Recording Audio for CommCare

Audio in CommCare can be used to help both the beneficiary or phone's user connect to messages on screen. This can be to guide an interaction for data collection or to provide counseling messages. Audio messages can often be longer than on-screen prompts, and can help serve as a "3rd party" expert for the people present. Audio can be recorded locally by people familiar with the program and incorporated directly into your application via CommCare HQ.

In order to minimize the time and data needed for installing an application over the network, it is essential to compress the audio to **56-**64**kbit mono ABR** encoding at **22 050Hz**, resulting in about **7kB per second** or **0.4MB per minute**. See the Audacity section below for more details.

File Types

Audio files must be in mp3 format. If your audio files are in another format, use this online tool to convert your files, at 128kbit/s. Then, import into Audacity to complete processing.

Guidelines for Good Recordings

Multimedia

Before deciding to include multimedia in your application, think carefully through what the goal of multimedia will be in your application. Some applications may not need multimedia (pure data collection or advanced users who don't need support).

Audio:

- **Writing a good audio script**:
    - **Audience**: Is the audio for the person using the phone or the beneficiary. This changes the message and phrasing of the audio messages.
    - **Counselling vs. Support:** Do you want to use audio to help user answer a question (for low literate users) or as counselling for the beneficiary
    - **Language and Dialect**: Try record audio from someone who speaks the local language/dialect. Use simple language that users/beneficiaries will understand.
- **Validate the audio script:** Before you start with the actual recording process, welcome feedback about the audio messages with your field team. Modify the phrasing of the audio messages based on feedback from FLWs, field staff and sector specific experts. Here are some things you can gather feedback on:
    - Verify local expressions being used are relevant, understandable and correct.
    - Confirm that the information in script coincides with field practices. If not, dispel any discrepancies.
    - Ensure comprehension of technical words, such as medical concepts.
- **Selecting the voice/speaker:**
    - Good qualities of a speaker include:
        - Native speaker of desired language
        - Clear voice and enunciation of words
        - Understanding of where to put emphasis in a phrase
        - Reads messages naturally
        - Speaks at a good pace (not too fast, not too slow)
    - Other considerations for persuasive behavior change communication include:
        - perceived influence or authority in certain kinds of voices (i.e. perceived education, or age of the speaker)
        - preference for male or female speakers
    - Ask speaker or a couple selected speakers to record a few messages. Compare the messages recorded by each and discuss with your team which voice(s) you would like to use in the application.
- **Prepare ahead with your speaker:**
    - Share the audio script with the speaker a couple of days prior to the recording if possible.
    - Let the speaker familiarize themselves with the text and give them an opportunity to ask questions.
    - If possible, make the person who developed the script available to answer questions.
    - In most cases all small discrepancies with the script is noticed at this point can be revised immediately before the real recording starts.
- **Time allocation for recording:** FLWs might not be used to doing such recordings, or are taking time off from their regular work to work with you. They need ample breaks between recordings. Given these reasons, recording may take longer than you expect, so allocate more time for recording than you might expect. Let the speaker know how much time is expected for the recordings. Plan breaks.

- **Equipment:**
    - We recommend you select a high quality microphone/recording device.
    - We discourage using laptop microphones. The audio is usually very poor and processing such as noise removal may not be able to improve the quality of recordings made through a laptop microphone.

- Some headphone microphones may be suitable.
- We suggest you test the quality of the audio recording for 1-2 files before purchasing the device and recording the full audio set. Test out the files on the mobile phone you will be using if possible.
- Some recording devices are highly sensitive and pick up background noise easily. If this is the case with your device, we suggest covering the mouth piece with foam or a cloth. This will help reduce background noise to a large extent.
- If your recording device records audio files in different audio formats, we suggest you switch the format to **mp3 mode** on the recording device itself. CommCare applications are compatible with the **mp3** format only. Recording the files in this format from the beginning, will save time later on.
    - If you are using existing media that you want to integrate into CommCare, there are online tools available to convert the audio files from one format into mp3 format.
    - If you forgot to save the audio files in the **mp3** format as you recorded them, you can change the format at the time of processing. Once processing is complete, you may export them in the **mp3** format.
- Record a short test take before starting to ensure equipment is functional and audio quality is good.
- Carry extra batteries on recording day!
- **Recording Set-up:**
    - Bring two printed copies of the audio the messages on paper: one for the speaker and one for the project staff to follow along. Project staff can listen for missed words or mistakes. You may also number the messages in large font for ease of reference.
    - The mouth piece of the microphone should be directly in front of the recording person, it should not be too close to her mouth also. The front face of the microphone should be facing the person whose voice is being recorded.
    - We recommend that the person who is recording the audio files (the one holding the device, not the person who's voice is being recorded), use headphones attached to the device to listen to the voice as it is recorded. We have found this is a helpful in determining the clarity of the recording, and will indicate to you whether any background noise or interferences were also captured. Be careful not to use a headphone that has a microphone attached to it, this sometimes creates a disturbance as two microphones are working simultaneously at the same time.
    - Depending on the pronunciation of the speaker you may need to adjust the position of the recorder. Generally, a 45 degree angle downwards from the mouth works well. Additionally, a 1 inch gap between the mouth and the recorder is recommended. Common problems with positioning of the equipment:

        - If recorder is too close to the mouth, you will likely see large sound peaks that are harder to remove for sounds like "bh", "ph" for example.
        - If recorder is too close to the mouth, you capture breathing in the script which is hard to remove or reduce in the processing phase.
        - If recorder is too far from the mouth, you capture other sounds (like your voice or breathing or those of people sitting around you) that you do not intend to.
- **Tips for the speaker during the recording:**
    - Have the speaker speak clearly and slowly — slower than feels natural, with good annunciation, discernible breaks between words, and plenty of pauses.
    - Have him/her speak slightly louder than usual ("project") but not so much that it sounds unnatural
    - Have them read each phrase in order with a short pause after each (~3 seconds).
    - Have them read the number (in English, if possible) before each phrase, with a short pause (~1 second) between the number and the phrase.
    - The numbers will aid greatly in identifying which phrase is which, especially if they were recorded in a language other than your own.
    - Do several takes. It's ok. Third time's the charm sometimes!

- **Recording Environment:**
    - Pick a quiet place with little background noise or disturbances.
    - Both recorder and 'speaker/talent' should remove all accessories which may interfere with the recordings (ex. jewelry, keys, phones). It will be very difficult to remove disturbances in the processing stage.
    - If recording outdoors, find shade. If in a room, find a well-ventilated room.
    - Make sure you bring sufficient water to the recording session for your speaker.
- **Processing**: For processing steps see Recording Audio for CommCare.
- **Managing a large number of audio files:** There are many ways one could manage a large number of audio files. Here are some tips to make it easier:
    - Have the speaker or the recorder say the number or title of the audio message that is being recorded at the beginning of each message. Or you can create a sound effect (i.e. a clap or tap on the table) during the recording itself that will denote the start and end of recording different messages or the perfect message. You will visually see this sound peak at the time of processing.
    - Approach 1: Record all the phrases in one take (one audio file). Don't use a separate file for each phrase. If the recording is interrupted with background noise and the speaker messes up, let the recorder keep running and continue on when possible, starting with that same phrase.
    - Approach 2. Record a separate file for each audio message in your script. Read out the message number at the beginning if you do this.
    - Regardless of which approach you take, first splice out the best audio recordings for each message. These are your rough cuts. Name them according to the question ID/keyword in the Definition File. You can do individual or bulk processing of audio files in Audacity. Instructions here: Recording Audio for CommCare

If using Zoom H1 Microphone

- Carry extra batteries.
- Turn On: Slide power switch down for 1 second
- Turn Off: Slide power switch down for 1 second
- Reducing Noise: Back of Recorder – Low Cut On will reduce any wind or background noise.
- Input Level: Can be automatic by switching Auto level – On (back of recorder). Or can be done manually (recommended settings to come)
- Output Level: Volume that will come through your headphones. Manually adjusted near headphone input.
- Recording Format: Wave Format is higher quality sound than MP3. Change the Bit Rate using the arrows (this will only change if SD card is inserted). Recommended settings would be 48/16 or 48/24. Higher bit rate results in highly quality audio recordings, however, decreases the recording time of the card. Depending on the length of the audio recording, the bit rate can be adjusted appropriately.
- Listening to Recordings: After recording, you can use the play button on the side to listen to any recording. The side arrows will allow you to choose which recording you would like to listen to. As the recording plays, the remaining time of the recording will be shown. The playback can be heard using your headphones or simply through the recorder.

- Deleting Recordings: While the playback is running or when it is complete, you can press the trash key to delete. Press the Record button for confirmation of the deletion. A message 'Done' should appear once the deletion is complete.

Processing Audio Using Audacity

Processing audio for CommCare involves five easy steps which include: (1) Splicing; (2) Background Noise Removal; (3) Blank Noise Removal; (4) Pauses; and (5) Amplification. Steps 1-3 can be done for each clip individually and steps 4 and 5 can be done in bulk. Please also see our video demonstration on how to use Audacity for processing to help you along the way!

Audio Processing Tutorial English	Audio Processing Tutorial French

Instructions

1. First install Audacity and follow the instructions for downloading. You will need the LAME mp3 encoder to use audacity. You can download it for free by searching for *lame_enc.dll.*

2. Configure the MP3 encoding settings in Audacity (Edit -> Preferences -> File Formats -> Bit Rate) (Please read *Note* below for newer versions of Audiocity). For speech, **56-**64**kbit mono ABR** encoding at **22 050Hz** should give excellent results. If the audio contains other noises or music, 96kbit mono could be considered. For very high quality applications (at a minimum, CommCare user will be using headphones) use 128kbit stereo. 64kbit mono requires **~7KB per second of audio**. We use average bit-rate encoding (better quality for a given file size and preferred over variable bit rate/VBR for low bit rates). Set the project's frequency to 22 050Hz in the bottom left corner of Audacity. See a discussion on encoding choices, including for voice.
   Note: In newer versions of Audiocity 2.0 and higher the option for MP3 encoding has changed to the export step. File -> Export -> Save as type -> MP3 -> Options Button.

3. Copy the files from the memory card in the recording device and save on your computer.

4. Open the recording file in Audacity.

5. **Splicing:** If you recorded a string of audio messages in one audio file, your sound peaks may look like the image below. If you decided to use unique sound effects like a clap or a tap on the table as a way to denote the start and stop of audio messages, then you will be able to visually decipher the different message segments. If you recorded one audio message in one file, you will hopefully have a shorter file.

   a. Play the file and find the best recordings for each audio message.

   b. Select the portion of audio (highlighted below) that you wish to process.

   c. If your file contains all the audio messages, then copy and paste this best audio recording for one message into a "new project" and complete your processing there.

   d. If you recorded separate clips for each message, then you may find it easier to process within the same file. Delete the recording segments that you do not want to keep.
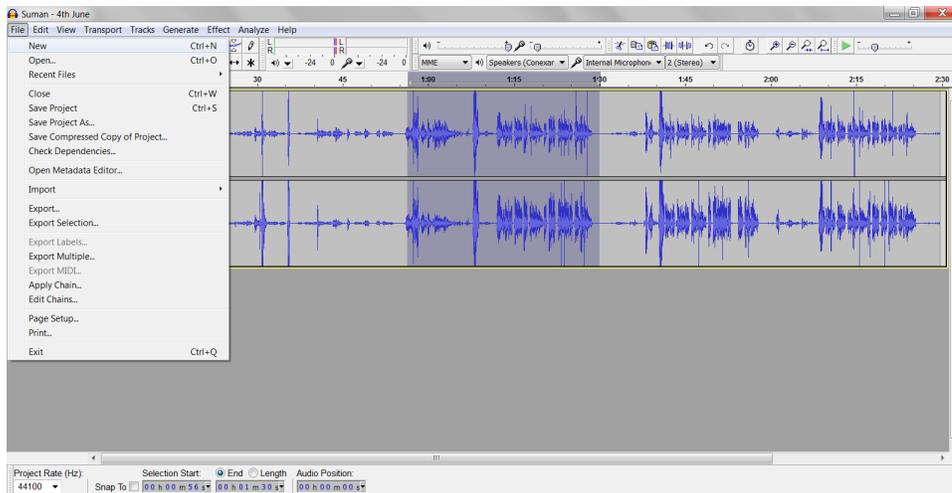
6. **Background Noise Removal:** Listen to the audio message. It is best to remove background noise to ensure we have good quality messages that are played from the application. See this video from 6:54 to 10.48 seconds to learn how to remove background noise. **NOTE:** This may not be the best way to do it and it can be time consuming, so the better thing to do is, obviously, be careful while recording to ensure there are no background noises.

7. **Blank Noises:** Now check for any blank noises in the audio clip and delete them. The objective is to keep the audio recording crisp and precise. To do this, drag the cursor and select that portion of the clip and press delete.

8. **Save/Export Individually:** Save this audio file as .mp3, click on File, **Export as mp3.** Save the file name as the questionID in the application. You may have to reference the definition file. (You can also select the processed portion of the audio clip in a larger file and select **Export Selection as mp3.**) Save the processed clips in a folder called "audio".

   This is the end of processing clips individually. You can complete the last two processing steps in bulk to save time. To begin, select all the mp3 files in your "audio" folder and drag them into Audacity. You can work in smaller sets of 10-15 files instead of copying hundreds of audio clips into Audacity.

9. **Pauses:** Note it's recommended to add a slight 0.5 - 1 sec delay at the beginning of each audio clip. Don't make the delay too long, otherwise FLWs will be inclined to try to re-play the message by hitting the hash button twice. This will actually pause the audio clip! To add intro pauses, place the cursor at the very beginning of your audio clip. Click on Generate Silence and enter the preferred seconds.

10. **Amplify:** The next step is to amplify the volume of the audio file because audio played in the field by FLWs needs to be audible by beneficiaries, in an environment that is prone to a lot of noise (i.e. the farm, a health center, a home with a crying baby, goats and chickens).

    a. Select the sound peaks of all of your audio files in Audacity
    b. Go to Effects and choose Amplify
    c. In the pop-up box, input 10 in the "amplification" field and below, check the box "allow clipping" (see pic below) and click Ok.
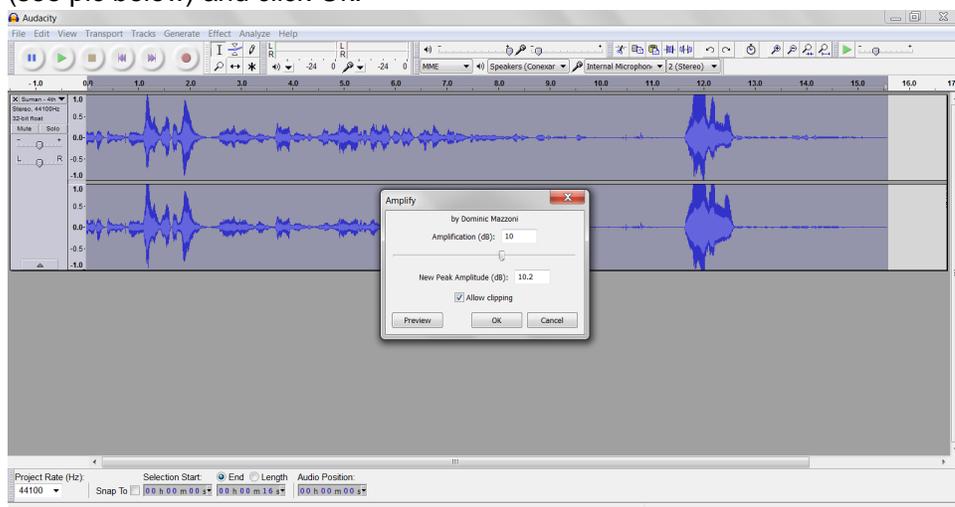
    - You will notice the audio file will have vertically stretched out bars indicating that the volume has been amplified.
    - If the audio is going to be played in a loud environment, you may need to increase the audio volume slightly more. More amplification is better than less amplification. The FLW can reduce the volume on the device if its too loud, but wouldn't be able to increase it if the audio was not amplified enough to begin with.

11. **Save/Export in Bulk:** Now that you have processed audio in bulk, you can export in bulk too. Select File > Export Multiple.

Post-processing

These steps are optional.

Making the volume equal using MP3 Gain

These steps makes all the recordings approximately the same volume. First download and install MP3Gain.

Graphical Instructions

1. Open MP3Gain and choose "open file/folder" and open all the clips you want to use.
2. Then do Gain --> Apply Constant Gain. Configure and tweak as needed.

Command Line Instructions

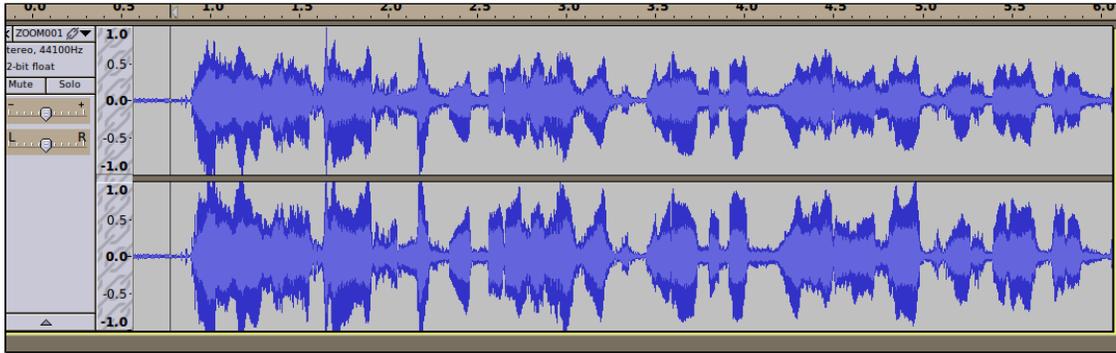**mp3gain -r -c -d 10 *.mp3** (assuming all the mp3s are in the current directory)

The -d 10 is a volume boost (here, 10dB) to give to all files *after* they have all been normalized to the same volume. This is because the default volume level tends to sound quiet on the phones. Tailor the amount of boost to your deployment and the devices you will use). Each 10 dB of boost approximately doubles perceived loudness.

Clipping Audio files in Audacity

You can also clip audio files in Audacity to control the loudness of the file. See below an example of good and bad clipping. If you recorded audio that is really loud or really soft, you might want to manually use the clipping feature.

Don't boost too much or clipping will occur (the stength of the signal is boosted beyond the maximum of what the sound file can represent; the rest is 'clipped' off). Excessive clipping will sound harsh and severely degrade sound quality. You can view the amount of clipping in audacity.

ok clipping

bad clipping