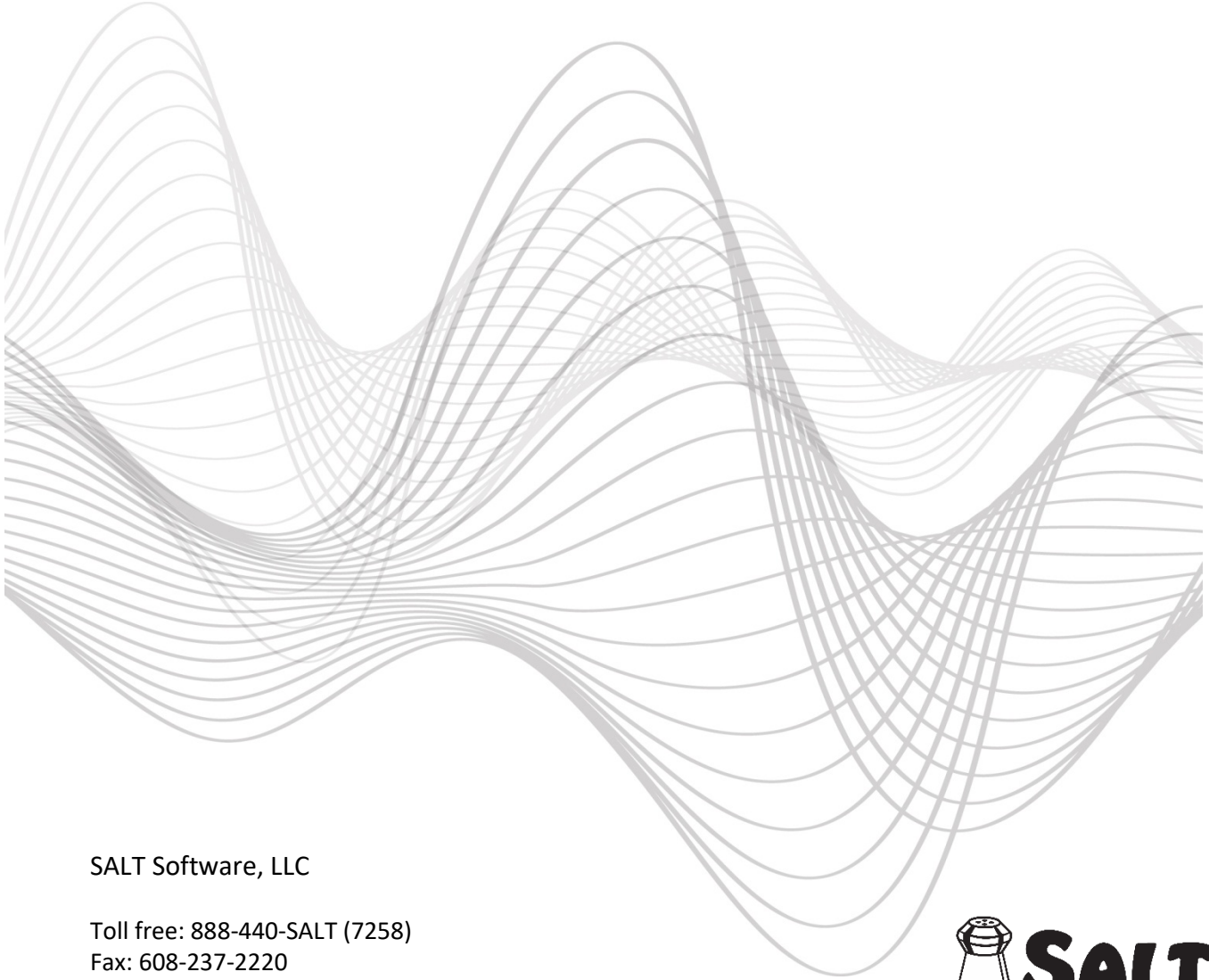# STEP-BY-STEP EXERCISE

Creating a Reference Database

SALT Software, LLC

Toll free: 888-440-SALT (7258)
Fax: 608-237-2220
support@saltsoftware.com
www.saltsoftware.com

Creating a Reference Database

In this exercise you will create a reference database and use it for comparison within the Database menu. This database will be built from four sample transcripts: "AA1Nar.slt", "AA2Nar.slt", "AA3Nar.slt" and "AA4Nar.slt". Each of these transcript files contains a narrative sample (speaker selects story) from a typically developing seven-year-old African American student. Although four transcripts are insufficient as a comparison set, they are used in this exercise to illustrate the process of creating and using your own reference database.

A. Start the SALT Program

The SALT program begins with the Getting Started window. Typically, you begin by creating a new transcript or opening an existing transcript. In this exercise, however, you will be using the Tools menu so click the **Close** button to close this dialogue box.

B. Make Sure the Database Transcripts are Ready to be Stored

1. Select **Tools menu → Build Reference Database** and you are presented with the Build dialogue box.

2. Select **Help** from the Build dialogue box and read the paragraph titled "Preparing SALT Transcripts". Then close the help window.

> "The transcripts must contain plus lines specifying either current age (+CA: or +DOB: and +DOE:) or grade in school (+Grade:). Transcripts without plus lines specifying current age or grade in school may not be added to the database. Transcripts should also contain plus lines specifying sampling context (+Context:) and, if appropriate, selection subgroup (+Subgroup:). In addition, gender, if found, is stored in the database and can be used for selection. Although timing lines are not required, the "Database menu --> Standard Measures Report" compares the timing and rate information if it is available."

The four transcripts that will be used to build this database should all contain the following plus lines:
   + Gender: M or F
   + CA: 7;0 to 7;5
   + Context: NAR
   + Subgroup: SSS

These samples are narratives where the speaker selects story (SSS).

3. Click **OK** to close the Build dialog box.

4. Check for required plus lines.

   - Select **Tools menu → Rectangular Data File** and you are presented with a dialogue box.

   - Select the transcripts.
      – 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.

    – Select four transcript files: **AA1Nar.slt**, **AA2Nar.slt**, **AA3Nar.slt**, and **AA4Nar.slt** and click **Open**. The four transcripts are listed in the dialogue box window.

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 → Open** and go back to step 1 (above).

    – Click **OK** to return to the "Rectangular Data File" dialogue box and notice the message "4 transcripts" underneath the "Select Transcripts" button.

- Select **Plus Lines**.

- Check **Gender**, **CA: (yy;mm)**, **Context**, and **Subgroup**. Click **OK** to close the dialogue box.

- Click **Generate Data**. Make sure all transcripts contain the required plus lines.

  Transcripts must contain either age (+CA, or +DOB & +DOE) or grade (+Grade) data. There is no grade data for these transcripts. Gender, sampling context (+Context), and selection subgroup (+Subgroup) are optional but highly recommended. You need to know which speaker is the target speaker (1st), what the sampling context is (Narrative), and what the subgroup is (SSS *for speaker selects story*).

C. Create a New Empty Database

Note: there are two folders where SALT expects to find the reference databases:
1. Built-in Databases: for databases that come with the software. By default, this folder is
   C:\ProgramData\SALT Software\SALT 20\Reference Databases
2. External Databases: for databases you add to the software. By default, this folder is
   Documents\My SALT Data\External Reference Databases

Although this exercise directs you to create the new database in the default External database folder, you may create it any folder you want. Just make sure you remember the name and location of that folder.

---

1. Select **Tools menu → Build Reference Database** and you are presented with the Build dialogue box.

2. Make sure the **Database Location** is set to **External Databases**.

3. Select **New Database** from the Build dialogue box to begin a new database. You are immediately prompted with the "Save As" dialogue box. The contents of the default External database folder are displayed. Initially this folder is empty.

4. Click in the "File name:" box and type the name of the new database file. You may use any name you want, but for this exercise use **SampleAA**. The database extension ".sltdb" (*for Salt database*) will be automatically added to the end of the filename.

5. Click the **Save** button and an empty database is created in the External database folder.

D. Build the Database

1. Change sampling context. Only one sampling context may be specified for the database. If you have transcripts from a variety of sampling contexts, you would create separate databases for each context since you would not want them grouped together for comparison.

   - The transcripts you are using to create this database are narratives. Change the **Sampling Context** to **Nar**.

2. Consider the "Default Age, Grade, and Gender Match" settings. These settings determine the default selection criteria for this database. "Age match" is set to plus or minus 6 months. Grade match and gender match are not selected. If you were creating a database with many samples in a narrow age range, you might change this setting. The Bilingual Spanish/English Story Retell database, for example, contains over 4,000 samples and defaults to an age match of plus or minus 2 months, with grade match. These are just the default settings; users have the option of changing them.

   Keep the default settings.

3. Consider the "Options for Multiple Subgroups" settings. Although all the transcripts stored in the database must be based on the same sampling context, you may choose to combine transcripts from different subgroups. For example, the TNL2 Narrative Samples database has a different subgroup for each task (McDonalds, Aliens, LateForSchool, and All 3 Stories). If you specify different subgroups when adding the transcripts, users have the option of selecting these subgroups when choosing database records to use for comparison. The options in this section affect the subgroup selections available when comparing transcripts to this database and only apply if the database transcripts have been added with two or more subgroups.
   - Allow <all> as a selection: if checked, users have the option of comparing their sample with all subgroups combined; if not checked, users must specify one, and only one, subgroup.
   - Default subgroup (if no match): default subgroup if the subgroup specified in the user's transcript (+Subgroup:) does not match any of the subgroups found in the database.

   Since the subgroup is the same for the samples in this database, these settings don't apply (*ignore them*).

4. Consider the "Options for Multiple Ethnicities" and "Options for Multiple Locations" settings. If you specify different ethnicities or locations when adding transcripts to the database, users have the option of selecting them when choosing database records to use for comparison. The options in these sections only apply if the database transcripts have been added with two or more ethnicities or locations. These options affect the selections available when comparing transcripts to this database.

   Since the ethnicities and locations will not be specified for the samples in this database, these settings don't apply (*ignore them*).

5. Consider the "Default method used to equate samples by length". When comparing transcripts to this database, this setting affects the default method used to equate samples by length.

    Keep the default setting, **Words**.

6. Consider the "Restricted Access" setting. Some databases are contributed for clinical use only. Samples from these databases may be selected for comparison but not selected using the *Rectangular Data File* tool. When building the database, you have the option of restricting access to the database by embedding a specific password. If a password is specified here, users will be required to enter this password in order to select transcripts from the database when using the *Rectangular Data File* tool. The checkbox indicates whether or not a password has been set. **Important note**: once a password has been specified, you must enter that password in order to remove it or change it. There is no recovery option.

    Keep the default setting. Do not enter a password to restrict access.

7. Enter the database description which is displayed if users view the "Database Snapshot" when selecting a database comparison set. Enter that description here.

    •   Click the **View or Edit Database Description** button in the Build dialogue box.

    •   Click in the edit window and type in the following description:

        This database contains narrative transcripts from several typically developing African American students.

    •   Click **OK** and the description is automatically stored in the database file, "SampleAA.sltdb".

8. View the list of error codes, fluency codes, and nonmainstream-form codes which are associated with this database. When this database is used for comparison, these codes are used when reporting on the contents of the selected database samples. These lists may vary depending on the transcripts used to build your database.

    •   Select **View or Edit Error Codes** and the current list of error codes is displayed. Do not change the list of error codes for this exercise.

    •   Click **OK** or **Cancel** to return.

    •   Select **View or Edit Fluency Codes** and the current list of fluency codes is displayed. Since the database samples do not contain fluency codes, you can either delete all the codes or leave it as is since these codes will not be found in the database samples.

        Delete all the fluency codes.

    •   Click **OK** to return.

    •   Select **View or Edit Nonmainstream-form Codes** and the current list of nonmainstream-form codes is displayed. The database samples use the code [D] to mark dialect. This is the first code in the list.

You can either delete the other codes or leave it as is since these codes will not be found in the database samples. Do not change the list of nonmainstream-form codes for this exercise.

- Click **OK** or **Cancel** to return.

9. Add transcript records to the database.

- Select **Add Transcripts** and you are prompted with the "Add Transcript Records" dialogue box.

- 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 you're Documents folder is displayed. Open the **Samples** folder and then the **Lesson Samples** folder.

- Select four transcript files: **AA1Nar.slt**, **AA2Nar.slt**, **AA3Nar.slt**, and **AA4Nar.slt** and click **Open**. The four transcripts are listed in the edit box.

- Click **OK** to return to the "Add Transcript Records" dialogue box with "4 transcripts" selected.

- The four transcripts you are storing all have the target speaker as the 1st speaker. (*This is determined by the $ speaker line at the beginning of each transcript.*) Make sure the target speaker is the **1st speaker**.

- For the **Selection subgroup**, enter **SSS** (speaker selects story). Leave **ethnicity** and **location** blank.

  All the samples for this database have the same subgroup (SSS). Consider the following example, however, which illustrates the use of the subgroup label when more than one subgroup is involved. You collect a set of narrative language samples based on several different stories (*as is the case with the Narrative Story Retell database*). The elicitation protocol is the same; only the story varies. You want to be able to select them for comparison based on a specific story. You would create only one database for these language samples with the sampling context "Narrative". Each transcript, however, would be identified by its subgroup, a label identifying the specific story. This subgroup label could later be used as an additional selection criterion. Similarly, the ethnicity and location labels would be used to identify sets of transcripts from different ethnicities and locations. The subgroup, ethnicity, and location labels, if entered here, may later be used for selection.

- **Within-utterance pauses** and **Between-utterance pauses**. Notice the two check boxes on the bottom of the left-hand side of the dialogue box. These are used to specify whether or not pause times were marked in the transcripts. These settings are needed because the software cannot determine whether or not pauses were marked by looking at the transcripts. It could be that there are no pauses marked because there were no significant pauses in the sample. If you have a set of database transcripts in which some of them contained pause times and some do not, you would add them to the database separately. When comparing pause data, only the samples added with these settings checked are used.

  Pause time, within and between utterances, was marked in these four transcripts so the two checkboxes in this section should be checked.

- **Display messages while scanning**. You want all warnings and error messages to be displayed so

keep all the display messages options checked. If the subgroup, for example, was not included as a plus line in the transcripts, you would uncheck the "Non-matching subgroup – stored" setting. If the subgroup, ethnicity, or location is not specified, it will not be checked for a match when the transcripts are added (*so it doesn't matter if the message option is checked or not*).

- You are now ready to add the transcript records to the database. Click **Add**.

- A progress window is displayed while the records for each transcript are stored in the database file. The transcripts are encrypted to protect their content.

- After a short time, you are informed that 4 records were added. Click **OK** to continue.

  Each transcript record is listed with the record ID, which is a combination of filename, speaker, sampling context, and selection subgroup. Had you specified an ethnicity or location, they would have been included in the record ID.

  The database is complete for the purposes of this exercise. You could, of course, add additional transcripts to this database.

- Click **Close Database** and the database is saved and ready to use. Note that even if you just closed this dialogue box, the database would be saved and closed.

- Click **OK** to close the build database dialogue box.

E. Compare a Transcript to the SampleAA database

1. Select the sample transcript file "Courtney.slt".
   - Click the open icon [icon] or select **File menu → Open...**.

   - Select the transcript file "**Courtney.slt**" from the "Documents\My SALT Data\Transcripts\Samples\Lesson Sample" folder and then click the **Open** button.

   - The "Courtney" transcript is displayed on your screen. Courtney is a 7-year-old African American student. The transcript is a narrative with Courtney telling the story of Goldilocks and the three bears.

2. Select samples from the SampleAA database to use or comparison.

   - Select **Database menu → Select Database Samples and Settings**. You are prompted to select the database comparison set for Courtney.

   - STEP 1: Select database

     – The default database for narratives is <u>Narrative SSS</u> which is stored within the default folder for Built-in databases. You should change this to the SampleAA database you just created.

     – Click the **Select External Database** button. The database files in the default folder for External databases are displayed.

     – Select the database file "SampleAA.sltdb" and then click the **Open** button. The sampling context is "NAR" which matches the sampling context of the transcript. Note that the "Subgroup",

though unavailable as it cannot be changed, is "SSS".

- Click the **Database Snapshot** button for a brief description of the SampleAA database and a summary of the database record types. Notice that there are four language samples in the age range 7;0 to 7;5. Click **OK**.

- STEP 2: Select age, grade, and/or gender criteria

  - "Age match" is set at plus or minus 6 months. "Grade match" is not available because the SampleAA database doesn't contain grade data (*because none of the database transcripts contained the +Grade plus line*). "Gender match" is available because the SampleAA database contains gender data (*because all or some of the database transcripts contained the +Gender plus line*).

  - Click **Find Matched Samples**. The four database samples all match Courtney's age.

- STEP 3: Select method used to equate samples by length

  - The target speaker is **1st speaker**. Notice the transcript information at the bottom of the dialogue box. Courtney's age is 7;2 and her language sample contains 304 NTW (number total words), 46 C&I Verbal utterances, and 3:01 elapsed time.

  - Keep the default setting, **Same number of total words**.

  - Click **Find Equated Samples** to select the best match.

  - You are presented with four choices. Only one of the database samples was as long, in terms of NTW, as Courtney. To use all four samples, they would all need to cut at 61 words. Select the 3rd option, **3 samples – 189 words** and click **OK**.

  - Click **OK** again to close the selection dialogue box.

3. View the *Standard Measures Report*.

   - Select **Database menu → Standard Measures Report**.

   - Select **Entire transcript** (recommended) and click **OK**. The *Standard Measures Report* is displayed.

F. Editing Database Content

Once a database has been created, you can make changes to the contents.

1. Select **Tools menu → Build Reference Database** and you are presented with the Build dialogue box.

2. Make sure the **Database Location** is set to **External Databases**.

3. Select **Open Database** and select the **SampleAA** database from the list of External databases.

   - Click the **View or Edit Database Description** button in the Build dialogue box.

   - Click in the edit window and type the following note at the beginning of the database description:

     Sample database - do not use for comparison!

- Click **OK** and the updated description is stored in the database file, "SampleAA.sltdb".

Notice that there are options for adding additional transcripts, removing existing transcripts, editing the code lists, and changing the sample context and other settings.

4. Click **OK** to close the build database dialogue box. Any changes you made to the database are automatically saved.

G. Changing the Default Databases

In the previous section, the Narrative SSS reference database was pre-selected based on the context. Suppose you wanted your own database to be the default database for SSS samples.

1. Select **Setup menu → Analysis Settings → Default Databases**.

2. Press the Help button to read about how you select the default databases. Close the Help window.

3. Notice that the 3rd row contains the default database for Nar SSS, Narrative SSS (B). The (B) indicates that this database is found in the default folder for Built-in databases. Click in the box containing this database. The two Browse buttons are now available for selecting the database.

   Click the **Browse External (E)** button and select the **SampleAA** database. Notice that this database is followed by (E) indicating that it is found in the default folder for External databases.

4. Since we are not going to change the defaults at this time, press **Cancel** to exit this dialogue box without changing the changes.

H. Converting Old Databases to Current Format

If you created your own reference databases using an older version of SALT, you may need to convert them to the current format before they can be used.

1. Select **Tools menu → Build Reference Database** and you are presented with the Build dialogue box.

2. Select **Open Database** and change the file type in the drop down list from "SALT Database File (*.SLTDB)" to "Older Format SALT Database File (*.SD1)". Select your database from the list of older databases and you will be prompted to convert your database.

3. Your converted database is opened in the dialogue box where it can be edited to suit your needs. Click **OK** to close the build database dialogue box. Any changes you made to the database are automatically saved.

I. Close All Windows

- Select **Window menu → Close All**.