

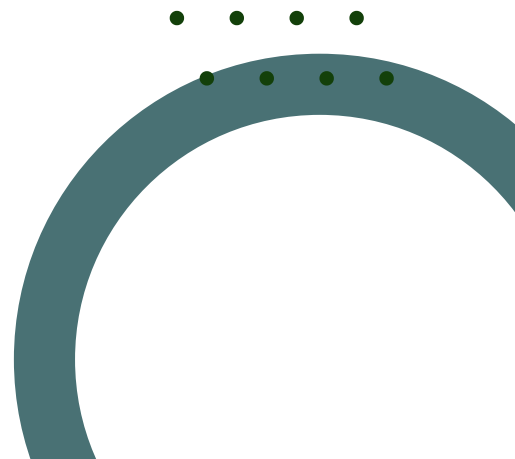
2
0
2
4



HALA
ONLINE

TASK MANUAL

**Carolyn Siegman, Joonhee Kim,
Tomomi Nishikawa, Ces Jocson, &
Theres Grüter**





HALA ONLINE TASK MANUAL

2024

URL: <https://hdl.handle.net/10125/108117>

Carolyn Siegman¹, Joonhee Kim¹, Tomomi Nishikawa²,
Ces Jocson¹, & Theres Grüter¹

¹ *University of Hawai'i at Mānoa*

² *Ochanomizu University, Japan*

*(c) 2024. The HALA Online, including all materials, scripts, and this manual, is
openly licensed via CC BY-NC 4.0*



ACKNOWLEDGMENTS

This version of the task was originally created as a class project in Professor Theres Grüter's seminar class on Heritage Language Acquisition at the University of Hawai'i at Mānoa in the Fall 2023 semester. While this task was created by the authors of this manual, we are very grateful for the feedback we received from our classmates and our friends who unofficially piloted the task as we were still constructing it. We also thank Cleo Jocson, the artist who created all of the illustrations used in this version of the HALA task.

We would like to express our sincere gratitude to Professor William O'Grady for sharing the materials from the original version of the HALA task which were created by him, Amy J. Schafer, Jawee Perla, On-Soon Lee, and Julia Wieting (O'Grady et al., 2009) with us, and for the permission to develop and share this tool further. Many thanks also to Professors Kitaek Kim (Seoul National University) and Hyunwoo Kim (Yonsei University) for sharing the materials from their adaptation of the HALA task with us.

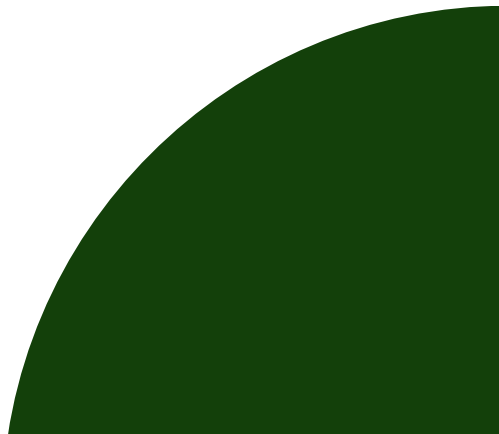
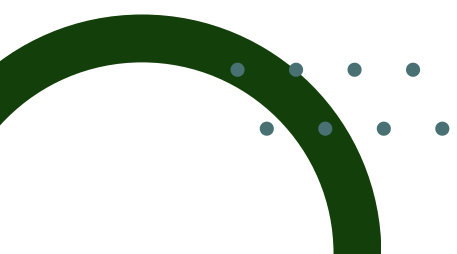




TABLE OF CONTENTS

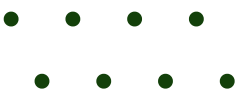
1. Introduction	1
1.1 Previous versions of the HALA Task	2
1.2 Present Version: The HALA Online	4
2. Instructions for Researchers	6
2.1 Creating Your Version of the Task	7
2.2 In Person Data Collection	13
2.3 Remote Data Collection	19
3. Data Analysis and Scoring Guidelines	22
3.1 Scoring for Accuracy	23
3.2 Converting Audio Files	25
3.3 Response Time Analysis	29
4. Concluding Remarks	31



INTRODUCTION

- 1.1 PREVIOUS VERSIONS OF THE HALA TASK
- 1.2 PRESENT VERSION: THE HALA ONLINE

This manual is for a new version of the body-part naming task from the Hawai'i Assessment of Language Access (HALA) project. Throughout the rest of this manual, this task will be referred to simply as the HALA task. Originally created by O'Grady et al. (2009) and further modified by Kang (2011) and Kim & Kim (2022), this task can be used to test relative language strength for bilingual speakers. This section will discuss the changes and motivations behind the creation of the current version of the task. Section 2 offers instructions for researchers for both remote and in-person data collection. Finally, Section 3 provides some suggestions for how the results of the task may be scored for data analysis.



1.1 PREVIOUS VERSIONS OF THE HALA TASK

The first iteration of this task by O’Grady et al. (2009) was designed as a production experiment in which participants name aloud a designated body part after being prompted with The recording begins automatically as soon as the body-part picture appears on the screen within PCIBex. a photo of it. Each trial began with a bell sound and then finished when the participant provided the term or chose to skip the trial if they did not know the word. A recording device would collect all the trial data in one single sound file and the bells would mark the beginning of each one so researchers could then measure response time.

The task used 43 stimuli images that were carefully chosen to represent body parts that commonly have designated names cross-linguistically as this task was designed to be used with different language combinations. Specific parts were indicated using a red circle to elicit words such as “lips” or “teeth” for an image of a mouth. The 43 items were then sorted into three categories based on their relative frequency of use as determined by intuitive ratings, naming times from pilot data, and HAL (Hyperspace Analogue to Language) log frequencies of words from the English Lexicon Project (Balota et al. 2007). Higher frequency words were presented first, then the lower frequency words followed. Words within categories were pseudo-randomized.

Body parts were chosen to be the focus of the task due to the higher likelihood of finding counterparts across languages, their likelihood of being acquired earlier by speakers, and their relative resistance to replacement by borrowed words from other languages. This project also included some stimuli to elicit other lexical items, such as words that name entities in the natural environment. Since landscape and flora vary geographically, the discussion of the HALA task here is limited to just the body part naming task as those terms are more broadly applicable across different populations, and has been the focus of other versions of the HALA task that have been created since.

Bilingual participants would complete this task twice, once in each language, and the performance in one language would be compared against the other to determine in which language the speaker had greater accuracy (i.e., could name the most body part terms correctly), and greater speed in naming (i.e., lower response time in naming the body part). Therefore, this task is designed to be an assessment of relative language strength within a speaker through these measures of lexical retrieval.

Kang (2011) used an adapted version of the HALA task in his dissertation. The overall design was the same as in O’Grady et al. (2009), though he made a few key changes. While O’Grady et al. had three categories of low, medium, and high frequency, Kang used a subset of 31 stimuli from the original task in order to reduce cognitive load on participants as his study involved children (ages 5-12 years old). He eliminated one high frequency term (hair), one medium frequency word (waist) and several low frequency words (arch, bicep, calf, cheekbone, eyelid, forearm, knuckle, pupil, shin, toenails). The remaining medium and low frequency words were collapsed into one single “low frequency” category of 14 items total, based on relative frequency measures again from the English Lexicon Project (Balota et al. 2007). Instead of using the self-paced task design originally reported by O’Grady et al., each trial was pre-timed to last either 4,000 milliseconds for higher frequency words or 4,500 milliseconds for lower frequency words, and the trials moved on to the next whether the participant provided a name for the image or not. Kim & Kim (2022) adopted Kang’s version of the task for their use with children 11-14 years old and additionally used a different set of color photos for their stimuli. Both the original task by O’Grady et al. and the adapted versions by Kang (2011) and Kim & Kim (2022) are administered using Shockwave Flash animations on a computer screen.

There has also been a version of the HALA task created specifically for children (Dubiel & Guilfoyle 2017). While this particular version was not used as a reference for the creation of the HALA Online version, it is worth mentioning for researchers who work with children. The Child HALA version uses images of cartoon characters and selective coloring to highlight the item or body part to be named. It also uses small animations between trials to keep children engaged in the task and orders some trials to better avoid confusion regarding which part they are supposed to name. This task uses a subset of 27 items from the original 43-item list created by O’Grady et al.

1.2 PRESENT VERSION THE HALA ONLINE

The version of the HALA task presented here, referred to as the HALA Online, makes some notable changes in the materials and implementation of the task. First, whereas the previous two iterations used Shockwave Flash animations to administer the task in-person on a computer, the present version was built using the online PCIBex platform (Zehr & Schwarz 2018). This format has several advantages. Not only is it a more updated format technologically, but it also allows for remote testing. In the wake of the COVID-19 pandemic, there has been a large shift to methods that allow data collection when the researcher and participant are not in the same location. Even after the most intense periods of lockdown from the pandemic, researchers still continue to use the new remote data collection methods for a number of reasons. The PCIBex platform allows for research to be collected both in-person and online, allowing the HALA Online to be available in a wider variety of contexts. This platform also saves each trial as a separate sound file, which can lead to more streamlined analysis as it eliminates the need for bells to mark the transition to new trials and the need to split one long recording into many. This platform is also free to use, so it is accessible for researchers who may not have access to software that requires a subscription or other payment to use.

This task also returns to the self-paced format first employed by O’Grady et al. Some participants may need more time to recall a word that they know, especially for the low frequency words. The priority for this design was to give speakers every opportunity to name the image even if it requires time to recall. We believe that these trials are still informative regarding the participant’s language use and therefore should be available for analysis. Researchers using these materials are free to impose their own cut-off point in their own data collection if they so choose, and may score trials that take longer than their cut-off point as a “miss.” Different researchers may have different goals with this task, and so this design preserves the option of leaving the trial durations open or imposing a cut-off up to the individual researcher. It should be noted that the circumstances of a self-paced trial with post-hoc cut-offs will differ from a task where the trial durations are pre-determined, and the results between the two may not be directly comparable as a consequence. Nevertheless, we still feel it is best to leave the option for implementing a time cut-off up to the researcher’s discretion.

This task also uses new black and white illustrations as stimuli instead of photos. This choice was meant to clarify the subject matter in each photo and to also use images that are more neutral with regard to race and ethnicity. As this task is meant for broad use in different language communities, this choice eschews the presentation of images of a specific racial or ethnic group as this may be a stumbling block for task administration, and it refrains from reinforcing ideas of who should or shouldn't be represented in media, even for the context of a linguistic task. The HALA Online otherwise draws on the same high-low frequency category distinctions of the 31 items used by Kang (2011) and Kim & Kim (2022). This task provides two primary measures of lexical retrieval: accuracy and response time (more information on scoring suggestions are available in Section 5). Bilingual participants are to complete the task twice, once in each of their languages (or perhaps more times if they are multilingual) with the goal of using these measures to assess their lexical retrieval ability in each language. This ability is then used as an indicator of relative language strength within that individual (see O'Grady et al. 2009 for more discussion of this topic).

The HALA Online is meant to be generalized enough that it can be used with any language combination that a speaker may have, though there may still be some instances where some adjustment is needed. For example, in Russian the words for "arm" and "hand" are the same, so a Russian speaker would provide the same word twice in this version of the task. Individual researchers may choose how best to deal with these cases in their own studies.

INSTRUCTIONS FOR RESEARCHERS

- 2.1 CREATING YOUR OWN VERSION OF THE TASK
- 2.2 IN PERSON DATA COLLECTION
- 2.3 REMOTE DATA COLLECTION

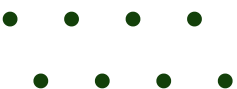
Before creating your own version of the task, you can try out the experiment by visiting the following demo link for the experiment. Please type in any numbers or letters for your participant code to proceed. (Versions A and B are identical except for the order in which the items appear.)

Version A:

<https://farm.pcibex.net/p/gkhfOs/>

Version B:

<https://farm.pcibex.net/p/tmuvUF/>



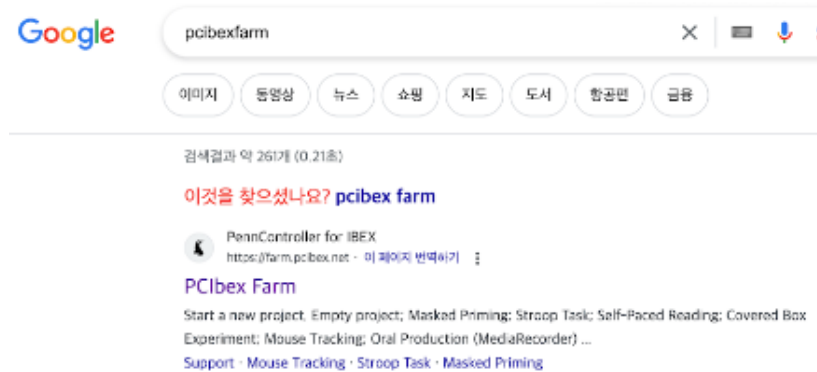
2.1 CREATING YOUR VERSION OF THE TASK

If you think that the experiment fits your needs, you can create your own version of the task on the PCIBex Farm platform. The following section illustrates the steps you can take to create your version of the task with accompanying pictures. Note that all the instructions for the experiment are written in English. For an extensive adaptation of the material beyond the scope dealt with in the present section of the manual (e.g., addition of experimental items, translation of English instructions to the language of interest), refer to the basic tutorial provided by the PCIBex Farm (<https://doc.pcibex.net/basic-tutorial/>) or consult a JavaScript expert.

Summarized Instructions:

1. Visit the “PCIBex Farm” website through searching it in your web browser.
2. Sign up and Login
3. Click “Oral Production (MediaRecorder) to start a new project.
4. Delete all the existing “Resources”.
5. Delete all the existing “Scripts”.
6. Download the folder “HALA Language A” or “HALA Language B” from Scholarspace. You will be creating two separate experiments for Language A and B.
7. Drag “data_includes” folder to “Scripts” section.
8. Drag “chunk_includes” folder to “Resources” section.
9. Publish the experiment for data collection.
10. Before sharing the experiment to the participants, make changes to the “instruction.html” to fit the language you are testing. Click on the “instruction.html” in the “Resources” section.
11. Change [Language A or B] in brackets to actual language you are testing your participants on (e.g., Korean, Japanese). They are currently on the 11th and 22nd line. You can navigate through the html file by scrolling down.
12. Now, you can share the “Data-collection link” to your participants, and start collecting data!

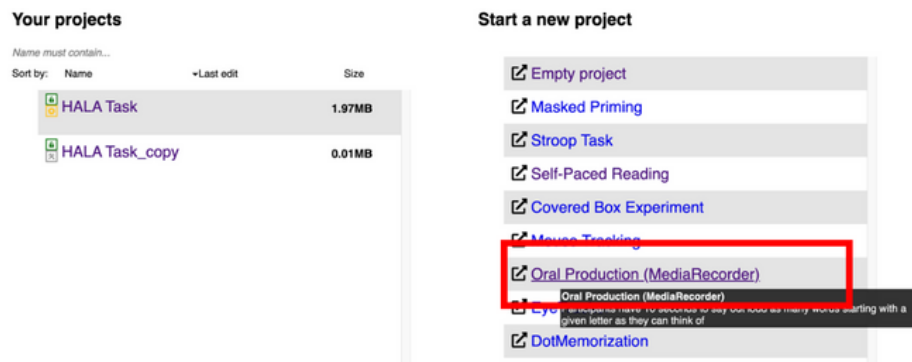
2.1.1 Visit the “PCibex Farm” website by going to the following link: <https://farm.pcibex.net/> or through searching it in your web browser.



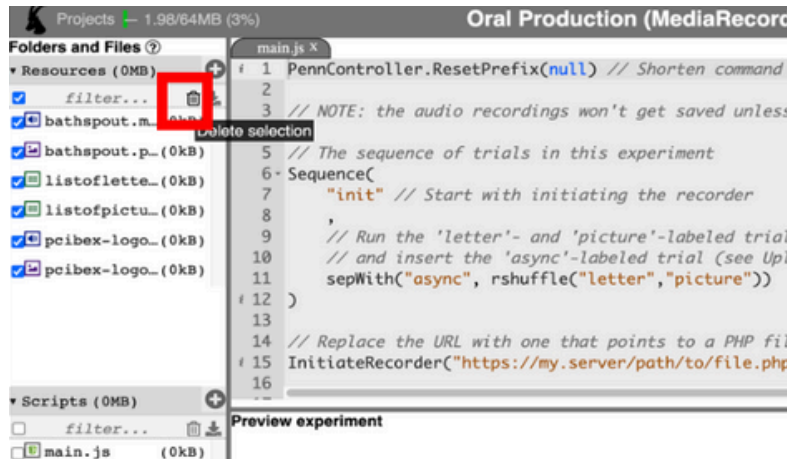
2.1.2 Sign up and login.

A screenshot of the 'SIGN UP' form on the PCibex Farm website. The form has four input fields: 'Username', 'Password', 'Password confirmation', and 'Email address' with a 'why' link. Below the fields is a 'SIGN UP' button. At the bottom, there is a link that says 'or Log in with an existing account'.

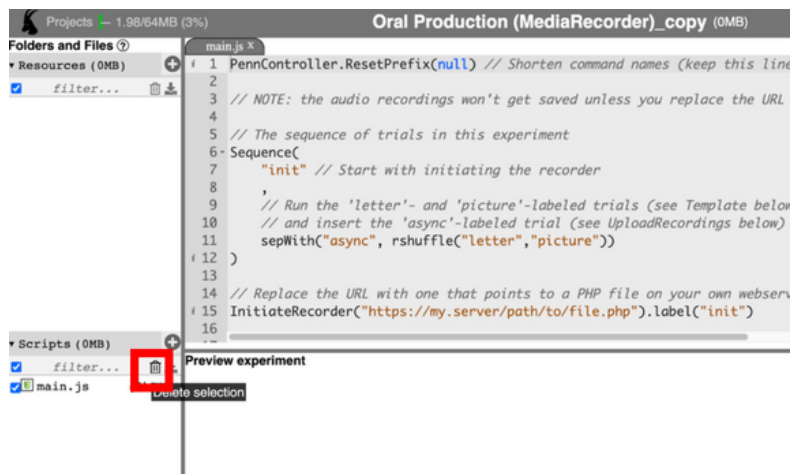
2.1.3 Click “Oral Production (MediaRecorder)” to start a new project.



- 2.1.4** Delete all the existing “Resources”. The deletion may take some time, so please be patient and try refreshing the website until all the resources are gone.

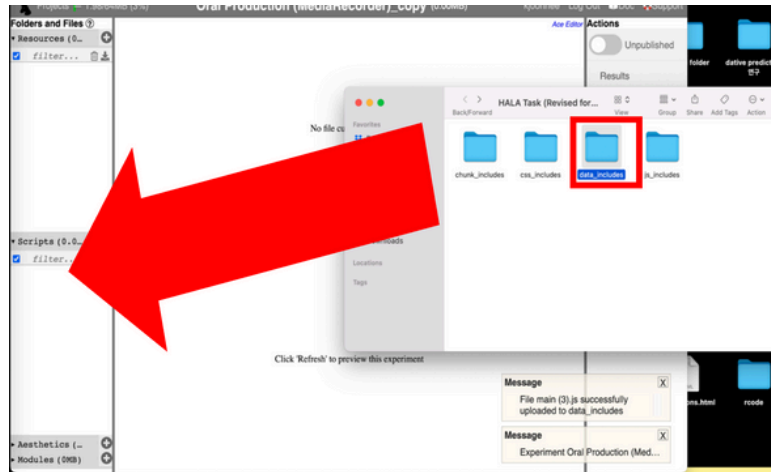


- 2.1.5** Delete all the existing “Scripts”. The deletion may take some time, so please be patient! The deletion may take some time, so please be patient and try refreshing the website until all the scripts are gone.

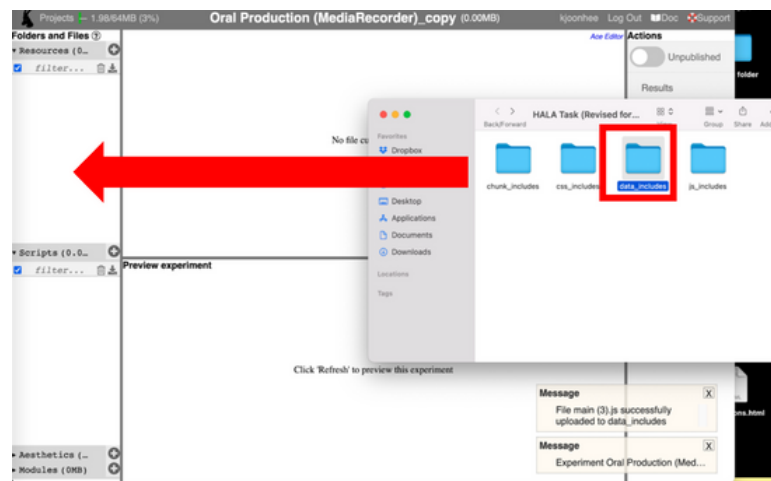


- 2.1.6** Download the folder “HALA Language A” or “HALA Language B” from Scholarspace. You will be creating two separate experiments for Language A and B.

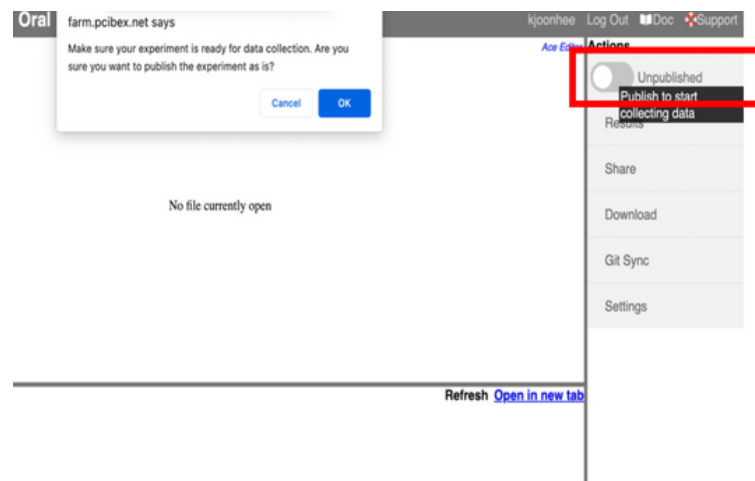
2.1.7 Drag “data_includes” folder to “Scripts” section.



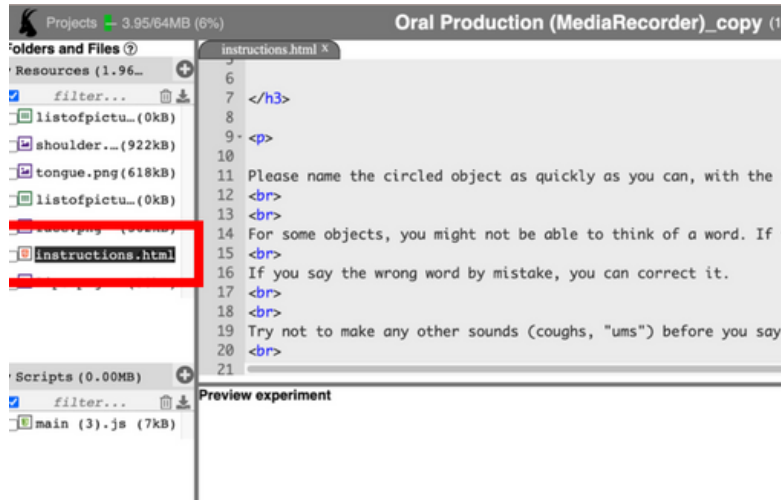
2.1.8 Drag “chunk_includes” folder to the “Resources” section.



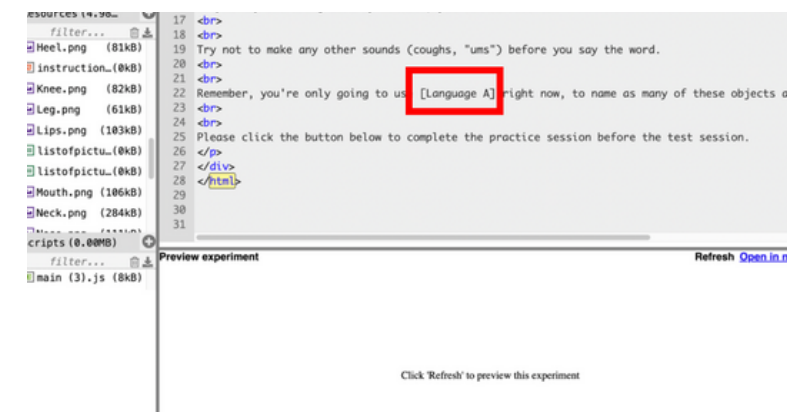
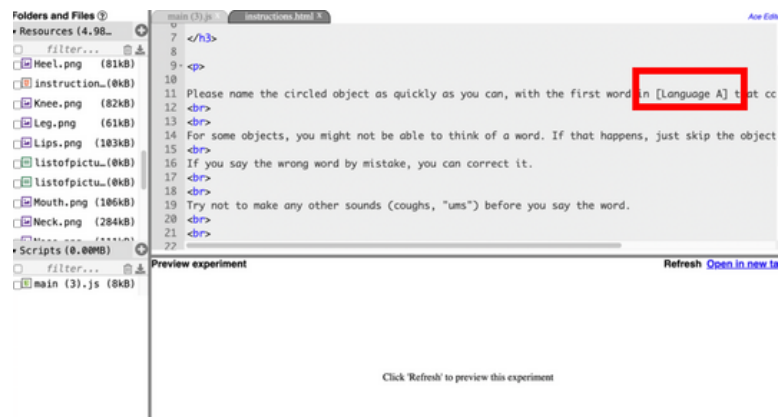
2.1.9 Publish the experiment for data collection.



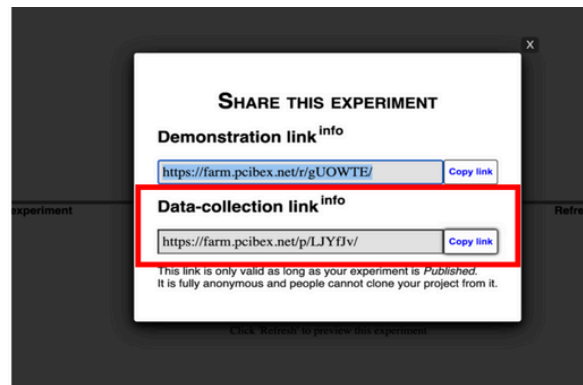
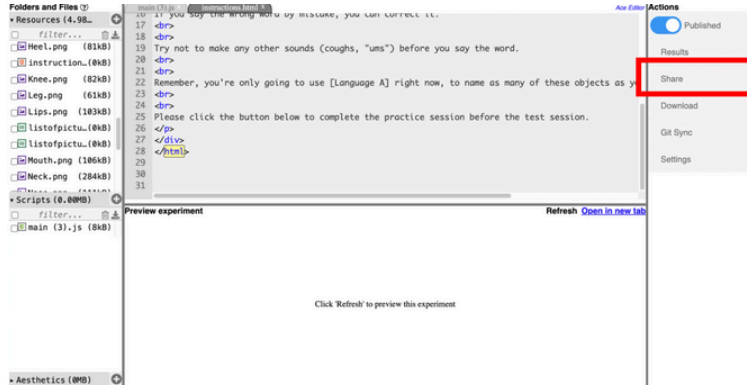
- 2.1.10** Before sharing the experiment to the participants, make changes to the “instruction.html” to fit the language you are testing. Click on the “instruction.html” in the “Resources” section.



- 2.1.11** Change [Language A or B] in brackets to actual language you are testing your participants on (e.g., Korean, Japanese). They are currently on the 11th and 22nd line. You can navigate through the html file by scrolling down.



2.1.11 Now, you can share the “Data-collection link” to your participants, and start collecting data!



2.2 IN PERSON DATA COLLECTION

The procedure for in-person data collection recommended here strongly relies on the procedure as described in O’Grady et al. (2009) for the original version of the HALA task.

2.2.1 How to Administer the HALA Online in an In-person Context

The HALA Online is conducted individually using a personal computer. To ensure accurate response time measurement and analysis, it is recommended to carry out the test sessions in a quiet environment with minimal ambient noise. Researchers need to be mindful of potential sources of ongoing noise, such as computer fans, fluorescent lights, refrigerators, or rain.

2.2.2 Technical Requirements

The HALA Online requires connection to the internet as it utilizes the PCibex Farm website. Please refer to the following website for general problems with utilizing the PCibex Farm (Documents in the PCibex website: <https://doc.pcibex.net/>; FAQ in the PCibex website: <https://farm.pcibex.net/help>). In order to start the experiment, click on the link to the data-collection link that you have created in a quiet internet-connected environment.

2.2.3 Equipment Set Up

The participant should sit directly facing the computer screen, ensuring that the computer microphone is positioned optimally for recording.

2.2.4 Procedure

The following section provides a walk-through of the experiment procedure with accompanying pictures that appear on the screen.

STEP 1 Type in the participant code and continue with the experiment.



progress

Welcome to my experiment!

Before beginning the task, please answer the following question.

What is your participant code?

[click here to continue](#)

STEP 2 Let the participants read the instructions provided.
Ask the participants if they have any questions and clarify any confusions they may have.



progress

Instructions

Please name the circled object as quickly as you can, with the first word in [Language A] that comes to mind.

For some objects, you might not be able to think of a word. If that happens, just skip the object and proceed to the next object. If you say the wrong word by mistake, you can correct it.

Try not to make any other sounds (coughs, "ums") before you say the word.

Remember, you're only going to use [Language A] right now, to name as many of these objects as you can.

Please click the button below to complete the practice session before the test session.

[click here to continue](#)

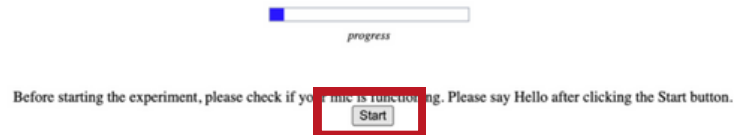
Please name the circled object as quickly as you can, with the first word in [Language A] that comes to mind.

For some objects, you might not be able to think of a word. If that happens, just skip the object and proceed to the next object. If you say the wrong word by mistake, you can correct it.

Try not to make any other sounds (coughs, "ums") before you say the word.

Remember, you're only going to use [Language A] right now, to name as many of these objects as you can.

Please click the button below to complete the practice session before the test session.

STEP 3 Check the recording quality through mic test.

By clicking the “Start” button, you can proceed to conduct a mic test and check the quality of recordings. Let the participant say Hello after clicking the “Start” button.

After 10 seconds, the following screen will pop up:



You will be able to check the recordings through the “Play recorder” button. Check for the signal to noise ratio, and if you think the signal sound is too low or the noise level too high, relocate the participant closer to the computer microphone or eliminate potential sources of noise (e.g., computer fans, fluorescent lights, refrigerators, or rain), reload the experiment link, and redo steps 1 to 3. Otherwise, click “Next” to proceed to practice trials.

STEP 4 Practice Trials

Navigate the participants through practice trials. Remind them not to press the “Next” button too fast before fully naming the circled object.



After the five practice trials are over, ask the participant if they have any questions and clarify any confusions they may have. Remind them again not to press the “Next” button too fast before fully naming the circled body parts in the test trials.



**The practice session is over.
Let's move onto the main task.**

[click here to continue](#)

STEP 5 Test Trials

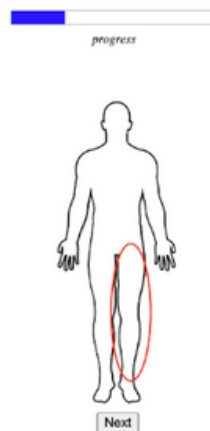
Have the participants continue through the test trials. The test trials consist of 31 items.

We recommend that the experimenter actively observe each test and record observations on tally sheets (see Appendix 1).

It is especially important to note the occurrence of 'filler sounds' (e.g., uhh, hmmm, coughs) as their presence may impact later response time analysis.

Experimenters can also use this opportunity to monitor the accuracy of participants' responses and make note of any peculiarities, such as the use of borrowed words, archaic language, neologisms, and so on.

To prevent any distraction or self-consciousness for the participant, experimenters should position themselves discreetly behind the participant while taking notes.



STEP 6 Download Recordings

After the participants are done with the test trials, download the recording files and save it on your computer. Click the sentence “Please click here to download a copy of your recordings in case you need to send them manually.” (You can ignore the error message on the second line.) Upon clicking the link, a zip file of the recordings will be downloaded to your computer.



2.3 REMOTE DATA COLLECTION

For remote data collection, you can choose to conduct an experiment over video conferencing software (e.g., Zoom, Skype, Teams) or through complete remote data collection by sending the data collection link to the participants. For remote data collection over Zoom, you can send the [Data-collection link] to your participants via Zoom chat and guide the participants through screen sharing in the same way as in-person data collection. For complete remote data collection, we recommend that you provide your participants with a version of the instructions below, adapted for your specific study. Replace the brackets with the information that fits your experimental settings (i.e., language, data-collection link, participant code, and e-mail).

2.3.1 Instructions for Participants

In this task, you will name the circled object in the picture as quickly as you can in [language].

Please move to a noise-free environment for the test session, since a high-quality sound recording is necessary for this task.

Click on the link below and proceed to the experiment on your computer!

[paste in your Data-collection link]

STEP 1 After you click the link, type in your participant code. Your participant code is [participant code].



Welcome to my experiment!

Before beginning the task, please answer the following question.

What is your participant code?

[click here to continue](#)

STEP 2 Carefully read the instructions, and press the “click here to continue” button.



Instructions

Please name the circled object as quickly as you can, with the first word in [Language A] that comes to mind.

For some objects, you might not be able to think of a word. If that happens, just skip the object and proceed to the next object. If you say the wrong word by mistake, you can correct it.

Try not to make any other sounds (coughs, "ums") before you say the word.

Remember, you're only going to use [Language A] right now, to name as many of these objects as you can.

Please click the button below to complete the practice session before the test session.

[click here to continue](#)

STEP 3 Follow the instructions and proceed to the practice and test trials.

STEP 4 After you are done with all the trials, you will see the following screen. Click the sentence “Please click here to download a copy of your recordings in case you need to send them manually.”. (You can ignore the error message on the second line.)



Please wait while the archive of your recordings is being uploaded to the server...

There was an error uploading the recordings: Could not GET http://localhost/Recordings/Guide.php;

[Please click here to download a copy of your recordings in case you need to send them manually.](#)

STEP 5 The program will automatically download a zip file of your recordings. Click on the zip file.



STEP 6 1. Please save the zip file of your recordings and send it to us at [email].



DATA ANALYSIS & SCORING GUIDELINES

- 3.1 SCORING FOR ACCURACY
- 3.2 CONVERTING AUDIO FILES
- 3.3 RESPONSE TIME ANALYSIS

The present section provides guidelines for analyzing and scoring your data from the HALA Online task. There is no one-size-fits-all approach to scoring, so each researcher should make the most reasonable decision for their work. We will provide some tips for making these decisions.

If you have audio-recorded participants' responses, you will deal with two types of data: the words produced by the participants, for which you will determine accuracy, and response time data.


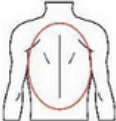
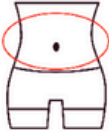
In the HALA Online using PCibex Farm (<https://farm.pcibex.net/>), a sound file is generated for each trial. Accuracy scores can be obtained by examining the responses in the audio files, and response time can be measured to determine how long it takes for participants to orally produce the word corresponding to each body part picture. This seemingly simple scoring process requires researchers to make their own decisions in obtaining data for subsequent analyses, as will be illustrated below.



3.1 SCORING FOR ACCURACY

The HALA task has a list of body-part vocabulary. The most straightforward method for scoring the participants' responses is to check if they match the words in the list.

However, participants sometimes produce words not on the list. Some may be considered acceptable answers, while others may be less so. Drawing a clear line is not always easy. Please see following table for examples.

Expected answer	Participant's response
arm 	elbow
back 	body
stomach 	belly button, tummy

When making decisions, considering the accompanying pictures is recommended. For example, the circled part of the “arm” might be less acceptable as an “elbow” compared to the circled part of the “back” as a “body.” For the “stomach” picture, it is true that there is a belly button in the center of the picture of “stomach.” Additionally, for young children, “tummy” may be a more familiar word than “stomach.” We recommend that you consider the population under investigation in your study to make decisions on what you consider an acceptable response in your specific study.

Some participants may give more detailed descriptions of the pictures, such as “left arm,” which accurately corresponds to the image. Sometimes, they may self-correct their responses. For these responses, you can decide whether to accept them as correct answers. Accepting these answers may have an impact on how we measure response times, a topic we will discuss in the next section.

Avoiding such responses may be possible by providing specific instructions to participants, such as informing participants that they do not need to specify right or left in their responses. However, overly detailed instructions might add unnecessary pressure to participants.

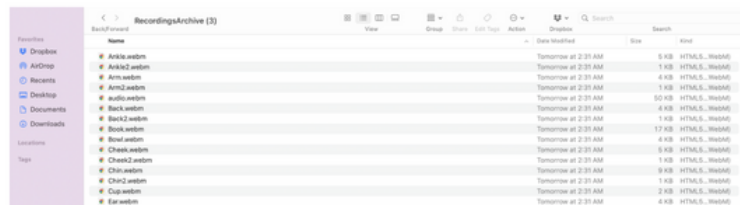
Once criteria for correct answers are established, consistency in scoring is crucial. Make sure that if a specific response is accepted as correct for one participant, the same standard is applied for others.

3.2 CONVERTING AUDIO FILES

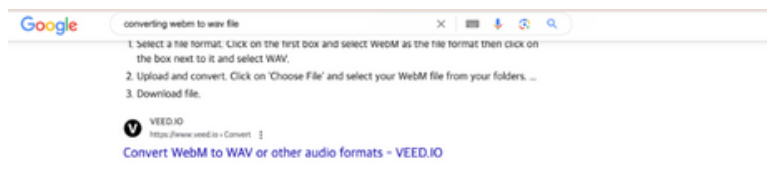
For measuring response time, we recommend using audio editing software. Many are available for free downloading. One caveat to be noted is that the PC Ibox software saves the audio files as .webm format, and analysis of this audio format is not supported through audio editing software. Therefore, before starting response time analysis, you need to convert the audio file format to .wav through the use of converter software. This can be done using WEBM to WAV converters on the internet. Follow the steps below for converting the audio files.

3.2.1. The Use of Free Conversion Websites

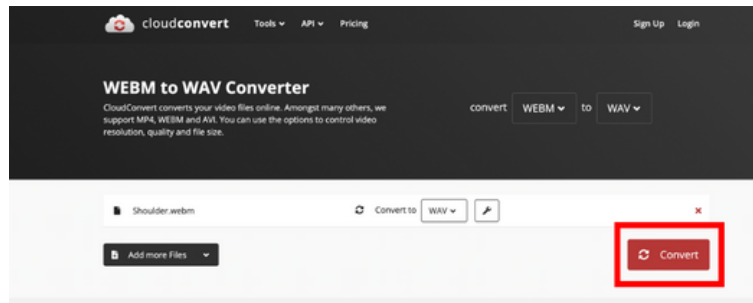
Step 1 Look into the files saved in the RecordingsArchive for each participant. Each audiofile with the name of the picture (e.g., Ankle.webm or Cheek.webm) is the audiofile saved for each trial. See example below:



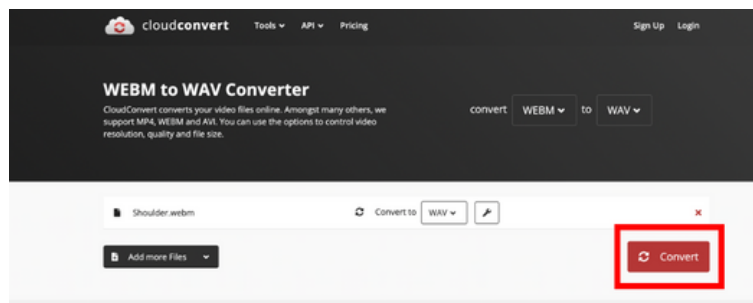
Step 2 Look up “converting webm to wav file” on Google, and access a converter website.



Step 3 “Cloudconvert” is one example of such website, and upload the file you wish to convert to .wav file and press “convert”.



Step 4 “Cloudconvert” is one example of such website, and upload the file you wish to convert to .wav file and press “convert”.

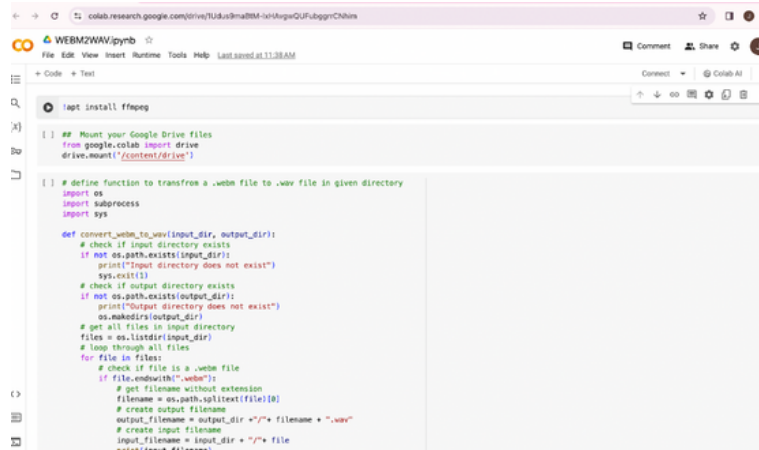


3.2.2 The Use of Google Colab

We provide here the alternative method of audio conversion through providing the python codes in the format of Google Colab. The link for the codes are provided below:

<https://colab.research.google.com/drive/1Udus9maBtM-lxHAvGwQUFubggrrCNhim?usp=sharing>

Step 1 Access the codes through a Google account. Some institutional Google accounts may not grant access to Google Colab. If this happens, try out a personal account. The image below is the interface for the Google Colab when you click the link.



```

!apt install ffmpeg

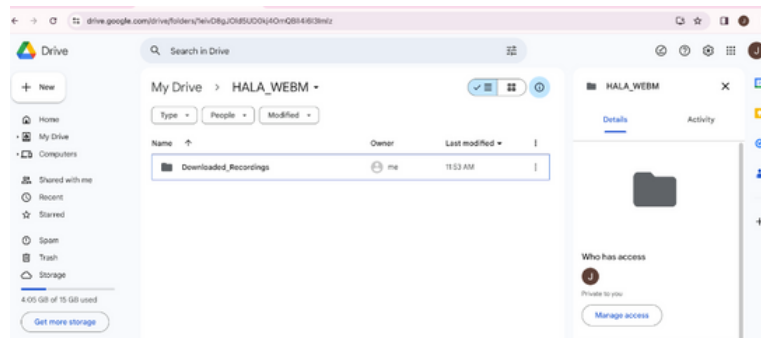
[] # Mount your Google Drive files
from google.colab import drive
drive.mount('/content/drive')

[] # define function to transform a .webm file to .wav file in given directory
import os
import subprocess
import sys


def convert_webm_to_wav(input_dir, output_dir):
    # check if input directory exists
    if not os.path.exists(input_dir):
        print("Input directory does not exist")
        sys.exit(1)
    # check if output directory exists
    if not os.path.exists(output_dir):
        print("Output directory does not exist")
        os.makedirs(output_dir)
    # get all files in input directory
    files = os.listdir(input_dir)
    # loop through all files
    for file in files:
        # check if file is a .webm file
        if file.endswith(".webm"):
            # get filename without extension
            filename = os.path.splitext(file)[0]
            # create output filename
            output_filename = output_dir + "/" + filename + ".wav"
            # create input filename
            input_filename = input_dir + "/" + file
            print(input_filename)

```

Step 2 Create a folder Named HALA_WEBM in your Google Drive, and upload a downloaded recording folder (with all 31 recordings from a participant) as the subfolder. Your Google Drive would look like the following image:



Step 3 Change the name of the input and output directory to the name of the recordings folder you have uploaded. You are changing the folder name after “HALA_WEBM/” to the actual folder name that you have uploaded. They are located under “#call function with input and output directory”, which are the last three lines of the codes.



```

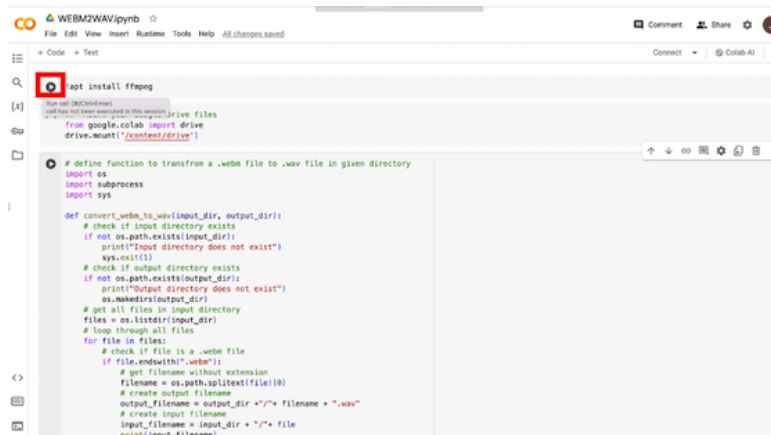
print(input_filename)
# call ffmpeg to convert .webm to .wav
subprocess.call(["ffmpeg", "-i", input_filename, output_filename])
print("Converted " + file + " to " + output_filename)

# call function with input and output directory
input_dir = "/content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings"
output_dir = "/content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output"
convert_webm_to_wav(input_dir, output_dir)

[] Start coding or generate with AI.

```

Step 4 After creating the HALA_WEBM folder, uploading your recording file, and changing the name of the input and output directory, run the codes in the order in which it appears. You can run the codes by clicking on the play button.



```

!pip install ffmpeg

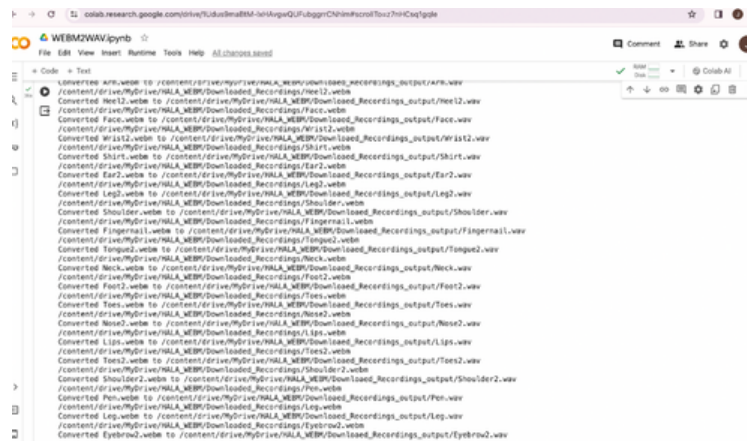
from google.colab import drive
drive.mount('/content/drive')

# define function to transform a .webm file to .wav file in given directory
import os
import subprocess
import sys

def convert_webm_to_wav(input_dir, output_dir):
    # check if input directory exists
    if not os.path.exists(input_dir):
        print("Input directory does not exist")
        sys.exit(1)
    # check if output directory exists
    if not os.path.exists(output_dir):
        print("Output directory does not exist")
        os.makedirs(output_dir)
    # get all files in input directory
    files = os.listdir(input_dir)
    # loop through all files
    for file in files:
        # check if file is a .webm file
        if file.endswith(".webm"):
            # get filename without extension
            filename = os.path.splitext(file)[0]
            # create output filename
            output_filename = output_dir + "/" + filename + ".wav"
            # create input filename
            input_filename = input_dir + "/" + file
            print(input_filename)

```

Step 5 After running all the codes in the Google Colab (pressing three play buttons in total) the screen will be as shown in the following picture:

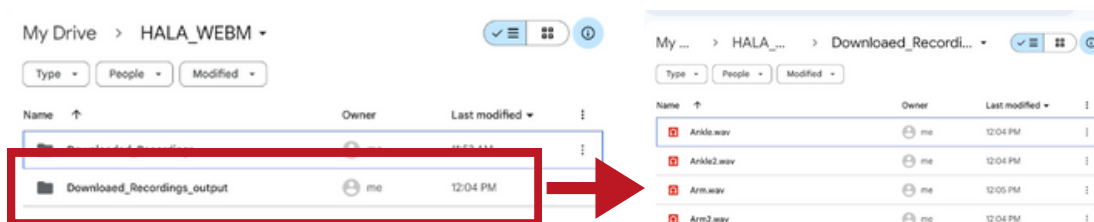


```

Converted Arm.wav to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Arm.wav
Converted Neck2.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Neck2.wav
Converted Neck1.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Neck1.wav
Converted Face.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Face.wav
Converted Wrist2.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Wrist2.wav
Converted Wrist1.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Wrist1.wav
Converted Shirt.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Shirt.wav
Converted Ear2.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Ear2.wav
Converted Ear1.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Ear1.wav
Converted Leg2.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Leg2.wav
Converted Leg1.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Leg1.wav
Converted Shoulder.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Shoulder.wav
Converted Fingerring1.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Fingerring1.wav
Converted Fingerring2.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Fingerring2.wav
Converted Tongue2.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Tongue2.wav
Converted Neck.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Neck.wav
Converted Foot2.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Foot2.wav
Converted Foot1.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Foot1.wav
Converted Ties.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Ties.wav
Converted Nose2.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Nose2.wav
Converted Lips.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Lips.wav
Converted Ties2.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Ties2.wav
Converted Ties3.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Ties3.wav
Converted Ties4.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Ties4.wav
Converted Shoulder2.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Shoulder2.wav
Converted Pen.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Pen.wav
Converted Ring.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Ring.wav
Converted Leg.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Leg.wav
Converted Eyebrow2.webm to /content/drive/MyDrive/HALA_WEBM/Downloaded_Recordings_output/Eyebrow2.wav

```

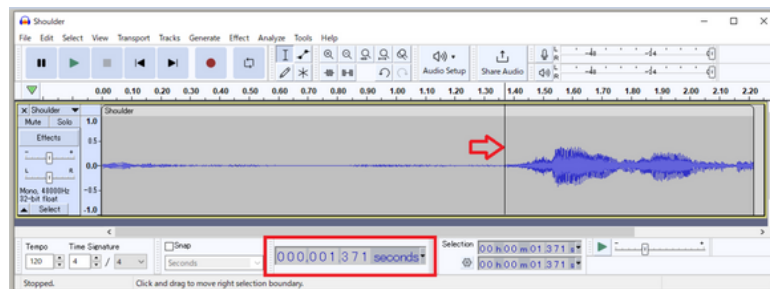
Step 6 Refresh your Google Drive and you will be able to see the output directory created for you with the converted .wav files inside.



3.3 RESPONSE TIME ANALYSIS

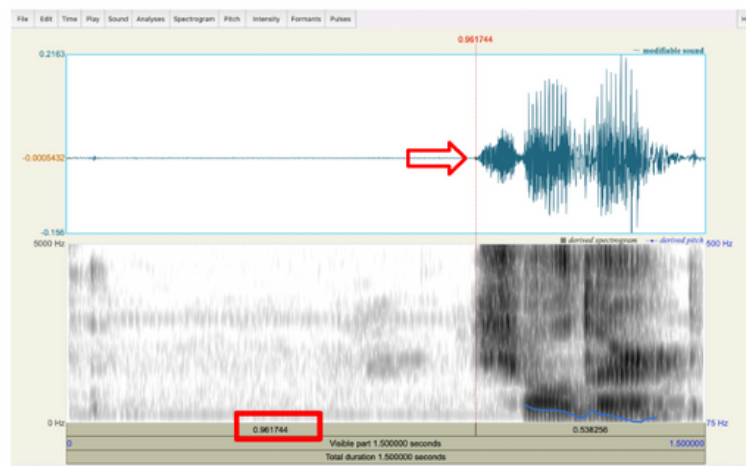
We recommend that response times be calculated and analyzed only for items with a correct naming response (See also O’Grady et al., 2009, p.107). This is standard in psycholinguistic research as it is more difficult to know what processes underlie an incorrect response.

Please see image below for a sample audio file for the target word “shoulder” opened in Audacity (<https://www.audacityteam.org/>).



The recording begins automatically as soon as the body-part picture appears on the screen within PCibex. In this sound file, the recording starts from the left end of the darker gray rectangular window featuring a blue waveform. You can navigate a cursor to the beginning of the blue waveform, starting at 1.371 seconds (or 1371 milliseconds), denoted by the black vertical line. This is the response time for the word “shoulder” in this sound file. (A red arrow and a red box are added for illustrative purposes.)

An alternative software you may use is Praat (<https://praat.en.softonic.com/mac>). Please see image below for a sample audio file for the target word “shoulder” opened in Praat.



In the sound file above, the recording starts from the left end of the blue rectangular window featuring a blue waveform. Again, you can navigate a cursor to the beginning of the blue waveform, starting at 0.9617 seconds (or 961.7 milliseconds), denoted by the red vertical line. This is the response time for the word “shoulder” in this sound file. (Again, a red arrow and a red box are added for illustrative purposes.)

In many instances, pinpointing the exact starting point can be challenging. In such cases, zooming in to see the enlarged waveform may help. Even then, the exact beginning may not be clear. Measuring response time twice and computing the average is one way to mitigate this uncertainty. It is important to acknowledge that achieving 100% accuracy may not be realistic, but consistency in how you measure response time within your own work is essential.

In some cases, determining the starting point can be challenging. Participants may hesitate before giving their responses (e.g., “Hmm..... shoulder”). Surely, these mumblings do not mark the beginning of the correct responses. If you see more than one waveform in your sound file, be sure to measure the beginning of the correct waveform.

In other cases, some participants give more detailed responses such as “left arm” or self-correct their responses, as previously discussed. For self-corrected responses—if you decide to accept them as correct answers—we would recommend that the beginning of the corrected response is used for measurement. Determining the appropriate starting point for more elaborated responses like “left arm” can be tricky. You can look at the onset of the first word “left” because that is when the participants start producing the target word they have in mind. However, some participants may give their responses more hesitantly, as in “left.....hmm....arm.” Again, it is your responsibility to make measuring decisions, and once you decide what to do, be consistent.

Another inherent concern for our version of the HALA task, which is self-paced, is that participants may click the “next” button too early, while they are still producing the target word. This will truncate the sound file of the target item and it may affect the recording of the following item. To avoid this problem, it is advisable to make sure that the participants are informed of the appropriate timing for clicking the “next” button. We want our participants to respond as quickly as possible, but that is by orally producing the words, not by clicking the “next” button.

CONCLUDING REMARKS

The updated version of the HALA task that we present here, the HALA Online, is designed based on the foundational ideas and assumptions of the original task by O'Grady et al. (2009, pp.109-110). While we have made changes to the format of the task, the original purpose and design remain the same. The most notable difference in the HALA Online compared to previous versions is the transition of the task to an online platform. One of the practical advantages of the original task was its full portability and ease of implementation in various environments. The HALA Online further enhances the practicality by not requiring physical and synchronous presence during data collection. However, the HALA Online requires internet access, and thus may not be a feasible option for data collection in more remote areas. Yet researchers may download the materials for the present version of the task from UH ScholarSpace and implement the task on a different platform that does not require internet connectivity. We thus hope that the current version of the task will serve researchers in as wide a range of contexts as possible, including not only language conservation and documentation, but also bi- and multilingualism in various heritage and diasporic language communities, where remote testing can offer a significant advantage for reaching a greater number of speakers. Additionally, the change to line drawings of stimuli makes the images more neutral in context and more appropriate for use with speakers of any age. This change also allows the HALA Online to be used in more contexts where photo images may have been distracting or culturally inappropriate. It is our hope that this version of the task can be more accessible to researchers and the bilingual communities with which they work while also being easy to use and score.

REFERENCES

- Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, 39(3), 445–459.
<https://doi.org/10.3758/BF03193014>
- Dubiel, B., & Guilfoyle, E. (2017). Language Strength in Bilingual Children: The Child HALA Test. *Heritage Language Journal*, 14(1), 1-29.
<https://doi.org/10.46538/hlj.14.1.1>
- Kang, S. G. (2011) *English attrition in Korean-English bilingual children*. (Unpublished doctoral dissertation). University of Hawaii at Manoa.
- Kim, K., & Kim, H. (2022). Sequential bilingual heritage children's L1 attrition in lexical retrieval: Age of acquisition versus language experience. *Bilingualism: Language and Cognition*, 25(4), 537–547.
<https://doi.org/10.1017/S1366728921001139>
- O'Grady, W., Schafer, A. J., Perla, J., Lee, O-S., & Wieting, J. (2009) A psycholinguistic tool for the assessment of language loss: The HALA project. *Language Documentation and Conservation* 3(1), 100–112.
- Zehr, J., & Schwarz, F. (2018). PennController for Internet Based Experiments (IBEX). <https://doi.org/10.17605/OSF.IO/MD832>

APPENDIX

Tally Sheets for the HALA Task Version A

Book (practice item)	
Cup (practice item)	
Hat (practice item)	
Pen (practice item)	
Shirt (practice item)	
Bowl (practice item)	
Leg	
Nose	
Ear	
Face	
Stomach	
Back	
Head	
Hand	
Shoulder	
Teeth	
Tongue	
Lips	
Mouth	
Eye	
Fingers	
Knee	
Foot	
Arm	
Wrist	
Toes	
Ankle	
Heel	
Elbow	
Neck	
Cheek	
Eyebrow	
Forehead	
Chin	
Thumb	
Fingernail	
Palm	

APPENDIX

Tally Sheets for the HALA Task Version B

Book (practice item)	
Cup (practice item)	
Hat (practice item)	
Pen (practice item)	
Shirt (practice item)	
Bowl (practice item)	
Tongue	
Back	
Ear	
Leg	
Eye	
Foot	
Lips	
Hand	
Shoulder	
Teeth	
Face	
Mouth	
Head	
Stomach	
Fingers	
Knee	
Nose	
Heel	
Forehead	
Toes	
Ankle	
Arm	
Palm	
Neck	
Chin	
Eyebrow	
Wrist	
Cheek	
Thumb	
Fingernail	
Elbow	