file to grab corresponding audio clips from.
- "Percentage of segments" – the percentage of segments to use (percentage of total number of segments in the TRS file) when running the check.
- "Segment limit (0=None)" – in some cases, files may have a large number of segments. In these cases, even specifying a percentage as low as 1% may still yield more tests than desired. This field can be used to limit the number of segments used for tests. If the limit is lower than the number of segments generated by the given percentage, the number of tests performed is equal to limit * 2 (each segment generates 2 tests - one with context, and one without). Otherwise (or if the limit is set to 0) the number of tests performed is: (percentage of segments) * (total segments in the file) * 2. To specify a specific number of segments rather than a percentage, set the percentage of segments to zero, and the segment limit to your desired number.
- "Default context padding (sec)" – This is the number of extra seconds on either side of the audio clip to play for tests with context. This is a default setting – you can adjust the amount of padding during each individual test that includes context.
  E.g. Suppose that this context padding is set to 2.0 sec. Now suppose that the program has randomly selected a segment that begins at an offset of 5.0 sec in the wav file, and ends at an offset of 7.0 sec. Tests without context will play the portion of audio from 5.0 sec to 7.0 sec. Tests with context will play the portion of audio from 3.0 sec to 9.0 sec.
- "Pick Segs Randomly" – if checked, segments will be picked in random order (though still within the criteria matching any filers (see below) that have been set. If unchecked, segments (again, matching filtering criteria) will be picked in ascending order, starting from the beginning of the file.
- "Filters" section – this section allows you to set up restrictions on the types of segments that can be picked for testing. You can restrict selection to a specific time interval, or to one or more specific types of segments.
 - Clicking the "Add" button brings up the "Add Filter" window:
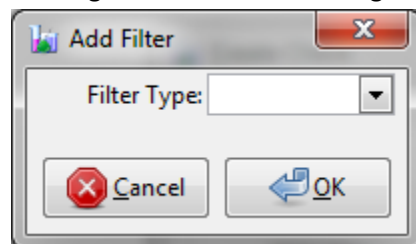
  - Select a filter type from the dropdown. Additional options will appear when you do. The two filter types are described below.

1. The "Speaker" filters allow you to restrict testing to segments of one or more speaker types. Selecting this option causes the following list to appear:
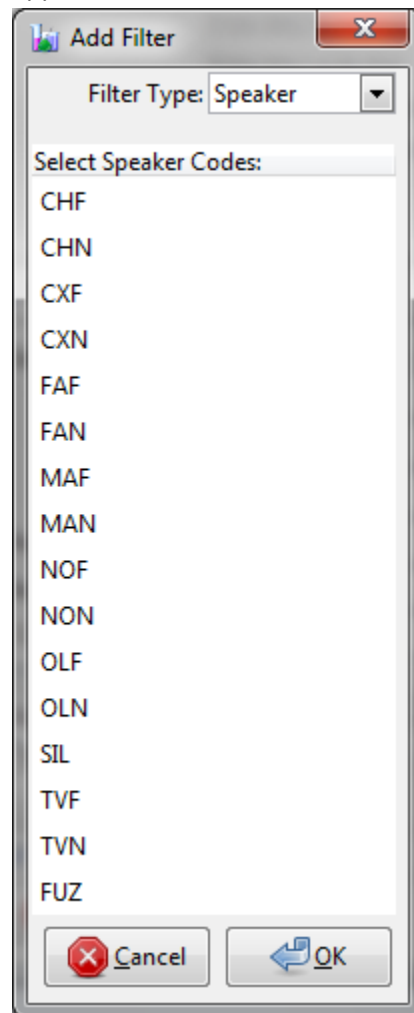
   Click on a code to select it. Hold Ctrl and click to select additional codes. These codes will form an OR relationship (segments with speakers that match any one of the selected codes will be eligible for selection).

2. "Time" filters – these can be used to restrict the segment selection to a specific time period within the TRS/CSV file. Selecting the "Time" option from the "Filter Type" dropdown causes the following options to appear:
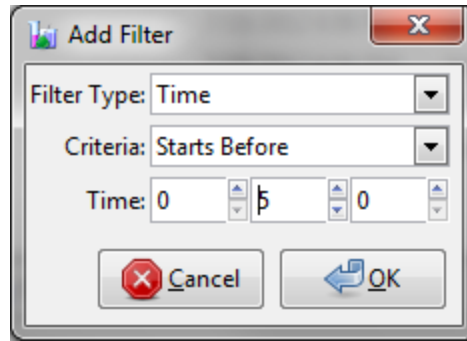
The "criteria" dropdown allows you to select the type of time cutoff you would like the filter to cause. The "Time" inputs allow you to specify the cutoff time.

For example, suppose the criteria dropdown is set to "Starts Before," and the "Time" inputs are set to 5 minutes (as in the image above). In this case, the filter will only allow tests to use segments that start within the first 5 minutes of the TRS file. The other "Criteria" options (Starts After, Ends Before, Ends After) work similarly.

You can set multiple time filters to create a bounded time period from which segments can be picked. For example, suppose we only wanted to allow segments in the period from 5:00-10:00. This could be accomplished by setting two time filters: one "Starts After" 5:00, and one "Ends Before" 10:00.

o   After you have set the filter options, click the "OK" button in the "Add Filter" window to set the filter. The window will close, and a new row will be added to the list in the "Create Filter" window.
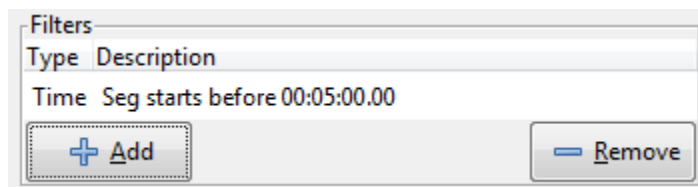
o   If you add multiple filters to the same check, these filters will be ANDed together. For example, suppose you set two filters; a speaker filter for the code "FAN", and a time filter "Starts Before" 00:05:00. This will cause the program to pick segments that have both a FAN speaker AND start before 00:05:00.
    Note: It is possible to create filters that do not allow any segments through. Equivalently, it is also possible that the TRS file does not contain any segments matching the criteria. In these cases, when the check is run, a popup will be displayed explaining the situation, and the check cannot be executed.

o   Ordinarily, the process of picking segments is done in such a way that the same segment is never picked twice. If it happens that there are not enough segments in the TRS/CSV file to satisfy the check's "percentage of segments" or "limit" criteria, an error message is displayed.

o   To remove a filter, select the corresponding row and click the "Remove" button.

- "Run Immediately" – If checked, the check will be run (i.e. tests will be performed) immediately after the OK button is clicked. In unchecked, the check will be stored in the program's database, and can be run from the main window via the "Run Check" button (see below).

2. "Run Check" – This button allows you to select a check and run it (i.e. run the tests). Clicking brings up the "Run Check" window:
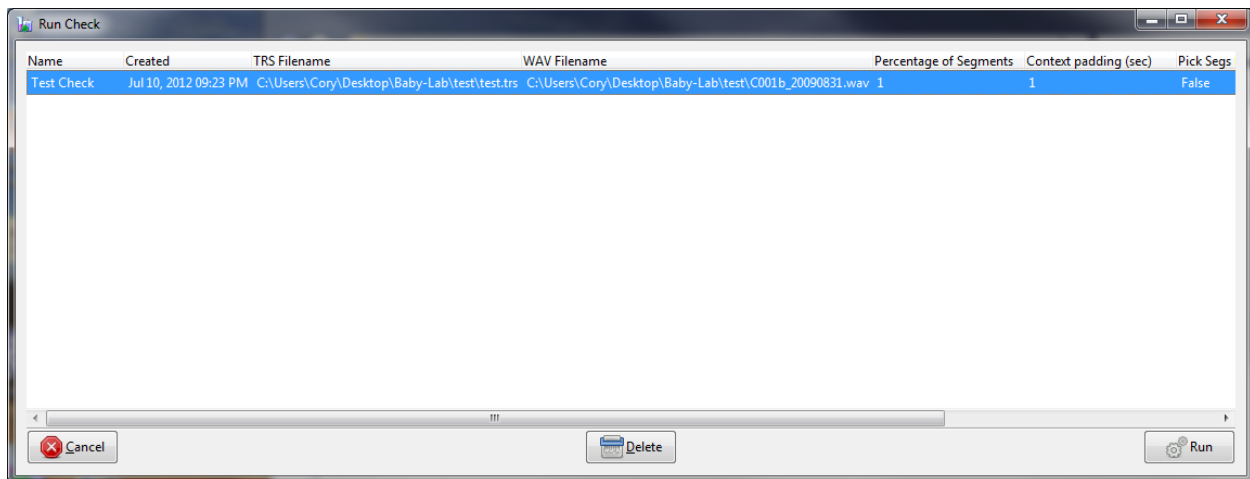
This window allows you to select the check you wish to run. The table displays one row for each check in the database. The check's settings are listed in the columns to the right. To select a check, click on the corresponding row and hit "Run". To delete a check, select a row and click the "Delete" button. A confirmation dialog will display.

Note: The application uses a "soft delete" technique to remove the check. This means that the check is simply marked as deleted. Checks that are marked in this way never show on the screen – however, all of the information associated with the check (their settings, runs, etc.) still remains in the application database. If you delete a check and later realize it was a mistake, it can be revived upon request.

After you click OK, a progress bar will be shown. During this time the application is going through the TRS file, selecting random segments. After this process has completed, the test window is displayed:
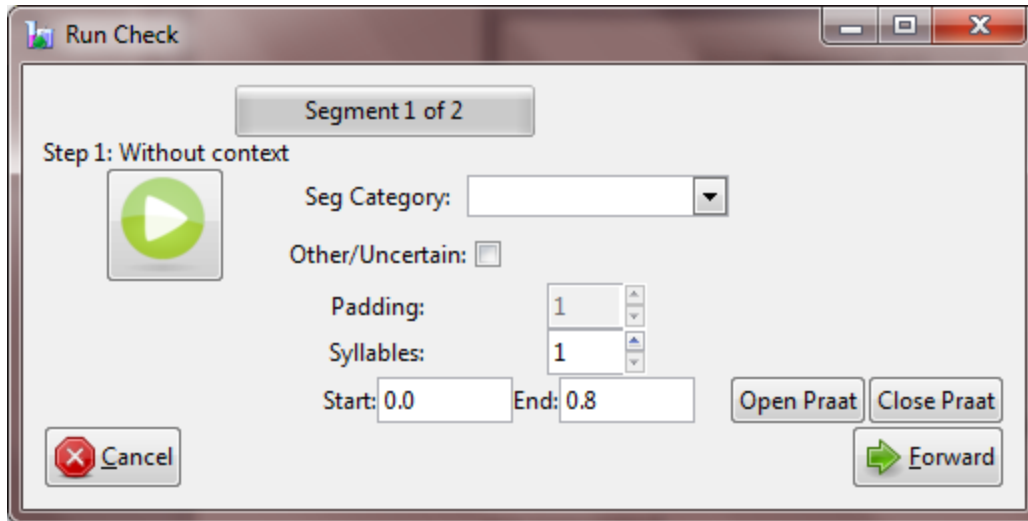
At the top of the test window, a progress bar is shown. This progress bar advances as the tests are completed.

Below the progress bar, a label is displayed, showing the step number for the current test. Each test goes through two steps: "without context" and "with context."

Click the play button to hear the sound from the current segment.

Select an option from the category dropdown indicating the type of segment you think that the sound represents.

If the clip's category is difficult to determine, you may also check the "Other/Uncertain" box. This marks the segment in the final output spreadsheet.

Padding may be individually adjusted for tests with context via the spin button entry.

Enter the number of syllables present in the clip using the syllables spin button entry.

Finally, adjust the absolute start and end boundaries of the clip so that they match the sound. This information will appear in the final spreadsheet alongside the LENA-defined boundaries. You can use Praat to select your boundaries. Click the "Open Praat" button. This causes a Praat spectrogram window to appear. Select the appropriate portion of the sound using the mouse (click and drag). Then, switch back to the Reliability App, and click the "Close Praat" button. This closes praat, and updates the Start and End inputs to the boundaries you selected in the spectrogram window. Due to limitations in the ways we can communicate between these programs, it is not possible for the reliability app to retrieve the values if you simply close Praat. Instead, you have to click the "Close Praat" button in the Reliability App.

After you have selected your options, click the "Forward" button to advance to the next step/test.

(Note: The application saves the dropdown values to the database after each step. If you click the "cancel" button at any point before the run is complete, all tests from the current run that have been recorded so far will be removed from the database. Tests from other runs will not be affected.)

On the last step of the last test, two elements of the test window will change. First, the "Forward" button will alter its text to "Finish," indicating that the last step is being displayed and the window will close when it is clicked. Second, an "Export Immediately" checkbox will appear.
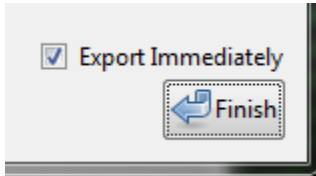
If the "Export Immediately" checkbox is checked when the user clicks the "Finish" button, a window will display prompting for a location in which to save the exported data. See the "Export Results" button description below for more information on exporting.

3. "Export Results" – this button allows you to export the results of a particular run in spreadsheet format. Clicking it causes a file selection window to display. This window allows you to select a location in which you'd like to save the file. In addition, the "Open Immediately" checkbox in the lower right-hand corner allows you to specify whether or not you'd like to open the file in Excel immediately after the save operation.
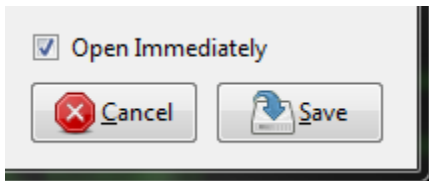
(Note: the path to the spreadsheet program is located in the "settings" table of the application database. It may need to be changed if the application is every migrated to a different machine where Excel resides in a different directory.)

The output format is "*.csv" (comma-separated value). This is a simple format that should be open-able in most spreadsheet programs.

4. "Exit" – exits the reliability application, closing all open windows.

Exported File Details:

Most of the information in the exported files should be self-explanatory; however there are a few columns that may require an additional explanation:

- Test Number – Each test has two steps; with and without context. Therefore each test number appears in two rows.

- Actual codes – this is a comma-delimited list of all of the actual codes that LENA determined the selected segment had (multiple codes indicate multiple speakers in the same segment).
- Ratio Correct – this row displays a value for the Category Correct column. The value is computed as:

  Value = (count of all rows marked TRUE in the column) / (total rows in column)

  Or (equivalently):

  Value = (number of responses that the user got correct) / (total number of responses)