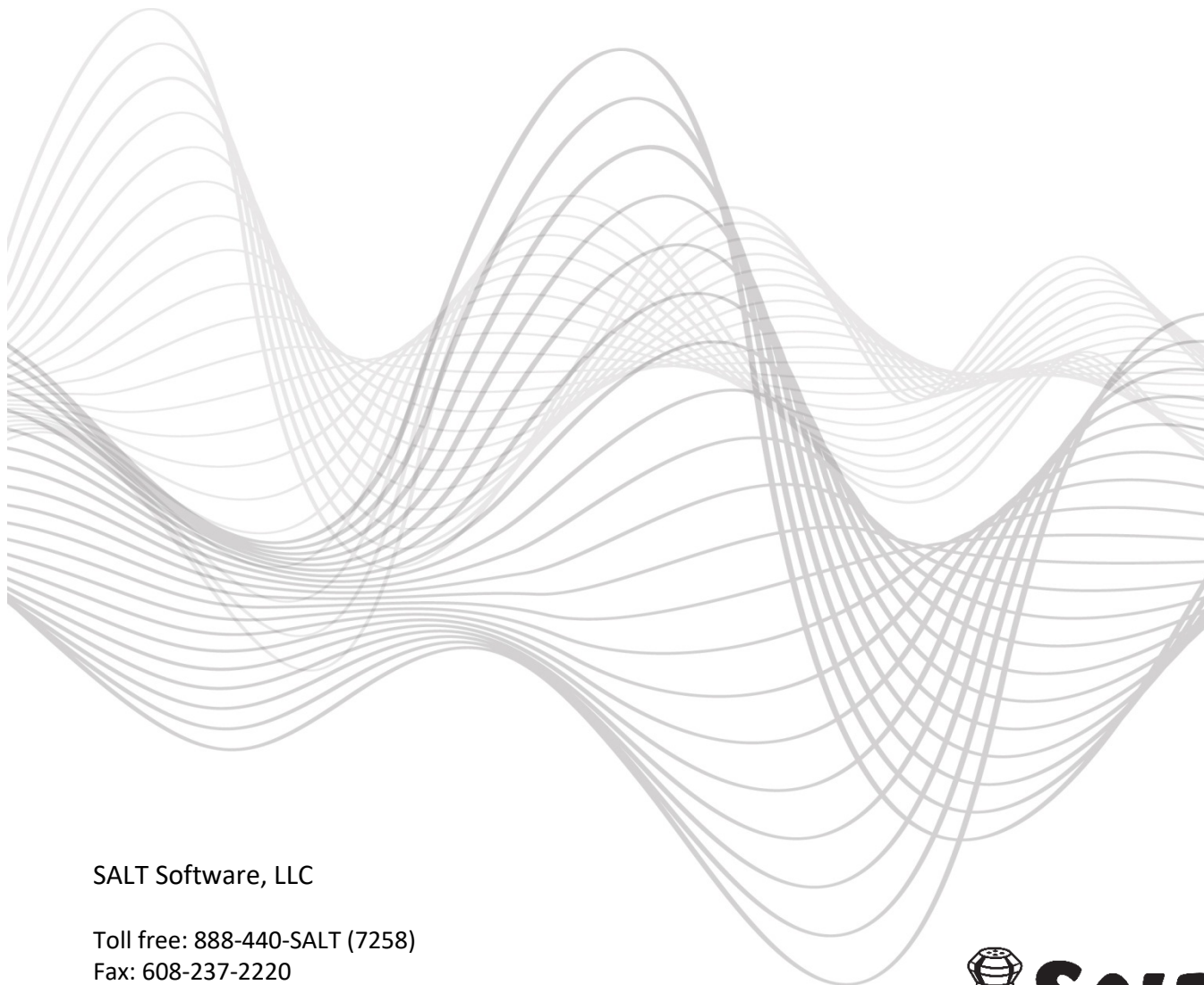


STEP-BY-STEP EXERCISE

More Research Tools



SALT Software, LLC

Toll free: 888-440-SALT (7258)

Fax: 608-237-2220

support@saltsoftware.com

www.saltsoftware.com



More Research Tools

This exercise uses four transcript files (AA1Nar.slt, AA2Nar.slt, AA3Nar.slt, and AA4Nar.slt) and covers five tools, (Insert Grammatical Category Codes, Explore Multiple Transcripts, Generate Word List, Generate Bound Morpheme List, and Generate Code List).

A. Start SALT with the AA1Nar Transcript



1. The SALT program begins with the Getting Started window. In this exercise you will be opening one of the sample transcripts so click the **Open** button and you are presented with the Open dialogue box.
2. Unless you changed the default transcript folder, the *My SALT Data\Transcripts* folder located within your *Documents* folder is displayed. Open the **Samples** folder and then the **Lesson Samples** folder.
3. Select the transcript file **AA1Nar** from the “Lesson Samples” folder and then click the **Open** button. You should have the transcript displayed on the screen.

B. Insert Grammatical Category Codes

SALT contains an algorithm for parsing words into 23 grammatical categories. This algorithm is based on English words and is not available for other languages. In this section, you will insert codes at the end of words identified as belonging to specific grammatical categories. Before doing this, however, you will look at several of the grammatical category reports.

1. Run the grammatical category reports.
 - Select **Analyze menu** → **Grammatical Categories**.

Notice that 23 grammatical categories are listed along with the number of words used by each speaker in each category. The list is repeated twice, first for all utterances and then for the analysis-set utterances. Notice that the child used 63 verbs and 1 copula form.

Suppose you want to see which verbs and copula forms were used.

- Select **Analyze menu** → **Grammatical Category Lists....**

You are presented with a dialogue box in which the 23 grammatical categories are listed.

- Check **Verbs** and **Copula Forms**.
- Keep **Analysis set** as the Utterance Base.
- Click **OK** to generate the list of verbs and copula forms. There were 24 verbs (63 occurrences) and 1 copula form (1 occurrence) identified in the transcript.
- Select **Window menu** → **Close All Reports**. The only window left should be the transcript.

2. Code the current transcript for grammatical categories.

Note that this edit option can only be used with one transcript at a time.

- Select **Edit menu** → **Insert Grammatical Category Codes**.

You are presented with a dialogue box in which the 23 grammatical categories are listed. Following each category is a code. This is the code which will be inserted at the end of each word in that category.

- Check **Verbs** and **Copula Forms**. Keep the default codes, [GC:VERB] and [GC:COP]
- Keep **Analysis set** as the Utterance Base and **1st speaker (Child)** as the speaker.
- Click **Insert GC Codes**.
- You are informed that 64 words were coded. Click **OK**.

Scroll through the transcript and look at the words which were coded as verbs and copula form. Although the parsing algorithm works well, you may find verbs that should have been coded but weren't or words that were coded as verbs but should not have been. By coding the words identified as belonging to one or more categories, you can "fix" the transcript by manually inserting and deleting these codes.

- Select **Window menu** → **Close All**. Do not save the transcript changes.

C. Explore Multiple Transcripts

The "Explore Multiple Transcripts" options provide a method for you to select, count, and code utterances and/or words you are interested in across a set of transcripts. They are similar to the Explore options selected from the main menu except these Explore options apply to a set of transcripts. Each transcript is explored independently, one after the other. This is a useful tool for looking at specific types of utterances across many transcripts. You may be interested, for example, in looking at all utterances which contain specific words, codes, or combination of words and codes. The selected utterances can be listed, located within the transcripts, coded, or saved as separate transcripts.

1. Select the Tool.

- Select **Tools menu** → **Explore Multiple Transcripts** → **List**.

Notice that, by default, no transcripts are selected; the target speaker is **1st speaker**, the utterance base is the current **Analysis set**, and the Include option is set to **Hits**. If you have questions on any of these settings, refer to the Help option in this dialogue box.

2. Select the Transcript Files.

- Choose **Select Transcripts** and you are presented with the "Select Transcripts" dialogue box.
- Click **Browse for Files**. Unless you changed the default transcript folder, the *My SALT Data\Transcripts* folder located within your *Documents* folder is displayed. Open the **Samples** folder and then the **Lesson Samples** folder.

Note: if the default transcript folder has been changed, follow these steps to reset the transcript folder to the default setting:

- Click the **Cancel** button to close the “Open” dialogue box.
- Select **Setup menu** → **File locations** → **Transcripts**. The current transcript folder is displayed.
- Click the **Reset** button to reset the transcript folder to the default setting.
- Click **OK**.
- You are asked if you want to save the folder setting as the program default. If you select **Yes**, the previous folder setting is lost. If you select **No**, the change is only in effect until you exit the SALT program. Select **Yes** or **No**.
- Select **File menu** → **Open** and go back to step 1 (above).

- Select four transcript files: **AA1Nar.slt**, **AA2Nar.slt**, **AA3Nar.slt**, and **AA4Nar.slt**. To select all four at the same time, either 1) hold down the <Ctrl> key on the keyboard and click once on each of the four transcript files, 2) click the first file (AA1Nar.slt) and then hold down the <Shift> key and click the last file (AA4Nar.slt), or 3) click the check box before each file.
 - After selecting the four transcript files, click **Open** and the four transcripts are listed in the dialogue box window.
 - Click **OK** to return to the “Rectangular Data File” dialogue box and notice the message “4 transcripts” next to the “Select Transcripts” button.
3. This exercise assumes that you are familiar with selecting words and/or utterances. In this section, you will repeat the exploration from the exercise titled “Exploring a Transcript”. Select all the omitted words and bound morphemes, and the words coded as errors with the [EW__] error code.

- Click the **Word and Code List** button.
- Type the following word and code list in the large edit box:


```
*=
/*=
[EW=]
```

where “*=” matches all omitted words, “/*=” matches all omitted bound morphemes, and “[EW=]” matches all codes beginning with “EW”. *For help with specifying word and code lists, select “Help” from this dialogue box.*

- Click **OK**.
- Click **List** to list the omissions and word errors for the four transcripts selected in step 2.

Notice that the results for each of the transcripts are listed separately. Also notice that only three of the four transcripts are listed. This is because the Include setting is “Hits” rather than “All Transcripts”. The missing transcript (AA3Nar.slt) did not contain any omissions or word errors.

4. Locate the omissions and errors in the transcripts.
- Select **Tools menu** → **Explore Multiple Transcripts** → **Find**.
 - Notice that the transcripts and previous selections are still in effect. Click **Find** and you are

presented with the Explore Multiple Transcripts – Find dialogue box.

The three transcripts with omissions and/or errors are listed in the bottom half of the dialogue box. The matching utterances for the first transcript are listed in the top half of the dialogue box. The first transcript is open and the first matching utterance is highlighted in the transcript.

- Move to the selected utterances in each of the transcripts. You have the option, at any time, of clicking in the current transcript and editing it.
- Click **Close** to close the dialogue box.
- Select **Window → Close All**. If prompted, do not save any of the transcripts.

5. Discuss the other Explore options.

- Select **Tools menu → Explore Multiple Transcripts → Insert Code**.

Notice that the transcripts and previous selections are still in effect. The “Insert Code” option is used to insert the code of your choice at the end of the selected utterances and/or words. An output window listing the number of codes inserted in each transcript is displayed after the codes have been inserted in all of the transcripts.

- Select **Cancel** to close this dialogue box without inserting any codes. *Refer to the exercise titled “Advanced Exploration – Coding & Saving as Transcript” for an example of inserting codes in a single transcript using the Explore option.*
- Select **Tools menu → Explore Multiple Transcripts → Delete Code**.

This option is used to delete the code of your choice from the end of the selected utterances and/or words. An output window listing the number of codes deleted in each transcript is displayed after the codes have been deleted from all specified transcripts.

- Select **Cancel** to close this dialogue box without deleting any codes. *Refer to the exercise titled “Advanced Exploration – Coding & Saving as Transcript” for an example of deleting codes from a single transcript using the Explore option.*
- Select **Tools menu → Explore Multiple Transcripts → Save as Separate Transcripts**.

This option is used to create a set of new transcripts which contains the selected utterances and/or words from each of the original transcripts. The new transcripts contain the \$ speaker line and the current selections for that transcript. Depending on the “Include” settings, each new transcript may also contain all or some of the non-utterance lines taken from the original transcript. Once the new transcripts are created, they can be analyzed as you would any other transcripts.

There are two settings in this dialogue box related to the location and naming of the new transcripts.

1. Specify the folder where you want the new transcripts stored. Because you have the potential for creating a large number of new transcripts, you may want to store them in a separate folder. You would either type in the name or use the Browse button to select that folder. If the folder does not exist, you are asked if you want it created.

2. Specify the filename prefix. In order to provide a new name for the transcripts, you have the option of providing a prefix for each new filename to begin with. The new filenames consist of the prefix you specify here plus the original filename. In our example, if you specify the prefix "OE_" (for omissions and errors), the new transcripts would be named OE_AA1nar.slt, OE_AA2nar.slt, OE_AA3nar.slt and OE_AA4nar.slt. If you do not provide a prefix, the new filenames are given the same name as the original transcripts.
- Select **Cancel** to close this dialogue box without creating any new transcripts. *Refer to the exercise titled "Advanced Exploration – Coding & Saving as Transcript" for an example of creating new transcripts using the Explore option.*

D. Generate Word List

The "Generate Word List" tool produces an alphabetical list of all the word roots used across a set of transcripts. The transcripts are combined and only one list of word roots is generated. You could use this tool to study an individual's total vocabulary since it produces one list of words generated from all the transcripts of that individual. You could also use this tool to study the vocabulary used by any number of speakers during a specific sampling context. Another use for this tool is to look for misspelled words across a set of transcripts. You would then use the "Explore Multiple Transcripts → Find" tool to locate and edit these misspelled words.

1. Select the Tool

- Select **Tools menu → Generate Word List**.

Notice that the same four transcripts are still selected. All four counts are checked, the target speaker is the 1st speaker, the utterance base is total utterances, and the area is main body. If the current settings are anything else, click the **Reset** button.

2. Generate the List

- Click **OK** to generate the list. A single alphabetical list is generated from the words found in all the transcripts. The words included in the list depend on the settings. Based on the current settings, they include all words from the main body (*excludes mazes*) of all the 1st speaker utterances.

Notice that the first word is preceded by an asterisk indicating that it was an omitted word. Only word roots are listed, no inflections or codes. For example, if the words "LOOK", "LOOK/ED", "LOOK/ING", and "LOOK[EW:SEE]" were found, they would all be counted under the word root "LOOK". The Speaker, Utterance Base, and Area settings affect which words are included in the list. The first column counts the number of transcripts (1 – 4) which contains each word. The second column is the percent of transcripts containing each word, the third column counts the total number of occurrences of each word, and the last column counts the number of transcripts (0-4) in which the word was mentioned first by the target speaker.

3. Generate a List of Words Coded as [EW:___].



Notice the two edit boxes near the bottom of the dialogue box. They are used to restrict the selected words based on the codes they may or may not have attached to them. Suppose, for example, you had a set of transcripts from bilingual children where all the code-switched words

were marked with the code [CS]. To generate a list of all the code-switched words, you would enter the code [CS] in the edit box labeled “Only include words containing at least one of the following codes”. To generate a list of all the words which were NOT code-switched, you would enter the code [CS] in the edit box labeled “Exclude words containing at least one of the following codes”.

In this exercise, you will generate a list of all the words which contain the [EW__] error code. The wildcard character, “=”, is used in the next step to match all variations of this code (*refer to the “Help” option for details*).

- Type the following code in the “Only include...” text box:
[EW=]
- Click **OK** to generate the list. The words included in this list were all coded as [EW].

Although counts are given for the number of transcripts containing each word, there is no way of knowing the specific transcripts involved.

4. Locate a specific word across the set of transcripts.

The previous step generated a list of all the words coded as [EW]. The first word listed was **DID**. Suppose you want to find the transcript and utterance containing this error.

- Select **Tools menu → Explore Multiple Transcripts → Find**.

The same four transcripts are still selected. The word list in step 3 was based on the 1st speaker, total utterances, main body words.

- Change the “Utterance Base” to **Total utterances**.
- Click the **Word and Code List** button.
- Click **Reset** to clear the previous search list. Do not save the current list.
- Type the following word specification in the large edit box:
DID[EW=]
- Click **OK**.
- Click **Find** to locate the word DID coded as an error in the four transcripts selected previously.

Notice that this word occurred in the first transcript, AA1nar.slt. This transcript is opened and the utterance is highlighted.

- Click **Close** to close the Find dialogue box.
- Select **Window menu → Close All**.

E. Generate Bound Morpheme List

"Generate Bound Morpheme List" is similar to "Generate Word List" except it produces an alphabetical list of all the bound morphemes used instead of the words used.

1. Select the Tool

- Select **Tools menu** → **Generate Bound Morpheme List**.

Notice that the same four transcripts are selected. The settings indicate that all three counts are selected, the list is based on bound morphemes rather than words containing bound morphemes, the target speaker is the 1st speaker, the utterance base is total utterances, and the area is main body. If the current settings are anything else, click the **Reset** button.

2. Generate the List

- Click **OK** to generate the list.

A single list is generated from the bound morphemes found in the four transcripts. The bound morphemes included in the list depend on the settings. Based on the current settings, they include all bound morphemes from the main body (*excludes mazes*) of all the 1st speaker utterances. Bound morphemes which begin with an asterisk were omitted in obligatory context.

3. Generate a List of Inflected Words



- Change the "List BMs or Words" setting to **Words** and click **OK** to generate the new list. This time all words containing bound morphemes are listed.
- Select **Window menu** → **Close All**.

F. Generate Code List

"Generate Code List" is similar to "Generate Word List" and "Generate Bound Morpheme List" except it produces an alphabetical list of all the codes used instead of words and bound morphemes used.

1. Select the Tool

- Select **Tools menu** → **Generate Code List**.

Notice that the same four transcripts are selected. The settings indicate that all three counts are selected, the target speaker is the 1st speaker, the utterance base is total utterances, and the area is main body. The code set is all codes and the code type is word and utterance codes. If the current settings are anything else, click the **Reset** button.

2. Generate the List

- Click **OK** to generate the list. One list is generated from the codes found in all the transcripts. The codes included in the list depend on the settings. Based on the current settings, they include all codes from the main body (*excludes mazes*) of all the 1st speaker utterances.

3. Generate a List of Word Codes



- Change the code type to **Word codes** and click **OK** to generate the new list. This time only word codes are included.
- Repeat the previous two steps changing the code type to **Utterance codes**.

Hint: Use the “Generate Word List”, “Generate Bound Morpheme List”, and “Generate Code List” tools to look for misspelled words and codes. Then use the “Explore Multiple Transcripts → Find” tool to locate the specific transcripts to be edited.

G. Close All Windows

- Select **Window menu → Close All**.

