SiMAmp Project: Simulated amplification in Parkinson’s disease

[Overview of the SiM project 1](#_Toc83587032)

[Overview of Stages 3 & 4: Amplification & Masks 2](#_Toc1078771487)

[Details for Data Collection of Stages 3 & 4: Simulated Recordings 2](#_Toc886760665)

[Conditions 3](#_Toc1419427816)

[Recording conditions 3](#_Toc2100781803)

[Stimuli 3](#_Toc1389475108)

[Filename conventions for audio files 5](#_Toc894746045)

[Summary for Stage 3: Amplification (no masks) 5](#_Toc2117201995)

[Summary for Stage 4: Amplification and masks 5](#_Toc283881189)

[Additional information about Stage 3 audio files 6](#_Toc1510068997)

[Plan for Trifecta project & TODOs 7](#_Toc968560054)

[Additional audio files & resources 7](#_Toc462012899)

# Overview of the SiM project

The purpose of this project will be to quantify auditory-perceptual and acoustic characteristics of speech produced with face masks and/or speech amplification in order to characterize impacts on spoken communication. **This project has three projected stages:**

1) Pre-recorded speech recordings in various mask conditions via HATS

Status: Analysis complete; manuscripts in prep

2) in-person speech recordings with young healthy adults, older healthy adults, and people with PD wearing masks

Status: Analyses complete; manuscripts in press or under review

3) pre-recorded speech from older adults and people with PD via HATS with amplification device (no mask)

4) pre-recorded speech from older adults and people with PD via HATS with amplification device AND masks

***This document outlines the plans for Stage 3 & 4***

**NOTE: The TRIFECTA funded project uses audio from Stage 3 (no masks; just amplified and unamplified speech)**

## Overview of Stages 3 & 4: Amplification & Masks

The current plan is to collect speech recordings via a B&K Head & Torso with mouth simulator device in 3 mask conditions & 2 speech amplification conditions. The goal of using a mouth simulator device will be to (somewhat) mimic the directionality/propagation of human anatomy as well as to be able to realistically place face masks. Stimuli will be recordings collected from the baseline condition of the SiM Stage 3 experiment (Harvard sentences read with no mask in a habitual speaking style by people with PD and older healthy controls).

# Details for Data Collection of Stages 3 & 4: Simulated Recordings

Original recordings from live talkers were collected at UB in Fall 2021

## Conditions

**Mask:** 4 mask conditions: No mask, Surgical mask, KN95, N95

**Amplification:** 2 amplification conditions: No amplification, speech amplified by MiniBuddy device

4 mask x 2 amp = 8 conditions

**Noise:** 2 noise conditions: No noise, 5dB SNR multitalker babble

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Project stage** | **Mask + No Amplification** | **Code** | **Mask + Amplification** | **Code** |
| 3 | 1. No mask | NM\_na | 5. No mask + Minibuddy | NM\_amp |
| 4 | 1. Surgical mask | SM\_na | 6. Surgical mask + Minibuddy | SM\_amp |
| 4 | 1. KN95 | KN\_na | 7. KN95 + Minibuddy | KN\_amp |
| 4 | 1. N95 | N95\_na | 8. N95 + Minibuddy | N95\_amp |

## Recording conditions

Each stimulus in each mask condition will be recorded at a **2 meter distance** using the B&K device in the following conditions

|  |  |
| --- | --- |
| **Recording condition** | **Code** |
| 1. No Noise | NN |
| 1. In 5 dB SNR multitalker babble noise | babble5 |

## Stimuli

We will use 11 PD and 10 OC (older control) talkers from the SiM study reading 1 Harvard list

Each group includes 6 male and 4 female talkers (PD group includes one extra male)

Sentence lists in the no mask, habitual speech condition

Utterances have been rescaled to 70 dB SPL

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Participant** | **Speaker gender** | **Speaker age** | **Harvard Lists** | **Missing files** |
| PD01 | M | 79 | H18 | - |
| PD04 | M | 63 | H15 | - |
| PD05 | M | 67 | H2 | - |
| PD06 | F | 69 | H16 | - |
| PD08 | F | 74 | H14 | Missing 1 sentence |
| PD09 | M | 68 | H14 | - |
| PD11 | M | 65 | H15 | - |
| PD12 | M | 66 | H10 | Missing 1 sentence |
| PD14 | F | 73 | H18 | - |
| PD15 | M | 67 | H15 | Missing 1 sentence |
| PD16 | F | 70 | H2 | Missing 1 sentence |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Participant** | **Speaker gender** | **Speaker age** | **Harvard Lists** | **Missing files** |
| OC02 | M | 66 | H17 | Missing 1 sentence |
| OC04 | M | 60 | H15 | Missing 1 sentence |
| OC05 | M | 64 | H2 | Missing 1 sentence |
| OC06 | F | 63 | H16 | - |
| OC07 | F | 73 | H12 | - |
| OC08 | F | 57 | H14 | - |
| OC09 | F | 64 | H14 | - |
| OC11 | M | 58 | H15 | - |
| OC16 | M | 67 | H2 | Missing 1 sentence |
| OC17 | M | 74 | H12 | - |

*8 sentence files removed in total (due to major hesitations, misreadings)*

## Transcriptions of stimuli

* Available in [Harvard\_Sentences\_Keywords\_CASALab\_homophones.csv](https://michiganstate.sharepoint.com/:x:/r/sites/CASALab/Shared%20Documents/Projects/simamp/Harvard_Sentences_Keywords_CASALab_homophones.csv?d=wb826d967e83d48e4ae3db68935c1218f&csf=1&web=1&e=9E5vmn)

## Filename conventions for audio files

Assuming each audio recording contains an entire speaker stimulus set, here is the suggestion (this can change):

**Original file name convention:**

* **project\_participant\_speechCondition\_maskCondition\_list\_sentence\_ch\_70dB**
* Ex: sim\_pd15\_habitual\_nm\_h15\_10\_ch1\_70db.wav

**Revised file name convention:**

* **project\_maskCondition\_ampCondition\_noiseCondition\_participant\_list\_sentence**
  + *Project is always sim*
  + *Mask condition is always NM for Stage 3*

**Example filenames:**

* sim\_NM\_na\_babble5\_PD01\_H07\_01.wav
  + SiM project, No Mask, No amplification, in +5dbSNR babble noise, PD01 participant reading from Harvard Sentence list 7, sentence #1
* sim\_n95\_amp\_nn\_OC05\_H18\_09.wav

## Summary for Stage 3: Amplification (no masks)

2 amplification conditions x 2 noise conditions: 4 conditions

21 talkers x 10 sentences – 8 missing sentences: 202 unique stimuli

4 x 202 = 808 total audio files

* Noise files used for intelligibilty experiment
* No noise files used for acoustic analysis

## Summary for Stage 4: Amplification and masks

3 mask conditions x 2 amplification conditions x 2 noise conditions: 12 conditions

12 conditions x 202 unique stimuli = 2424 total audio files

# Additional information about Stage 3 audio files

* Original recordings had 0.05 seconds of silence added to beginning and end as a buffer
* HATS recordings had additional ~ 1 second of silence added; this appears to be asymmetrical due to the way the recording loop cut the audio; about 75ms at the beginning and 25 – 35 ms at the end

# Plan for Trifecta project & TODOs

1. Collect ASR accuracy measures (Qiben)

2. Collect human intelligibility accuracy measures (Thea)

3. Collect acoustic measures (as per Gaballah et al., 2019) (Q/T)

4. Prepare audio processing plans, such as amplifying specific frequency ranges

* TBD:
  + do we want to modify the unamplified or amplified recordings?
  + Planned acoustic enhancements:
    - Increased spectral tilt (parameters TBD)
    - Increased energy in 1-3khz

# Additional audio files & resources

This details additional recordings that are related to the project but are not the central recordings of interest

[Baseline recordings of people with and without PD](https://michiganstate.sharepoint.com/:f:/r/sites/CASALab/Shared%20Documents/Projects/simamp/sim_production_audio_rescaled_70dB?csf=1&web=1&e=gq48iV): These recordings were collected in my lab last year and include a subset of individuals with and without PD reading aloud [Harvard Sentences](https://www.cs.columbia.edu/~hgs/audio/harvard.html) (phonetically balanced sentence lists). These audio files will be the input for the Head & Torso Simulator which will be fitted with an amplification device and re-recorded at a 6 foot distance.

* Note that these recordings are also part of a larger study that also looks at the effects of face masks on speech in PD.
* These will give you a sense of what the recordings will contain (but not the resulting quality).
* The version of these that we will be working with will include re-recorded versions from the HATS in 2 conditions: unamplified and amplified.
  + We will also have several face mask conditions as well, which we can also talk about if you would be interested in being involved with the larger project too.

[Amplification device recordings of a single healthy talker](https://michiganstate.sharepoint.com/:f:/r/sites/CASALab/Shared%20Documents/Projects/kamp/5_KAMP_extracted_sentences?csf=1&web=1&e=rOoacT): These include recordings of a single, neurologically healthy, older adult male speaker reading aloud sentences from the [CAPE-V](https://www.asha.org/siteassets/uploadedfiles/asha/sig/03/cape-v-procedures-and-form.pdf). These recordings were collected and rescaled to 70 dB, then played via a Head & Torso simulator wearing several different amplification devices. The project name is KAmp, which refers to Kemar Amplification (Kemar is the name of the HATS used here; we will be using B&K for the current project).

* These recordings will give you a sense of what amplified speech sounds like.
  + In the filenames, MAX = the maximum volume possible, whereas GBF = the gain-before-feedback; some of the devices high volume levels would result in squealing feedback, so we also recorded the utterances with the devices set to the GBF levels, which is what an actual user would likely select as the highest usable volume.
* We also have subjective auditory-perceptual data for these utterances that I have begun exploring (listener estimates of loudness, sound quality, understandability, and effort to understand).
* The amplification device we selected was the MiniBuddy, which was selected to account for a trade-off in sound quality and user qualities like portability. Happy to go into more detail on this decision if you’d like.

**Background readings**

* Andreetta, M. D., Adams, S. G., Dykstra, A. D., & Jog, M. (2016). Evaluation of Speech Amplification Devices in Parkinson’s Disease. *American Journal of Speech-Language Pathology*, *25*(1), 29–45.<https://doi.org/10.1044/2015_AJSLP-15-0008>
* Gaballah, A., Parsa, V., Andreetta, M., & Adams, S. (2019). Objective and Subjective Speech Quality Assessment of Amplification Devices for Patients With Parkinson’s Disease. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *27*(6), 1226–1235.<https://doi.org/10.1109/TNSRE.2019.2915172>
* Greene, M. C. L., & Watson, B. W. (1968). The Value of Speech Amplification in Parkinson’s Disease Patients. *Folia Phoniatrica et Logopaedica*, *20*(4), 250–257.<https://doi.org/10.1159/000263203>
* Greene, M. C., Watson, B., Gay, P., & Townsend, D. (1972). A therapeutic speech amplifier and its use in speech therapy. *The Journal of Laryngology & Otology*, *86*(6), 595–605.
* Iddon, H., Read, J., & Miller, N. (2015). Does voice amplification increase intelligibility in people with Parkinson’s disease? *International Journal of Therapy and Rehabilitation*, *22*(10), 479–486.<https://doi.org/10.12968/ijtr.2015.22.10.479>
* Knowles, T., Adams, S. G., Page, A., Cushnie-Sparrow, D., & Jog, M. (2020). A Comparison of Speech Amplification and Personal Communication Devices for Hypophonia. *Journal of Speech, Language, and Hearing Research*, *63*(8), 2695–2712.<https://doi.org/10.1044/2020_JSLHR-20-00085>