<div align="center">**TRANSCRIPTION PROTOCOL**</div>

Organizing the files:

1. You will receive a folder with the following files:
    a. Sound File
    b. Demographic information
    c. Bilingual Language Profile
    d. Interviewer Information
    e. Field Notes

2. You will also receive a number for the interviewer (i.e. CESA050, CoBiVa050)

3. Each file should be labeled using the CESA/CoBiVa number, round number, and your last name:
    a. Sound File = CESA0XX_SF_R1_LASTNAME
                    = CoBiVa0XX_SF_R1_LASTNAME
    b. Demographic information = CESA0XX_DI_R1_LASTNAME
                    = CoBiVa0XX_DI_R1_LASTNAME
    c. Bilingual Language Profile = CESA0XX_BLP_R1_LASTNAME
                    = CoBiVa0XX_BLP_R1_LASTNAME
    d. Interviewer Information = CESA0XX_II_R1_LASTNAME
                    = CoBiVa0XX_II_R1_LASTNAME
    e. Field Notes = CESA0XX_FN_R1_LASTNAME
                    = CoBiVa0XX_FN_R1_LASTNAME
    f. Transcription = CESA0XX_TR_R1_LASTNAME
                    = CoBiVa0XX_TR_R1_LASTNAME

Transcription Software:

How to Transcribe with ExpressScribe
(Insert Time Stamps and/or Transcribe)

Caution: Save Frequenty!

1. Go to https://www.nch.com.au/scribe/index.html

2. Click on the 'Get free version of Express Scribe here.'

3. Set up your 'hot keys' for the keys that you prefer. (Click on Options or Preferences to get to 'hot key' set up.)

4. Add a key for 'copy time.'

5. Use the hot keys to slow, rewind, fast forward, pause, and play the audio.

6. Type into a text editor.

7. Click the hot key for 'copy time' and then ctrl+v to paste into your text editor.

8. Reformat according to WEBVTT format and submit.

Transcription Instructions:

1. Working on the transcript
   a. Always use headphones to transcribe
   b. Rest every 10 minutes
   c. Transcribe EVERYTHING you hear, even the interviewer and any hesitations or repeats

2. Signaling turns in speech
   a. For the CESA corpus, E is used for the interviewer (entrevistador/a) and P is used for the participant.
       E:
       P:
       E:
       P:
       P 2 P 3 P 4, etc: (for anyone else who enters the conversation)

   b. Fort the CoBiVa corpus, <v PAR> is used for the participant and <v INV> is used for the interviwer
       <v INV>
       <v PAR>
       <v INV>
       <v PAR>

3. De-identify any identifying information in the transcript as XY, including names, schools attended, street names where the participant has lived.
   a. Do **not** de-identify names of towns or cities where they live or places where they went just once to get a coffee, for example. If it is a place of work, consider whether there are several or just one in the city/town and whether they mentioned the cross-streets or location. Think about whether this identifies who the person is. If you de-identify too much, then we have to go back to the original and re-do the work. **Please ask if you have a question.**
   b. Also, you will need to silence this segment in the audio file. This can be done using the free program Audacity, or other similar programs.

4. Orthography
   a. Unintelligible speech   xxx   quiero xxx mañana

b. Lengthened vowel   :   pue:s

c. Comments added by transcriber   [ ]   [risas] [suena el teléfono]

d. False start   -   pue- pues no

e. Direct/quoted speech   " "   y dijo, "sí"

f. Mark Other Language using **(OL)** before the word(s) that are switched. Include borrowings (such as biles, parqueadero) and code-switches.

g. Use standard spelling in the transcription. This is for the purpose of consistency, searchability, readability, and in order to avoid 'eye dialect.' (For example, write pues instead of pos.)

h. For morphological variation, you may write that as pronounced. (For example, fuiste, fuistes). **Ask if you have a question.**

i. Do not make notes about phonological, lexical or morphosyntactic features within the transcription. You may note thse in the 'Fieldnotes.' Note interesting and/or frequent phonological, lexical or morphosyntactic features in the 'Fieldnotes.' (For example, no pronunciation of the 's', code-switching for 'so', etc) This is **not** required. The fieldnotes are also for making notes about the context.

j. Represent short interjections (2-3 words) without starting a new turn. See example below for /yeah/

00:00:00.00 --> 00:00.05.00
<v PAR> So I didn't know what we were going to do like (OL) no sé /yeah/ But there we were.

In this example, 'yeah' is a remark from the interviewer. Since it is a conversation between two individuals, it is not necessary to switch to a separate line and speaker code each time.

k. Mark pauses in the following manner:
       ,   Short breath (May also be used for clarity)
       .   Slightly longer breath (May also be used for clarity)
       …   Slightly longer pause (Under 1 minute)

      [2.0]   Pause over 1 second. This example is for 2 seconds.

      Do **not** spend too much time on this. Just do the best you can.

l. Represent interjections consistently. (For example, hm, mhm, uh, ehm, aha, um.) Again, do not worry about this much. Just put what it sounds like to you.

5. WEBVTT Format:

    a. WEBVTT Format allows the transcripts to be time-aligned and clickable on the website. See examples below.

    b. Before each line of typed speech, include the time stamp (start and stop) in the format as seen above- 00:00:00.000 --> 00:00:18.000 (You may use ExpressScribe to get the timestamps.

    c. For each line of speech, you must attribute the speech to the correct speaker as seen above.
        i. For the Participant, use <v PAR>
        ii. For the Investigator, use <v INV>

    d. See more details in Transcription Instructions on the following pages and Transcription Training materials/videos.

How to Transcribe Using Stream:

NATIONAL ENDOWMENT FOR THE HUMANITIES

1. Convert your audio into a video using VLC.
    a. Check if your device has VLC Media Player. If not, download here: https://www.videolan.org/vlc/index.html (You don't need to donate. It's optional.)
    b. Click on the Media tab.
    c. Then, click Convert/Save from the drop-down menu.
    d. On the Open Media window, click 'Add.'
    e. Browse to select the correct audio file.
    f. At the bottom right side of the Window next to Cancel, there is a dropdown menu with 'Convert/Save,' click 'Convert.'
    g. In the 'Convert' window, next to Profile, use the dropdown window to select 'Video- H.265 + MP3 (MP4).
    h. In the same file, next to 'Destination File' click 'Browse' to decide where you want to save the file.
    i. Rename the file as COBIVA001_Video.

      j.   Click 'Save.'
      k. Then, click 'Start' to start the file conversion.
      l.   See short video tutorial.

2. Go to Stream by accessing Microsoft Office 365 through your university's apps.

3. Click on the Stream App. It is a pink arrow. Note that you may have to click 'Other Apps.'

4. Click +Create.

5. Click Upload the Video file that you just created.

6. Drag the file to the screen or browse to upload the file.

7. Let the processing bar finish, and the click 'Publish.'

8. Go to 'My Content' and then 'Videos.'

9. Then, click on the pencil next to the video to 'Update Video Details.'

10. Under 'Details,' make sure that under video language, it has the correct language selected Spanish OR English. (You have to choose one.)

11. Also, check that under 'Options,' 'Autogenerate a Caption File' has a check mark, and the box is pink to show that it is selected.

12. Under 'Permissions,' de-select 'Allow everyone in your company to view this video.'

13. The video will take some time to process, but please check to try to estimate as closely as we can how long the captioning takes. Sometimes the caption file has been created before it is available to the right of the video. If there is a CC (closed captions), you will already have access to the caption file under 'Update Video Details.'

14. When the video has processed, click on 'My Content' and then 'Videos.' Then select the video to listen to and edit the transcript that shows up in the right side 'transcript' box. Edit the transcript and add the participant codes. (You can copy and paste them if it is faster.) Note: Alternatively, you can try editing and adding participant codes after you download the file.

15. When the video has processed, you can click the pencil next to the video for 'Update Video Details' and click 'Download file' next to Captions under 'Options.' If you haven't yet, add the participant codes and edit the transcript.

16. Once you have the transcript, follow the steps in the next section to clean up the transcription file.