Comments on the annotation process

Non-speech
I've used this label for all instances when a sound cannot or should not be annotated. Instances of this category include: noise, background voices, laughter, coughing, hum and haw, inaudible speech, singing*

*Singing
- Under non-speech when signing does not contain words
- Under the appropriate language when containing lyrics

Precision of pauses
When the same person speaks in continuous speech, I still consider a 1sec segment of non-speech as part of the language currently spoken. Whereas, when the speaker or subject changes, I mark the pause (usually) as noise.

Txt files
I assume the txt format is the best for scripting, but if you need it in another format, let me know. I can split and export each entity (man, woman, non-speech, other) if that would facilitate your work. The way I've structured the script goes as follows:
- Identity of the interlocutor (man/ woman/ child/ other/ non-speech)
- Language or kind of sound or just *"child"* when a child speaks
- Beginning time of the utterance/sound
- End time of the utterance/sound

Each variable is separated by a semicolon from the next for easier parsing

Files not included in the transcription
I didn't transcribe audio files that only contain a few seconds of noise or inaudible files or files featuring the experimenter