UDSv4 Digital Voice Data Collection Manual

Created in collaboration with the National Alzheimer's Coordinating Center (NACC), the Clinical Task Force (CTF) Technology Workgroup, and the Framingham Heart Study Brain Aging Program at Boston University. 2024.

Table of Contents

Introduction

Uniform Data Set Version 4 (UDSv4) will give centers the option to collect digital audio recordings of the cognition section of the UDS study. NACC and the Clinical Task Force (CTF) Technology Workgroup are collaborating to support Alzheimer's Disease Research Center (ADRC) adoption of digital voice as part of the UDSv4 rollout. At this time, digital voice capture is highly encouraged but not required.

Why digital voice?

The implementation of digitally recording participant responses to neuropsychological tests is a cost-effective way to detect early changes in cognition. As our cognitive capabilities shift, we express them through vocal responses in subtle ways, such as changing word choices or sentence structures because of word finding problems, pausing, hesitating, and shifting as memory, attention, and executive functions are compromised.

Currently, there are no gold standards in methods for analyzing voice recordings in relation to cognition. However, just as with blood-based biomarkers, there is a growing, albeit still limited, set of literature suggesting that analysis of digital voice recordings as a method for differentiating those with and without cognitive impairment is promising.

Benefits of digital voice

- **A non-diluting resource:** Digital data can be repurposed for different purposes as algorithms and analysis techniques improve.
- **Minimal participant burden:** NP tests are already being conducted; digital voice collection allows for scientific enablement at no additional burden to participants.
- **Low cost and inclusive:** Penetration of recording devices allows for easy, low-cost collection of voice data that can be done in the person's native language.
- **Novel analytics:** Natural Language Processing (NLP) and advanced machine-learning methods offer opportunities to explore acoustic and semantic features in novel forms.
- **Quality control (QC):** Digitally recorded voice tasks can act as a QC tool to determine natural drift in standardization in any longitudinal study.

About This Manual

This data collection manual is intended to be an in-depth guide for Alzheimer's Disease Research Centers (ADRCs) in the process of collecting, naming, storing, and processing audio recordings of the Uniform Data Set (UDS) Version 4 cognitive battery.

We highly recommend ADRCs utilize the UDSv4 Digital Voice Start-up Checklist to get started with collecting digital voice data at their center.

At this time, digital voice capture is highly encouraged but not required.

Created in collaboration with the National Alzheimer's Coordinating Center (NACC), the Clinical Task Force (CTF) Technology Workgroup, and Framingham Heart Study Brain Aging Program at Boston University. 2024.

This document is adapted from the manual published in collaboration between the Framingham Heart Study Brain Aging Program at Boston University and the Indiana Alzheimer's Disease Research Center. The content for which was adapted from the manual published by the ADRC Clinical Task Force Cognitive Working Group.

Step 1: Regulatory Compliance

Informed Consent

It is crucial to provide an in-depth description of the audio recording implementation and use of the data as part of the informed consent process. In addition, it is important to provide detailed information regarding:

- **Purpose:** Clearly explain the reason for the audio recording and how it will be used.

  Helpful Suggestion!
  Keep the purpose and details of the study procedure general and non-specific to avoid limiting yourself in the future.

- **Confidentiality:** Describe how the recordings will be kept confidential and the measures in place to protect a participant's privacy.
- **Data Sharing:** Specify who will have access to the recordings and who the data will be shared with.
- **Benefits and Risks:** Explain any potential benefits and risks associated with the audio recording.

Additional consent is required if the audio recording is not an essential part of the study. Consider adding the following language to your consent documentation:

Do you give us permission to audio record your responses to some of the memory and language tests?

☑ Yes, you have my permission
☐ No, you do not have my permission:

IRB Protocols

Consult with the Institutional Review Board (IRB) at your site to determine if your IRB will need to be updated to collect and share digital voice data. It is important to provide detailed information regarding:

- **Objective:** Clearly state the purpose of the study and the role of audio recordings in achieving the study objectives.

IRB Protocols (continued)

- **Storage:** Describe the measures that will be put in place to protect participant confidentiality.

  Helpful Suggestion!

  Privacy and security of data should be protected in the same manner as other data types like imaging and genotyping. Such as local storage in HIPAA compliant solutions, access limited to IRB-approved study staff, and data transmitted securely to NACC.

- **Recording Procedure:** Specify 1) what will be recorded, 2) the setting and method of data collection, and 3) the equipment being used to record.
- **Data Sharing:** Specify who will have access to the recordings and who the data will be shared with.

PLEASE NOTE: De-identification tools that will be used for de-identifying digital voice data are still being reviewed and developed. This data will not be shared by NACC until we are confident that the data is de-identified.

Example Language for IRB and Informed Consent

Digital voice data capture has been successfully adopted and implemented within the ADRC program through research studies such as the Framingham Heart Study Brain Aging Program (FHS-BAP), Longitudinal Early-onset AD Study (LEADS), and Internet-Based Conversational Engagement Clinical Trial Study (I-CONECT).

The excerpts below, provided by groups who have successfully navigated regulatory compliance, can help researchers develop their own language to include in IRB and consent forms for digital voice data collection and sharing as part of UDSv4.

Purpose or Objective

> "We will audio record your responses to some of the memory and language tests. Recording your responses will allow researchers to write out (transcribe) your responses exactly as they were spoken. These audio recordings and transcriptions will allow researchers to analyze your responses to answer important questions such as behavioral markers of speech, language and cognitive decline." - *LEADS Study*

"The administration of select language and memory tests will be audio recorded for every participant, at every visit. This data will be collected to allow for additional analyses." - *I-CONECT Study*

Recording Procedure

"Prior to beginning the consent process, participants will be asked for their permission for the investigators to audio record the consent process. These audio recordings of the consent process will be used for quality control purposes only. The investigators will also ask permission to audio record the study examinations as well. Audio recordings of study exams will be used for research analysis. Participants will also be asked if a portion of the motor exam can be recorded for research purposes. The consent process will not be video recorded. If a participant does not agree to audio and video recording, the recordings will not be done. Participants have the right to refuse being recorded at any time during the exam." - *Framingham Heart Study Brain Aging Program*

"This session will be recorded using a digital audio recorder. Recordings will be analyzed in conjunction with other study information. We will also use recordings to make sure that your responses are accurately documented." - *Framingham Heart Study Brain Aging Program, within the study description*

"The administration of select language and memory tests will be audio recorded for every participant, at every visit." "A recording device, such as the Olympus VN-722PC digital voice recorder or a similar digital recording device, will be used." "The I-CONECT video chat system is a secure, scalable, and cost-effective video chat platform for conducting in-home telehealth with geriatric populations." - *I-CONECT Study*

Confidentiality

"Your recordings will be maintained on secure servers used for this study and only those with privacy training and permission from the study team will have access to them." - *LEADS Study*

"We will protect your privacy when sharing audio recordings by only sharing the recordings with approved researchers. We will not share any parts of the audio recordings that contain identifying information like your name, telephone number, or other information we think could be used to tell your identity." - *I-CONECT Study*

Storage

"The audio files acquired in this study will be securely transmitted and stored in databases being used for the study. Sites will be responsible for ensuring that no PHI is included in the recordings prior to upload." - *LEADS Study*

Data Sharing

"In the future, your recordings and information may be given to researchers for other research studies. However, you will not be identified in any of these recordings. Your data will be given a unique study identification number." - *Example language developed by NACC*

Benefits and Risks

"Use of this method of interview is expected to have minimal risks. The risk involves potential for breach of confidentiality. The risk is minimized by using a secure web platform that is used by many hospitals and clinics for doctors to communicate with patients, hence we believe it will be secure for this research interview. The visits will be digitally recorded (1x participant/month will be video recorded via Zoom recording and each participant will be audio recorded) and kept on FHS servers similar to our audio recordings of in person cognitive testing. The video recordings will be retained for short-term storage, long enough for QC purposes, and will then be destroyed/removed. The FHS forms are filled and stored at the main FHS facilities along with all other FHS records. Participants can always decline to answer any question or decline to complete any test even if the participant consents and completes to the rest of the questions or test within the examinations." - *Framingham Heart Study Brain Aging Program*

"However, because everyone's voice is unique, it's possible that someone listening to the recording could identify you from the sound or pattern of your voice." - *I-CONECT Study*

Virtual Visits

"Remote Neuropsychological Examination (Remote Cognitive Testing): You will be offered the option to continue participating remotely if you are unable to participate in-person. If this occurs the following additional procedures will take place.

Before Your Virtual Visit: The virtual visit/remote testing will take place using your electronic device. You will need to have steady access to the internet and will use your own device (computer/laptop/iPad/tablet). There will be technical support for any questions prior to or during the encounter. The remote testing will be done using a BU approved platform for interviews such as Microsoft Teams and/or BU HIPAA Zoom.

During Your Virtual Visit: You will be asked similar questions and administered the same tests that you would encounter during an in-person neuropsychological exam. The tests given virtually will mimic in-person tests as much as possible with modifications only made to facilitate virtual use. You may be offered a second in-person evaluation at your convenience, in your home or at our office, in the future." - Framingham Heart Study Brain Aging Program, language for virtual visits

Additional Examples

The FHS-BAP Informed Consent Form contains the following questions and statements about audio recordings:

Audio and Video Recording: The neuropsychological and neurological examinations will be audio recorded. The motor assessment of the neurological examination and remote neuropsychological exam will be video recorded. The audio and video recordings will be analyzed in conjunction with other study information.

Do you agree to have the neuropsychological and neurological examinations audio recorded?

_____YES_____NO _____N/A _____INITIALS

Do you agree to have the neuropsychological and neurological examinations video recorded?

_____YES_____NO _____N/A _____INITIALS

Step 2: Identify Participants

ADRCs are encouraged to collect digital voice recordings of the UDSv4 cognitive battery on all eligible participants.

The following eligibility criteria apply:

1. Ensure your participants are adequately consented (see Part 1)
2. All participants must be enrolled at your ADRC
3. All participants must have a NACCID
    a. If participant has a submitted and finalized UDS packet, they have a NACCID
    b. If participant does not yet have a submitted and finalized UDS packet they may need to be assigned a new NACCID – learn more here

Step 3: Select Audio Recording Equipment

The audio equipment selected by a researcher will greatly affect the quality of recordings and can have implications for data analysis down the line. In this section we have outlined the important details to consider in this selection process.

Factors to consider when choosing a recorder:

1. **Portability** (if off-site testing is done)
2. **Compatibility** with lab computers (Mac/PC/Linux)
3. **Ease of setup:** This will help inform your site's training process for test administrators.
4. **Data storage space:** Determine whether the recorder can collect multiple sessions before downloading the files to a computer or whether it needs to be downloaded after each testing session.
5. **Battery life & power options:** We highly recommend finding a recorder that can capture an entire session's digital voice recording using battery power. NOTE: It is recommended that you use the recorder in battery mode as some users have experienced power source background noise in their recordings.
6. **Microphone:** We recommend choosing a recording device that can lay on the table in front of the participant.
    a. Lapel mics, which clip to the participant's shirt, may pick up too much rustling noise if the participant moves.
    b. Headsets might be too cumbersome or uncomfortable. Hanging (ceiling mounted) mics are not portable.
    c. If using external mics, avoid using condenser microphones (they pick up more room reverb); use cardioid mics instead.
7. **Sampling rate:** The sampling rate is one measure of audio quality, expressed in Hertz (Hz). The lowest acceptable sampling rate is 16,000 Hz, and we recommend capturing audio at a sampling rate of 48,000 Hz.
8. **Format/encoding:** We highly recommend collecting voice data in the WAV format with LINEAR16 PCM encoding with at least 16000 Hz sampling rate.
9. **Quantity:** Your center may want to acquire multiple devices to accommodate multiple simultaneous testing sessions, offsite testing, or lost/broken devices.

10. **Additional equipment:** The recording device may require an SD card, charger, and/or a cable to connect to a computer for download.

Recommended Device(s)

The recommended device for in-person recording is the Zoom H4n Pro, along with the accessories below:

- SD Card (e.g., SanDisk 32GB Extreme PRO UHS-I SDHC Memory Card)
    - Provides additional storage on the recorder.
- Zoom AD-14 AC Adapter
    - Power cable for the recorder that can be used in lieu of batteries.
- USB to Mini USB Cable
    - Cable that connects between a computer (USB) and the recorder (Mini USB).

See Appendix B for detailed instructions on the Zoom H4n Pro setup and usage. Additional recording device recommendations can be found in Appendix C.

Step 4: Setup Equipment & Testing Space

Recorder placement and the sound quality of the testing room will greatly impact the recording quality. Sound reverberating off bare walls & floors, humming equipment, external noises (e.g., loud colleagues, birds chirping, etc.), and microphone direction are some of the factors that can diminish sound quality of voice recordings. The quality of audio can also be affected if the participant and/or examiner are wearing masks. It is recommended to track whether masks were worn for each testing session and if the participant, examiner, or both were wearing a mask.

A test recording in every testing room should be performed beforehand to confirm the quality of sound recording for data extraction.

If you have the option to make changes to the testing environment, consider the following:

Ideal Testing Room

1. Small to medium size
2. Multiple soft surfaces like carpet, couches, pillows, etc.
3. Avoid rooms that have a lot of hard surfaces that will make sound bounce around, such as windows, bare walls, and hard floors
4. Minimal exposure to external sounds (e.g., street noise, a conference room, loud colleagues, phones ringing, plumbing, weather)
5. Turn off noisy things in the room (e.g., fan, phone, air conditioner, computer in overdrive)
6. Lay a towel or piece of cloth under the recorder

Placement of recorder

1. Point the mic(s) of the recorder toward the participant (and away from tester)
2. Place it in a location where it will be out of the way of testing (once you start recording, you don't want to be moving the recorder around)
3. If possible, place the recorder on furniture that is NOT the desk/table you are working on (because sounds such as pages turning, bangs on the table, etc. get picked up), but be sure it is close to the participant

Sound-treating

The recording quality can be improved by sound-treating the testing room(s):

1. Floors
    a. Carpet/rug
2. Ceiling/walls

a. Bass Traps
b. Acoustic Panels
c. Alternatively, can use packing blankets or mattress foam
d. Or do it yourself (DIY)
    i. How to build a sound absorbing panel in 5 easy steps
    ii. How to build your own acoustic panels
    iii. Budget Audio Treatment
    iv. Cheap Sound Treatment Tests in a Commercial Office
    v. How to install acoustic foam without damaging your walls
    vi. Tips for DIY

Step 5: Train Your Tester Administrators

Ensure every tester is trained in how to use the recording equipment, how to properly save recordings and data logs, and what to do with the audio recording(s) after the testing session(s).

Recommendation: The test administrator should take time before the participant session to determine the optimal gain levels and microphone distance. Gain is the amount of amplification applied to an audio signal before it's processed, and it's measured in decibels (dB).

> EXAMPLE: When the gain is too high - the audio becomes “overdriven” and gets clipped. The gain meters (if available) should stay mostly in the green zone with very occasional yellow/red splashes when the participant laughs, coughs or exclaims something.

Step 6: Collect Digital Voice Data

Following the audio recording device and testing room setup parameters above, collect participant voice recordings during the UDSv4 cognitive battery.

We recommend centers record each cognitive test as a separate audio file. Centers still have the option to record the entire UDS4 cognitive battery in a single recording. In a later section, we provide instructions on how to split recordings by test using post-processing software such as Audacity or mark the timestamps for each individual test – see post-processing section below.

Some tests are virtually silent, such as Trails, but we encourage researchers to continue recording even during these silences. Audio recordings can be used to analyze many different angles of testing, so it can be useful to record during quiet tests to catch if people speak or make noise during the test (such as verbalizing during Trails).

General Digital Voice Recording Process

Prior to conducting the cognitive test

1. Verify that the recorder's battery is full, regardless of whether an AC adaptor is being used. This will help to prevent data loss.

2. Turn the recorder on (The Zoom H4n device has a switch on the left panel of the device, turn this to "ON")

3. Confirm all settings are correct
    a. Recording format = Uncompressed WAV file is recommended
    b. Sampling rate: 16KHz at a minimum (48KHz is recommended)
    c. Bit depth: 16-bit at a minimum (32-bit is recommended)
    d. Gain level = Gain is the amount of amplification applied to an audio signal before it's processed, and it's measured in decibels (dB). When the gain is too high - the audio may become "overdriven" and get clipped.

4. Lay the recorder down with the head of the recorder pointed directly at where the participant will be seated. If you are not in a testing room, place the recorder on a

different surface, but still close to the participant, so it does not pick up all the paper shuffling, table jarring, etc.
PLEASE NOTE: It is essential to place your paper holder on the opposite side of the table relative to the recorder, because the recorder is very sensitive, and paper shuffling will muddle the audio.

5. **(Optional QC measure)** Use the Play/Pause [►/||] button to record the following information:
    a. NACC ID
    b. Date of testing
    c. Examiner Initials

6. **(If applicable)** Put the recorder in "Recording Standby Mode" by pressing the "REC" button. Recording standby means the mic is on but is not yet recording.

During the cognitive test

7. After the participant has been consented and has signed the consent form, you may begin recording the examination.
    IMPORTANT NOTE: If the participant reports that they do not want to be audio recorded, turn off the recorder, remove it from the table, and proceed with testing (unrecorded).

8. When ready to begin the cognitive test and recording, start recording by pressing the Play/Pause [►/||] button.

9. After each test is complete, stop the recording.

10. Repeat steps 8-9 for each cognitive test.

    PLEASE NOTE: We recommend centers record each cognitive test as a separate audio file. If this is not conducive to the testing environment, centers have the option to record the entire UDS4 cognitive battery in a single recording.

Step 7: Save Audio Files & Data Logs

After testing, the examiner will download the recording from the recorder to a computer and save it.

General Naming and Storage Process

1. Connect device to a computer (e.g., the Zoom H4n Pro uses a USB cable)

2. Save the file in the appropriate file location using a standardized file naming convention for unedited recordings. See Appendix B for detailed instructions on the Zoom H4n Pro setup and usage.

   Recommended Naming Convention: NACCID_DATE_COGNITIVETEST.wav

3. It is imperative that centers maintain a data log with the following recommended metadata variables for each recording:

   View Digital Voice Metadata Dictionary

   Download template data log file

Minimum recommended data variables:

- NACCID
- Visit date and number
- Cognitive tests (NACC Code) and timestamps for each cognitive test recorded (if captured in a single recording)
- Whether the participant and/or interviewer were wearing masks
- The setting in which the visit took place
- Microphone location
- Device manufacturer and model
- Number of voices present in the recording and whose they are

If your center is processing data before uploading it to NACC, these additional variables should be tracked in the data log:

- If PHI was removed
- If recording was spliced
- Date of processing
- Processing program used

- Type of quality check(s) performed
- Whether recording passed quality check
- Why QC failed, if failed

Step 8: Post-processing Digital Voice Data

This step is optional, but highly recommended. Centers are not expected to process voice recording data; we want to encourage sites to get started with collecting and properly storing digital voice data for future processing and research analysis.

Researchers can choose their preferred processing software as long as it can silence PII in such a way that it cannot be reversed and can save audio files in an uncompressed (e.g., WAV) format.

After processing the recording, the edited audio file should be saved with a standardized file naming convention and corresponding details should be entered in the data log for sharing with external investigators. **Do not save over the unedited audio file, save as a new file**.

Recommended Software

We recommend using the software Audacity, which is a free audio recording and editing software for Windows, Mac, and Linux. Researchers can download Audacity here: https://www.audacityteam.org/download/

Audacity can be used to record audio with a connected microphone, or researchers can use a separate device to record and then download the recording to a computer that has the Audacity program. To open an audio file in Audacity, you can drag-and-drop the file or open Audacity, go to **File** > **Open,** and navigate to the appropriate file.

There are many online tutorials and instructional videos for Audacity users, so looking up your questions online will usually yield a solution.

For an introduction to the program, watch the first 4 minutes of this video: https://www.audacityteam.org/download/.

(Recommended) Flagging Personally Identifiable Information

Personally Identifiable Information (PII) is information that can be used to identify, locate, or contact a single individual. **It is essential that all PII is removed from the recording prior to sharing with outside investigators.** Examiners should try to avoid using a participant's name during testing; however, it is not uncommon for a participant to say something in the middle of testing that would be considered PII and therefore must be removed.

Procedure for Testers to Flag PII

PII may not occur often during the recorded testing, and testers can limit PII by collecting all participant-related information before beginning the recording and being mindful of speaking PII such as referring to the participant by name.

It is extremely important that the examiner pays attention to every time PII is spoken by either themselves or the participant.

- Upon speaking/hearing any such information, the examiner should mark the active battery page at the time of the PII.
- If the examiner is not using paper for the test (for example, the use of a computer or tablet instead), a standardized way to mark when PII is spoken should be decided upon.

All examiners should use the same way of marking PII, in case the examiner who conducted the tests is not the same person who removes PII from the recording.

Types of PII to Flag

1. 18 HIPAA identifiers[1] of the individual or of relatives, employers, or household members of the individual:
    a. Name (including maiden name)
    b. All geographic subdivisions smaller than a state, including street address, city, county, precinct or neighborhood area, ZIP code, and their equivalent geocodes.
    c. All elements of dates (except year) for dates directly related to an individual:
        i. Birth date
        ii. Admission date
        iii. Discharge date
        iv. Date of death
        v. All ages over 89 (as well as the year of birth for this age group)
    d. Telephone numbers
    e. Fax numbers
    f. Email addresses
    g. Social Security numbers
    h. Medical Record numbers
    i. Health plan beneficiary numbers
    j. Account numbers
    k. Certificate/license numbers
    l. Vehicle identifiers (e.g., serial numbers, license plate numbers)
    m. Device identifiers and serial numbers
    n. Web URLs

Types of PII to Flag (continued)

o. Internet protocol (IP) address
p. Biometric identifiers, including finger and voice prints
q. Full face photographic images and any comparable images
r. Any other unique identifying number, characteristic, or code

2. Research-related Identifiers
    1. Start-of-Exam Recorded Identifiers: Participant ID, tester ID, and date

3. Regional Identifiers
    1. Schools attended
    2. Place of work
    3. City of birth

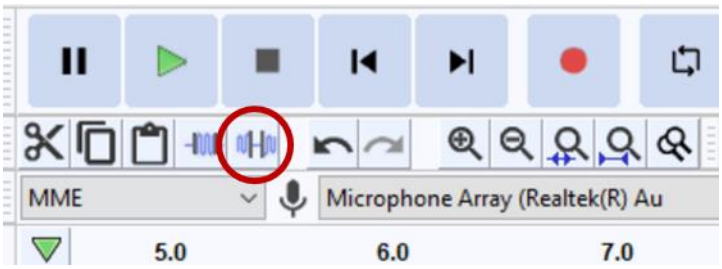
Labeling and Silencing PII in Audacity

Audacity has a Silence tool that you can use to remove Personally Identifiable Information (PII) from the recording. To learn how to use the tool, watch this video (https://youtu.be/VgI6PUNv0fY) and then follow the steps below.

- After opening the appropriate audio recording in Audacity, go to the drop-down menu at the top of the program and find **Tracks.** Select **Add New** > **Label Track**. You will use the label track to mark when PII occurred in each recording.
- Listen to the audio to find the first occurrence of PII in the recording. If you click and drag on the audio track, you can highlight a portion of audio; if you press play, it will only play the highlighted section of audio. (This also works if you highlight part of the label track.)
- You can use the Zoom tools to zoom in on the audio: 
- Once you find PII, click and drag on the **label** track until you have highlighted the area with PII (you can press play to check). You can click and drag the start and end points of the highlighted section and keep replaying the segment until you have isolated the PII.

Tip: It might be hard to avoid including words on either side of the PII: for example, if someone says "Yep, my brother's name is John Smith and oh, um..." - It might be hard to not also grab when they say "is" and "and" on either side of "John Smith", depending on how fast they speak. We want to limit non-PII speech in the segments, but don't spend too much time trying to avoid capturing a word or two on either side of the segment. Overall, the idea is "do your best effort" on limiting non-PII speech.

- Press **ctrl + B** or **cmd + B**. This will create a label at the location you have highlighted. Type “PII”.
- When you click on the “PII” label you made, it will highlight the selected audio. Click on the Silence icon or use **ctrl+L** to silence the selected audio:



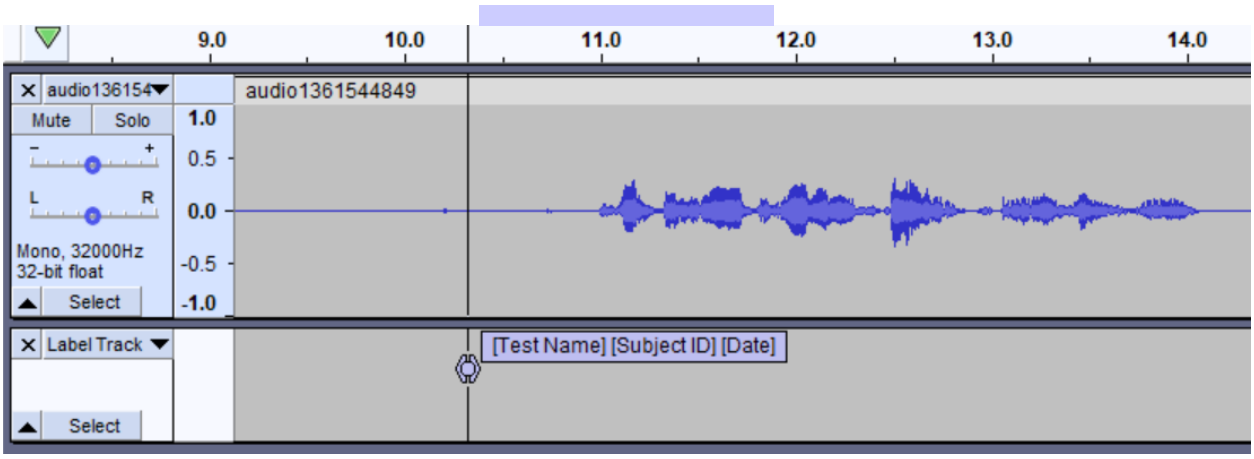
- Repeat these steps with each segment of PII until all are labeled and silenced.

Using Labels to Save Cognitive Tests as Individual Files

If your ADRC chooses to record all cognitive tests in a session as a single audio file, it is recommended that you use labels to mark or save cognitive tests as separate files as this is highly conducive to later data processing and analysis.

You can use Audacity to easily split and save each test as an individual audio file. Here is a video explaining this feature (https://youtu.be/72ewbraagj8).

- At the top of the Audacity program, find the drop-down menu for **Tracks**. Select **Add New > Label Track**. (Researchers can also use a label track to save timestamps of PII. This step will create a second label track. **It is important to use separate label tracks for the two tasks**.)
- Find the beginning of the first cognitive test. Click on the second label track so the vertical line is positioned before the first test begins.
- Press **ctrl + B** or **cmd + B**. This will create a label at the location you have selected. (If it creates a label on the PII label track instead of the cognitive test label track, it’s because you need to click on the second label track before pressing **ctrl + B** or **cmd + B**.)
- Type in using the standardized naming convention (this will eventually become the audio file name). Each test should have a different Accession # with the corresponding test noted in the data log.



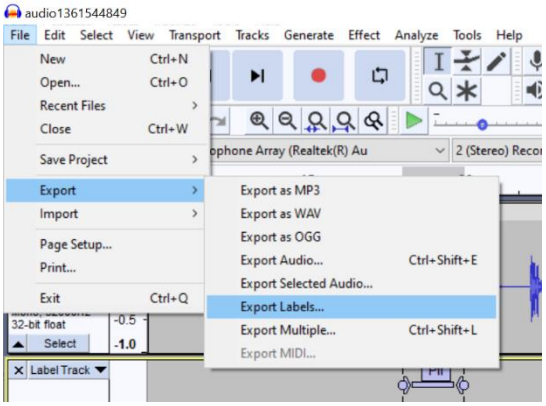
- Find the space between the end of the first test and the beginning of the second test. Click on the second label track at that location and use **ctrl + B** or **cmd + B** to create another label. Name it appropriately. Do this at the beginning of every test.

Saving Processed Audio Files

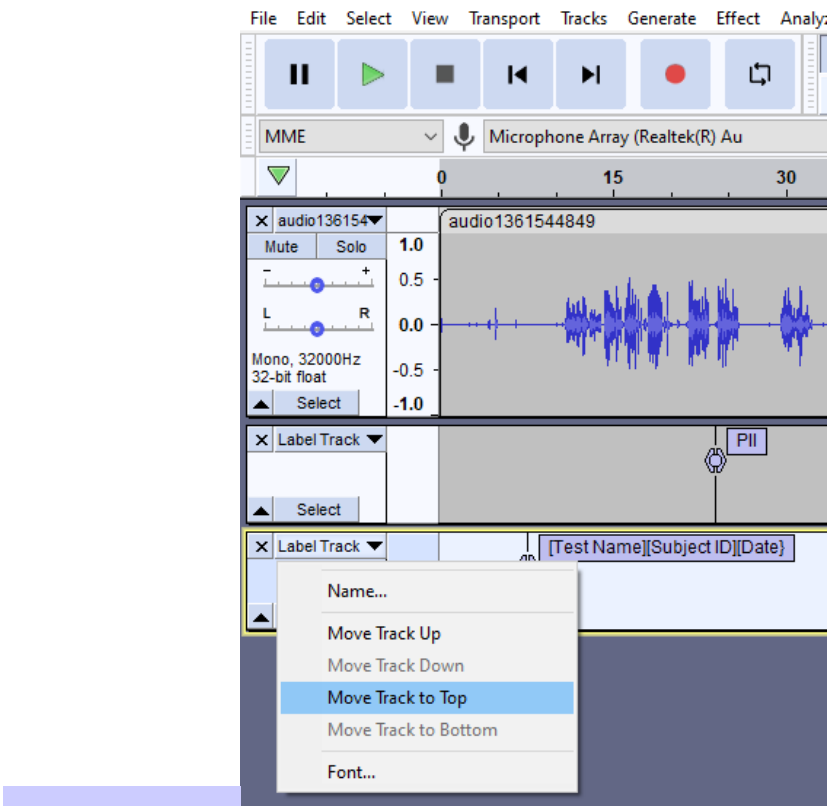
After the PII and cognitive tests have been labeled and the PII is silenced, you need to save the labels and audio files separately.

To do so, follow these steps:

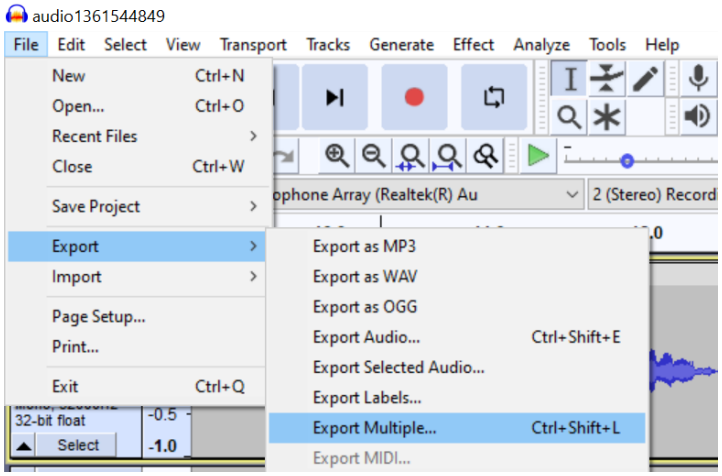
- To save the timestamps, click on **File** > **Export** > **Export labels**



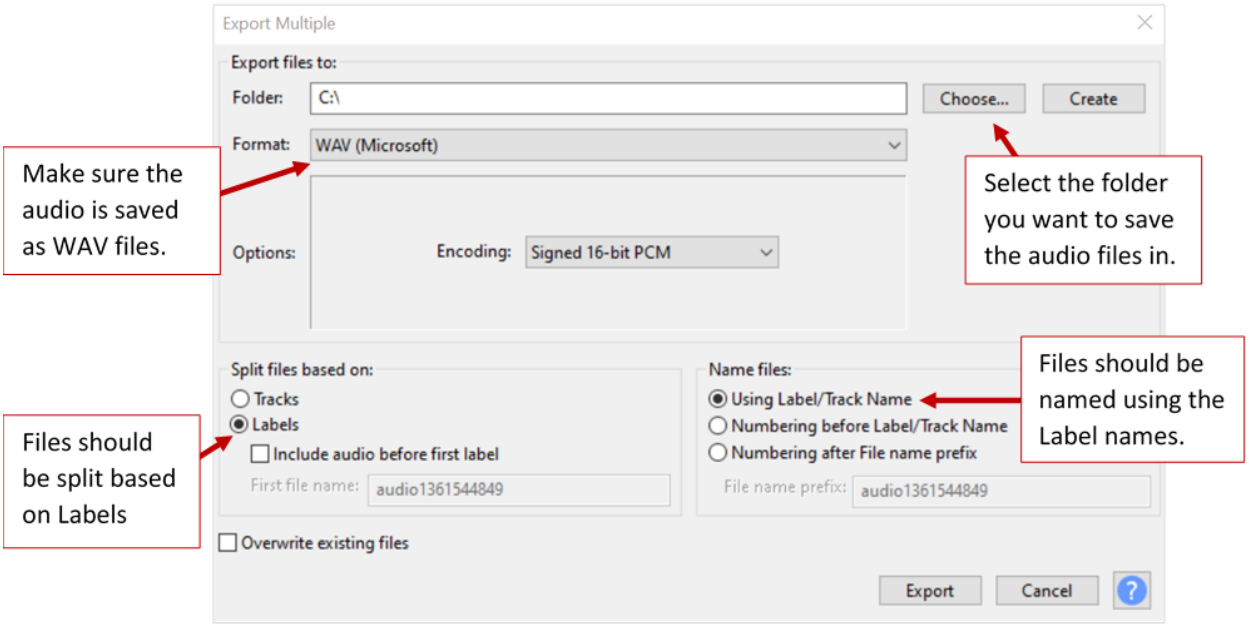
- Save the timestamp file (.txt is the default file type) in the designated file location with the standardized naming convention corresponding with the audio file. The text file will contain timestamps for both PII and the cognitive tests.
- Now it is time to save the audio files. Go to the label track that has labels for the cognitive tests. On the far right side of the track, click on the black triangle next to "Label Track" and select **Move Track to Top**:



- At the top of the program, click on **File > Export > Export Multiple**:



- The following menu will appear. Make sure you have these options selected:



- Click **Export**. Several windows will pop up, click **Ok** for all of them. The program will save the audio as individual files.
- Finally, save the Audacity project file in the designated file location. To do this, click on **File > Save Project > Save Project As...** Use the Audacity project file to conduct QC and make any necessary changes to the labels or audio.

All audio files should be stored in a secure location and backed up regularly.

Quality Control (QC)

The exact process for ensuring data quality is subject to individual research protocol. We recommend researchers integrating efforts with the existing QC procedures and documentation. Below we offer the suggested best practices that researchers may want to adopt if feasible.

Researchers should develop a system for tracking QC activities. This can be done using your preferred program such as REDCap or Excel. It should, at minimum, track the following variables:

1. Participant ID (NACCID and/or PTID)
2. Tester ID
3. Recording date
4. Name of audio file
5. ID of person conducting QC
6. Date of QC activity

7. What type of QC is being done (as outlined below: Supervisor, Peer, Intra, Data Integrity, etc.)
8. Whether the QC passed or failed
9. Why the QC failed, if applicable

We encourage researchers to create a feedback loop for the QC process. This means that the people who conducted testing and processed the audio (this might be the same person or different people) are sent the QC results so they can tell if they made any mistakes. The testers and audio processers can sign off on the QC to confirm that they have reviewed any errors.

Manual QC for PII

Given the importance of maintaining the confidentiality of research participants, it is essential that QC measures are implemented to ensure consistent and accurate removal of PII from the recordings. While we provide a framework for QC below, we encourage researchers to adapt and/or develop a process that works best with their existing infrastructure. Our recommended format includes three levels of QC: supervisor, peer, and self (intra).

- **Supervisor QC**: A supervisor should regularly choose recordings at random to undergo QC. We recommend selecting at least one recording from each tester every month. They should listen to the entire recording to check that all PII has been labeled and silenced. If they find PII, they should flag it and ask a staff member proficient with Audacity to label and silence the PII, then save the audio file and timestamp (.txt file) as updated versions.
- **Peer QC:** A Peer Reviewer should review 5-10% of the completed exams done by each tester on a regular basis. The recordings should be chosen at random. This does not have to be a trained NP tester; it just needs to be someone trained to listen for PII. The Peer Reviewer should listen to the recordings and make sure all PII has been labeled and silenced. If they find PII, they should flag it and ask a staff member proficient with Audacity to label and silence the PII, then save the audio file and timestamp (.txt file) as updated versions.
- **Intra QC**: Each person processing the audio should review their own test recordings, chosen at random, to ensure all PII was accurately identified (we recommend reviewing one recording a quarter). If they find PII, they should label and silence the PII, then save the audio file and timestamp (.txt file) as updated versions.

Each recording that undergoes QC should be logged in the designated QC tracking system. On a regular basis (we recommend quarterly), the tracking system entries should be reviewed to ensure that there are not (1) common pitfalls or (2) problems with the accuracy of any examiner or person processing the audio. Any common pitfalls or differences of

opinion will be discussed with the team and steps taken to resolve them. If the person processing the audio is not consistently accurate, they should be given feedback and additional recordings from that person should be reviewed. The number of additional recordings that will be reviewed will be decided upon by a supervisor based on the given circumstances.

QC for File Labels, Locations, and Data Logs

We strongly recommend that researchers implement a QC process that will ensure data files are labeled correctly and located in the appropriate folders. This is essential because the recording should contain no PII, which means the file name will be the only way to identify the recording.

Audio files must be labeled according to the prescribed naming convention that corresponds with the correct row in the data log. Without a correct file name and match in the data log, it will be extremely difficult to align the voice data with participant phenotypic data. The file name and location should correspond with the data log with no discrepancies.

Recommended Naming Convention: NACCID_DATE_COGNITIVETEST.wav

Step 9: Upload Digital Voice Data To NACC

COMING SOON!

NACC will be providing ADRC's with a digital voice data submission option via the ADRC Portals hosted on NACC's Data Platform. Learn more about the ADRC Portals.

PLEASE NOTE: NACC will not share any raw voice recordings until robust voice de-identification standards have been determined by NACC, NIA, ADRCs, and the CTF Technology Workgroup. NACC will provide additional guidance on IRB and consent documentation, data submission, and sharing as this process is further defined with input from ADRCs, the CTF Technology WG, and NACC.

In preparation for data submission, we encourage sites to:

a. Confirm the PTID and NACCID in your data logs are correct.

   To access your site's PTID and NACCID pairs, please work with your ADRC site data manager to utilize the PTID to NACCID Map tool available via the NACC portal.

b. Ensure voice recording files are properly named and correct.

   Recommended naming convention:

   NACCID_DATE_COGNITIVETEST.wav

c. Ensure you have collected the minimum metadata variables required to accompany voice recordings (see recommendations regarding data logs above - Item #7)

   View Digital Voice Metadata Dictionary

   Download template data log file

Appendix A: Frequently Asked Questions

1.) **How long should we expect audio recordings to be?**
    a. This depends entirely on the content that is being recorded. A set of neuropsychological tests could take an hour, whereas other shorter cognitive batteries may take less time.

2.) **What sort of file sizes should we expect for audio data?**
    a. That depends on the format (e.g., uncompressed vs. compressed), duration of the audio, sampling rate, bit-depth, and number of channels.

       For example, a WAV (uncompressed format) audio file that is one hour long, has 16kHZ sampling rate, 16 bit-depth, and one (mono) channel is roughly 115.2 MB.

3.) **How does a manual transcription process compare to an automated transcription process?**
    a. Manual transcriptions take a significant amount of time to produce. It could take about 7-8 times the length of the speech time to manually transcribe a recording. It takes significant effort to define a robust set of guidelines that transcribers can be oriented to – especially because the guidelines will likely adapt and change as different scenarios are encountered.
    b. Automated transcriptions are much faster to produce and are mostly limited by available computational resources. However, the quality of the transcription depends on the audio and the automated method being utilized.

Please visit the following webpage for an up-to-date list of FAQs provided by NACC and the CTF Technology Workgroup: https://naccdata.org/adrc-resources/udsv4-digital-voice

Appendix B: Using the ZOOM H4N Pro Device

Initial Set-up

1. Turn the power on by moving the power switch on the left panel of the device to "ON".
2. Press the menu button on the right side of the recorder.
3. Scroll to "SYSTEM" and enter it.
4. Enter "DATE/TIME" and set the date and time. The recorder uses military time. Press "OK".
5. Return to the menu and enter "REC".
6. Change "REC FORMAT" to "**WAV48kHz/24bit**". Exit out of the menu completely by repeatedly pressing the menu button.
7. Change the recording level to **50** using the "REC LEVEL" rocker on the right side of the recorder.
8. Set microphones to **120º**.
9. Initial set-up is complete. You may turn off the device.

Loading the SD Card

1. Be sure the power is OFF when inserting or removing the SD card to avoid destroying data.
2. Insert the SD card into the slot on the left panel of the device
   If "Format Card" appears on the display screen after inserting the card, it means that the SD card has not been formatted in the H4n Pro device. To format it, use the dial to select "YES".
3. To check the remaining capacity of the SD card, press "MENU" and select "SD CARD". Select "REMAIN" which will then display the remaining capacity meter, remaining space, and remaining recording time using the current settings.

Recording Instructions

1. The recorder should always have functioning batteries installed, regardless of whether an AC adapter is being used.
2. Plug recorder into the AC adapter in the testing room (or, if testing elsewhere, have an AC adapter with you and try to arrange the testing location so you can plug in the recorder).
3. Turn the power on by moving the power switch on the left panel of the device to "ON"
4. Be sure the "Stereo Mode" indicator is lit.

5. Put the recorder in “Recording Standby Mode” by pressing the “REC” button.
    a. Recording standby means the mic is on but is not yet recording.
    b. The red light on the recorder blinks when in standby mode.
6. Confirm all settings are correct (recording level = 50; recording format = WAV48kHz/24bit; microphones are at 120º). See Initial Set-up section above for instructions.
7. Make sure that the MIC button is pushed on the front of the recorder (NOT the “1” or “2” buttons); see the ZOOM H4N DVR Image below.
8. Start recording by pressing the Play/Pause [►/||] button. The time counter on the screen will advance, the recording symbol [●] will appear next to it, and the red light will stop blinking and remain on.
9. Record the following information: Participant ID, Date of testing, and Examiner ID.
10. Press the Play/Pause [►/||] button again to pause recording until ready to start recording the participant.
11. Lay the recorder down with the head of the recorder pointed directly at where the participant will be seated. In testing rooms, place the recorder on the file cabinet next to the testing table, as close to the participant as possible. It’s absolutely essential to place your paper holder on the opposite side of the table relative to the recorder, because the recorder is very sensitive, and paper shuffling will muddle audio. If you are not in a testing room, try to arrange to place the recorder on a different surface, but still close to the participant, so it does not pick up all the paper shuffling, table jarring, etc.
12. After the participant has been consented and has signed the consent form, you may begin recording the examination. **NOTE:** If the participant reports that they do not want to be audio recorded, turn off the recorder, remove it from the table, and proceed with consenting/testing (unrecorded).
13. Since you are currently in “Standby” mode, press the Play [►/||] button. Again, the time counter on the screen will advance, the recording symbol [●] will appear next to it, and the red light will stop blinking and remain on. MAKE SURE THIS IS ALL HAPPENING BEFORE PROCEEDING WITH TESTING.
14. After each cognitive test is complete, stop the recording by pressing the stop button [■]. Repeat steps 13-14 for each cognitive test if recording each cognitive test as a separate audio file.

    PLEASE NOTE: We recommend centers record each cognitive test as a separate audio file. However, centers still have the option to record the entire UDS4 cognitive battery in a single recording and can split audio recordings by test during pos-processing.

15. YOU MUST TURN OFF THE RECORDER BEFORE UNPLUGGING THE A/C ADAPTER OR ELSE THE RECORDING MAY BE LOST. (This is only true if the batteries in your recorder are dead, but you should always follow this procedure to ensure data is not lost.)

16. Although you are unlikely to need to play the recording back on the DVR device itself, this can be done by pressing the, [►/||] button to play and the, [■] button to stop.
17. To play an older recording back, press “MENU” then select “FILE” using the dial. Select the file to play and press. Select “SELECT” and press. Press the [►/||] button to start playback.

Using USB to Transfer Files

1. Connect device to computer with USB cable.
2. Press the “MENU” button on the right panel of the device.
3. Select “USB” using the dial and press.
4. Select “STORAGE” and press.
5. The device is now connected to the computer and the files can be transferred
6. Save the file in the appropriate file location with the file naming convention for unedited recordings

Dividing or Deleting a File

It is unlikely you will need to use these features; however, in the rare case that it may be necessary (e.g., two participants were accidently recording in the same file), follow these directions:

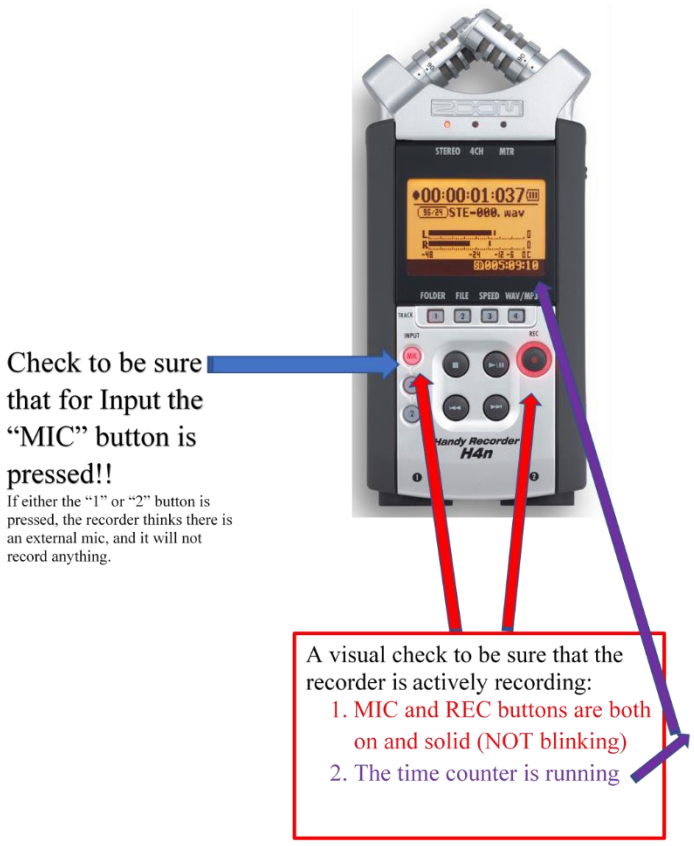
1. Press the “MENU” button on the right panel of the device.
2. Select “FOLDER” using the dial and press.
3. Select a folder using the dial and press.
    a. To divide a file and a desired position, select “DIVIDE” and press. Press to start the playback and press again at the division point. Select “YES” to confirm the divide.
    b. To delete a file, select “DELETE” using the dial and press. Select “YES” to confirm deleting. **Never delete files from the recorders until you are 100% certain they are correctly stored elsewhere**

Battery Type

1. To display the remaining battery life when using batteries, press “MENU”
2. Select “SYSTEM” using the dial and press.
3. Select “BATTERY” using the dial and press.
4. Select the battery type: Alkaline or Ni-MH.

Software Update

1. To download the most recent system software, the device with an SD card must be connected to a computer with access to the internet.
2. Open the ZOOM website (http://www.zoom.co.jp)
3. Connect the H4n Pro to the computer with the USB cable
4. Copy the downloaded software to the root directory of the SD card
5. Disconnect the H4n Pro
6. Turn it on while holding down the [►/||] button. Select “OK” when prompted to upgrade the version.



Appendix C. Digital Voice Recorder Alternatives

This list was compiled by the Framingham Heart Study Brain Aging Program and was last updated in June 2020. It can serve as a resource for researchers that are "shopping around" to find a recording device that best suits them. Please note that some details may become outdated over time – we will do our best to keep this list updated as device specifications and models change.

32GB Digital Voice Recorder, Homder Voice Activated Recorder

- Dynamic noise reduction chip & dual microphones to capture sound clearly, gives you a really clear and natural audio.
- All recordings are named with a timestamp, convenient to find the file you are looking for.
- Built-in 32gb flash memory stores up to 2,000+ hours of maximum recording time
- utilizes DSP digital & AGC noise reduction technology to enhance human speech recordings and filter out background noise, to give a really full clear and warmer vocal recording.
- A single full charge (about 4 hrs) could be continuously used 60+ hrs.
- Multiple high-fidelity speakers ensure a crisp & loud enough playback even without headphones.
- Password function keeps your files far away from leaking.

EVISTR 16GB Digital Voice Recorder Voice Activated Recorder with Playback

- Voice Activated Record
- Reduce blank and whispering snippet
- Voice Recorder USB Rechargeable
- File name with Year, Month, Day, Hour, Seconds
- Dynamic noise cancellation microphone, capture 1536kpbs crystal clear audio
- Voice Recorder MAC Compatible (WIN Compatible)
- Easy to figure out, press REC: starts to record; press STOP, save the recordings safely. Small Voice Recorders with A-B repeat, fast forward, rewind function during playback, a helpful recorder for lectures, meetings, interviews, speeches, class
- Voice Activated Recorder: set the AVR voice activated function, record only when the teacher is speaking, reduce blank and whispering snippets, save space and time. Recording your appointment, meetings, interviews, speeches, lectures easily.
- Easy File Management: recordings with time stamp, easy to find out when you recorded, what it recorded.
- #1 best-seller on amazon
- **Does not have Autosave feature**

- Do not shut down the device until you press STOP to confirm the file saved properly, it will show “Saved!”
- Do not shut down the device, while you are formatting, wait until it shows “format completed”

16GB Digital Voice Activated Recorder - aiworth 1160 Hours Sound Audio Recorder Dictaphone

- E36 voice recorder equipped with dual sensitive microphone and professional recording IC, support up to 1536Kbps PCM recording, provide a super clear recorded voice
- Built-in 800mAh rechargeable battery, support up to 45 hours continuous recording.16Gb flash memory could save 1160 hours recording files at most,
- in addition to this can support up to 32GB TF card(In addition to purchase) expansion and voice activated recording.
- The most user friendly voice recorder designed by aiworth, all operation buttons on the front side, operational logic like smart phone.
- With graphic user guide
- Lifetime software update
- Power-on password protection- 3-digit password,8000 combinations; without the password, no one could turn on the device and overheard your recorded files. After three trial and errors, the device will auto turn off.
- 16 levels to adjust the play speed; play faster, jump to the point you exactly want to playback; play slowly let you hear every single word clearly.

Aomago 8GB Audio Recorder Mini Portable Tape Dictaphone with Playback, USB, MP3

- This recorder upgraded its higher sensitive microphones, meaning that you can enjoy premium quality sound.
- Simple three-click recording, saving and playing, make it super user friendly.
- Set the recorder to voice activated recording, catch the speaking words only.
- A-B REPEAT FUNCTION: This is a great feature to help you study language, review lessons from selected starting point A to ending point B. You don’t have to go back or forward to listen to the words anymore.
- Easy transfer files: Voice recorder mac compatible. It supports recording files in MP3 or WAV format. You can transfer files easily by connecting to a computer via supplied Micro USB cable.
- 8GB MEMORY CAPACITY
- High Quality (128 kbps): 7680 Mins
- Short Play (64 kbps): 16920 Mins
- Long Play (32 kbps): 33120 Mins
- 7 EQ modes
- Different languages

- USB connection, for uploads and downloads
- Battery life expectancy: up to 12 hours continuous recording
- Warning:
- Do not use the right “POWER” button to totally shut down your voice recorder, or it will reset your voice recorder system time to default.
- We suggest you press the PLAY/PAUSE button for two seconds to power off your voice recorder, and next time you just need to press “PLAY/PAUSE” again to wake up your voice recorder.
- When Battery is almost exhausted or too weak, functions may be limited, please recharge!
- Charge time between 3 to 4 hours, turn on the voice recorder before charging.
- Press the REC button while recording to pause or resume recording. The LED will flash when the recording is paused.

Wohlman. Digital Voice Recorder 16GB 1536kbps Touch Screen High Recording Quality Noise Reduction Easy Operation Auto Activation MP3 Voice Recorder

- Clear recording with a resolution of 1536 Kbps and microphones with dynamic noise reduction, higher bit rate, higher recording quality, crystal-clear recordings and MP3 player.
- The built-in 180mAh battery can record 12 hours continuously. With 16 GB of internal storage, you can save up to 145 hours of recordings or 1500 songs.
- Automatic recording is possible with the preset time. Simply record and save with the "REC / Save" button. Simple operations with touch buttons.
- With the automatic voice recognition function, the recorder automatically starts recording when the sound is recognized. Without sound, it will be in standby mode to reduce recording capacity and power consumption. The detecting distance can reach up to 10m.
- With the A-B repeat play function, the recording can play back within a certain period of time. You can also fast forward and rewind during playback, which is useful for reviewing lessons, meeting records, songs, interviews, etc. We offer a one-year guarantee.
- With the USB cable, you can easily transfer the files to the computer as well as delete the files directly. Compatible with Windows and IOS systems.
- The password setting secures your recording data
- Tschisen V93 is embedded with AGC noise reduction design and will give you high quality recordings.
- Speech recognition automatically picks up detected sounds and stops recording when it is quiet.

Sony ICDUX560BLK Digital Voice Recorder 1" Black

- NYtime #1 pick for voice recorders
- Built in stereo microphone and voice operated recording
- Three recording options: wide/stereo, narrow/focus and normal
- Quick charge; up to 1 hour recording time, with 3-minute charge
- Easy to use user interface and recording level indicator
- Micro SD memory card slot, headphone jack & mic input. LCD backlight
- Record in MP3/LPCM with a high-sensitivity S-Microphone
- Up to 4 GB of built-in storage, expandable via MicroSD (SDHC/SDXC) cards
- Focus and wide microphone modes to suit lectures or meetings
- Direct USB built-in for easy connection to PC
- FM radio to listen to or record radio broadcasts
- Normal, focus, and wide-stereo recording provide you with the opportunity to record the audio that you need to capture in any environment, while the slim and lightweight build make it easy to take with you wherever you go and the easy to use up makes file searching simple.
- UX560 received the highest overall ratings from our panel of test listeners (nytimes). It produces clear, understandable audio in the classroom, quiet office, and noisy coffee shop settings. It also offers a better collection of features than the other models we tested, with an easy-to-navigate menu system, a bright backlit screen, 39 hours of recording time (in MP3 format), 27-hour battery life, voice-activated recording to pause and restart after silences, and a pop-out USB 3.0 connector that lets you recharge the recorder and transfer files to a computer easily. Like many of the other recorders we looked at, it comes with an adequate amount of onboard storage (4 GB) but accepts microSD cards, so you can record and store hundreds of hours of recorded audio should you need it. The UX560 is also the slimmest recorder we tested—at 0.43 inch thick it can easily fit in a shirt or pants pocket.

SONY PCM-D10

- Reliable hi-res recordings of up to 192kHz/24-bit
- 3-way adjustable high-resolution 40K frequency response microphones
- 2 XLR-TRS combo jacks with 48V phantom power
- Digital dual-path limiter function prevents distortion
- Bluetooth capability for both remote control and playback via Sony's free REC Remote app
- More expensive than most ($500)
- Commonly used for podcasters, radio, amateur film makers
- Capture flawless Hi-Res sound anytime, anywhere with the pcm-d10 portable recorder. Record professional sound with Hi-Res Audio at up to 192kHz/24-bit. Whether it's your live music set, new podcast episode or breaking news report, the pcm-d10 unlocks a new level of detail and texture. The three-way adjustable

microphones adapt to your situation, while the twin XLR/TRS combo jack lets you plug in your choice of input. High-quality dual ADCs maximize S/N and independent analog volume dials give you precise control of your inputs.

Tascam DR-05 recorders

- The dual internal condenser microphones can handle anything from subtle to loud, with sensitivity to capture every detail
- A revamped layout means operations like recording, adjusting levels, deleting bad takes and adding Markers are quick and easy
- Uses only two AA batteries but can record for an outstanding 17. 5 hours; It can also be powered by a USB mobile battery
- Connect to a PC using USB Audio Interface Mode for voiceover work, live streaming, podcasting and songwriting with studio-quality audio
- Used in a study investigating pitch modulation in human mate choice
    - In this study the researchers used a sampling rate of 96 kHz and 24-bit amplitude quantization. Recordings were stored onto microSDHC media cards as uncompressed WAV files and later transferred to a laptop computer for editing and analysis. This method allowed us to obtain high-quality, directional voice recordings that would otherwise be difficult to obtain in a noisy environment using a stationary microphone.
    - Acoustic editing and analysis were performed in Praat v. 6.0.21 [32]. Fragments of silence, acute noise, non-verbal vocalizations (e.g. laughter) and multi-voicing (e.g. the voice of the dating partner) were first manually removed from audio files. Recordings were then segmented into multiple parts each corresponding to a given participant and a single speed date. We further split each sound file into three equal time segments (beginning, middle and end of the date; mean segment duration 50.6 ± 23 s), resulting in a total of 726 voice clips for acoustic analysis.

Contact Us

Have questions about getting started with digital voice?

Email the NACC team at naccmail@uw.edu

References

1. Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule (Office for Civil Rights) 7-8 (2012).

2. Best Practices. Google Cloud. 2022. Updated 2/10/2022. 2022. https://cloud.google.com/speech-to-text/docs/best-practices

3. Introduction to audio encoding. Google Cloud. 2022. Updated 2/10/2022. 2022.

End of Document.

Created in collaboration with the National Alzheimer's Coordinating Center (NACC), the Clinical Task Force (CTF) Technology Workgroup, and the Framingham Heart Study Brain Aging Program at Boston University. 2024.