**PLEASE READ ALL OF THIS CAREFULLY!**

Last semester, DIS corrected AI-generated transcripts of conversations so that we can accurately know what both the husbands and wives in these videos say!

These transcripts will be used by a group in University of Buffalo, and before we send them, I want to be reasonably confident that they are mostly accurate. I want to know how accurate they are – 90%, 95% … nearly perfect? Therefore, you have been assigned the task of quality control. I have selected a sample of the transcripts worked on last semester, and I need you to do the following:

First, go to the following excel sheet: [simple\_transcripts\_qc.xlsx](https://fsu-my.sharepoint.com/:x:/g/personal/mi23h_fsu_edu/EQhpbrugy2tJi5V-9_LmftgB-Lk3M7N0WshdQB5DKzFdlQ?e=mdv10n)

Here, you will see that you have been assigned 6-7 videos. The names of the videos are coded such that they are unique to a conversation. For example, 4003 refers to a couple, C1 means that this was the first of four conversations couples in this study had. And T1 refers to the timepoint of the study. There are T4 and T7 as well, but we are not doing those right now. You are just checking a sample **of the T1 transcripts.**

In the top left of this little app, you can input the name of the video you have been assigned, and this will load both the video and the corresponding transcript.

**IMPORTANT NOTE:** Please make sure that you load up the correct video and transcript. It is very easy to make the mistake of loading up the wrong video! Make sure that what you work on matches the excel sheet with your assignment.

You have what you need to verify that the transcript is accurate! **I need you to do two things in this task:**

1. Make necessary corrections to transcripts
2. **TRACK** the number of word modifications made (additions, removals, or changes)

**Transcript Correction Guidelines**

**Punctuation**

* **Fix major punctuation errors only** 
  + Add spaces after periods (e.g., change "sentence.But" to "sentence. But"). These spacing corrections should be noted as they affect word count
  + Correct major issues. It has to be really noticeable for you to change it. For example, if there is no punctuation **at all**, you should add commas or periods to denote pauses. There are no strict rules for punctuation. As long as it reasonably reflects the way the people are speaking, that’s good enough. If you are doubting whether to correct it, it’s probably fine to leave it. Don’t spend too much time on anything punctuation-related.

**Word Accuracy Guidelines – MAIN point of task**

* **Focus on identifying and correcting:** 
  + Missing words (especially in unclear speech or overlapping dialogue)
  + Incorrectly added words that should be removed
  + Mistaken words where you're confident of the correct term
  + (UNINTELLIGIBLE) markers - decode if possible, but okay to leave if unclear
  + Leave text as-is if the original transcriber's interpretation of a word seems reasonable
* Words like “hm” or “hmm” can be tricky. I think that if it is something you think the other person would hear and contributes to the conversation, then you should add it. You shouldn’t include things like sounds people make, or laughter.

**Video Review Process**

* **You will watch the video twice!** One time focusing on the husband, and another on the wife - this should make things easier for you
* However, use your own judgment - if you happen to get a couple where they talk for long periods of time and alternate clearly, then it might be better to alternate as you check (this is likely rare)
* You can speed things up by listening to the video at a different playback speed, which the video player should allow you to change, **but don't compromise your ability to hear the correct words**
* I imagine that something like 1.25x speed shouldn't affect listening ability too much, and if you get used to that perhaps even 1.5x - some couples differ on how quickly they talk, just please don't compromise the accuracy, which is what we are mainly going for here!

**Tracking Modifications**

* **Keep a running tally of all word-level changes:** 
  + Words added
  + Words removed
  + Words modified

Note: I don’t need to know a breakdown of each – I just want **ONE** number that tells me how many errors you had to modify throughout the transcript.

* **Do NOT include punctuation changes in this count**
* **Suggestion: Use a tally sheet while working to track modifications**A number written in black

  Description automatically generated

**Saving Your Work**

* Hit the “Save” button regularly. This creates a 'corrected' version of the transcript in a google cloud storage bucket - you don't have to worry about saving the transcript in a particular place, as I will have access to it once you save and it says it saved successfully.
* WHEN you are done: Tell me how many changes you made to the transcript by putting it in the excel sheet where you found your video assignment.
* If you encounter anything noteworthy, you can note it in the Notes column in the spreadsheet also. For example, maybe all the modifications you made were adding meaningful “hmms” that the original transcriber omitted. Anything that you think would help contextualize the nature of your corrections would be good to note!

Note: Notice that the app allows you to load the corrected version of the transcript. Remember, when you make changes to the original and save them, this creates the corrected version of the transcript. So, in the case where maybe you are halfway through and want to return to a transcript later for further editing, you would quit the app, and when you return, load **the corrected version**, **not the original**. From there, you could continue working on and save like normal! **Be careful here - if you load the original by mistake and save it when a corrected version exists, this overwrites the current corrected version and you would lose all changes you made.** I can’t revert this, so please just understand that saving things save them to the corrected version of the transcript.