## Summary Generator

1. **Overview:**

The purpose of the summary generator application is to provide a tool to combine ADEX summary, special case file and transcribed TRS files into one “overall” summary together.

1. **Terminology:**

*ADEX summary files* are groups of csv files which generated from LENA directly. Generally there are 6 different types of data defined in it. They are Child Data, Female Adult Data, Male Adult Data, Other Data, AVA Data, Time Data and Signal Level Data. Each type of data is composed by a small number of criteria. For example, signal Level Data includes Average Signal Level and Peak Signal Level. You can check APPENDIX I for details. Currently we only consider calculate the average from **“AWC, Turn\_Count, Child\_Voc\_Duration, FAN\_Word\_Count, FAN, MAN\_Word\_Count, MAN, CXN, OLN, TVN, NON, SIL”** these criteria. If you need to calculate others, please contact programmer.

*Special case file* is a spreadsheet file which used to record abnormal situations or handwriting notes for the participants. This file is used to create a series of filters to filter out some participants that we don’t want to put into the final summary file. For example, we can create a filter to find out the participants who only spoke French or a filter to find out participants who were not sick during recording. These filters will affect the data in ADEX summary and transcribed TRS files.

Transcribed TRS files are generated from the “statics program” separately. In order to generate these files, you need to choose generate an

1. **Usage:**

Special case needs a sheet named “Special Cases” to work. And must have “ITS File” and “Study Number” as heads.

1. **Outputs:**

# Introduction (from Elizabeth):

ADEX Summary Spreadsheet:

The most current set of spreadsheets is saved here:

Experimental Data\Daycare Study\ADEX Output\ADEXAnalysis

-They are saved in folders for the appropriate childcare setting

\* The most recent example of this is called SummarySpreadsheetJuly2014 and is saved here:

C:\Experimental Data\Daycare Study\ADEX Output\Elizabeth\June 2014 ADEX Analysis

## Instruction (updated with Dr. Melanie):

- 1. Remove 1st and last 30 minutes of data with Audio\_Duration column. At least 1800 seconds, so 6-7 rows usually, sometimes you have to delete an extra row to get to 1800.

- 2. Filter out naptime using the naptime filter.

Naptime Database is saved here:

C:\Experimental Data\Daycare Study\ADEX Output\Elizabeth\2014 AdultChildRatio Project\Naptime Data

- 3. Organize summary spreadsheet into the 3 separate childcare settings

- 4. Pull out Participant code (ex. C001a), age of participant at first recording, and gender into the summary spreadsheet

- 5. Take averages of the following columns for each spreadsheet:

AWC, Turn\_Count, Child\_Voc\_Duration, FAN\_Word\_Count, FAN, MAN\_Word\_Count, MAN, CXN, OLN, TVN, NON, SIL

- 6. Add the averages of these columns from each participant separately, and then get an average per participant. (i.e. add all of the C001a averages together and find an average for C001a, do the same for C001b, etc.)

- 7. Have a column called # of recordings, and have the number of spreadsheets for that participant in that column. (i.e. C001a has 2 spreadsheets, so that number would be 2)

# Progress:

## ADEX Processor:

- 1. read one CSV file as a list of list. [V]

- 2. get required columns. [V]

- 3. process filters on data. [V]

- 4. calculate the average. [V]

- 5. write results into excel. [V]

- 6. get filenames from other processors to exclude.

- 7. save preliminary results. [V]

## Comments Processor:

- 1. filter required column. [V]

- 2. generate outputs. [V]

- 3. write results into excel. [V]

## Transcribed data handler:

- 1. load information from csv [V]

- 2. merge with same ID [V]

- 3. save results to excel. [V]

## load/save configuration:

**APPENDIX I: ADEX INFOMATION**

AWC - Adult Word Count

Turn\_Count - Conversational Turns

-Key Child Data:

Child\_Voc\_Count - Key Child Vocalization Count

CHN - Key Child Segment Duration

Child\_Voc\_Duration - Key Child vocalization Duration

Child\_NonVoc\_Duration - key Child Non-vocalization Duration

-Female Adult Data:

FAN\_Word\_Count - Female Adult Word Count

FAN - Female Adult Segment Duration

FAN\_NonVoc\_Duration - Female Adule Non-Vocalization Duration

-Male Adult Data:

MAN\_Word\_Count - Male Adult Word Count

MAN - Male Adult Segment Duration

MAN\_Nonvoc\_Duration - Male Adult Non-Vocalization Duration

-Other Data:

CXN - Other Child Segment Duration

OLN - Overlap Segment Duration

TVN - TV Segment Duration

NON - Noise Segment Duration

SIL - Silence Segment Duration

FUZ - Uncertain Segment Duration

-AVA Data:

AVA\_RS - AVA Raw Score

AVA\_SS - AVA Standar Score

EMLU - Estimated Mean Length of Utterance

AVA\_DA - AVA Develpmental Age

-Time Data:

Recording\_Index Recording Index

Elapsed\_Time - Elapsed Time

Clock\_Time - Clock Time

Audio\_Duration - Audio Duration

-Signal Level Data:

Average\_Signal\_Level - Average Signal Level

Peak\_SignalLevel - Peak Signal Level