

**Office of the Vice Chancellor for Research & Innovation**

Office for the Protection of Research Subjects  
1901 S. First St., Suite A, MC-685  
Champaign, IL 61820

**Notice of Approval: Amendment 24**

April 16, 2024

<b>Principal Investigator</b>	Mark Hasegawa-Johnson
<b>CC</b>	Heejin Kim, Chris Zwilling
<b>Protocol Title</b>	<i>NOVA HABILITAS (NOVA H) Consortium</i>
<b>Protocol Number</b>	23183
<b>Funding Source</b>	Industry Consortium
<b>Review Type</b>	Expedited 6, 7
<b>Amendment Requested</b>	Adding ADAPT Community Network as an approved site
<b>Status</b>	Active
<b>Risk Determination</b>	No more than minimal risk
<b>Amendment Approval Date</b>	April 16, 2024
<b>Expiration Date</b>	March 2, 2028
<b>Sites Approved</b>	University of Illinois Urbana-Champaign Davis Phinney Foundation LSVT Global Team Gleason Lingraphica ADAPT Community Network

This letter authorizes the use of human subjects in the above protocol. The University of Illinois at Urbana-Champaign Institutional Review Board (IRB) has reviewed and approved the research study as described.

The Principal Investigator of this study is responsible for:

- Conducting research in a manner consistent with the requirements of the University and federal regulations found at 45 CFR 46.
- Using the approved consent documents, with the footer, from this approved package.
- Requesting approval from the IRB prior to implementing modifications.
- Notifying OPRS of any problems involving human subjects, including unanticipated events, participant complaints, or protocol deviations.
- Notifying OPRS of the completion of the study.



---

IRB Number: 23183

## Human Subjects Research – Protocol Form

**Guidelines for completing this research protocol:**

- Please submit typed applications via email. Handwritten forms and hard copy forms will not be accepted.
- For items and questions that do not apply to the research, indicate as “not applicable.”
- Provide information for all other items clearly and avoid using discipline specific jargon.
- Please only include text in the provided boxes. The text boxes will expand as they are typed in to accommodate large amounts of text.

**Before submitting this application, ensure that the following have been completed.**

- Protocol Form is complete.
- Relevant CITI modules have been completed for all members of the research team at [www.citiprogram.org](http://www.citiprogram.org).
- Informed consent/assent/parental permission document(s) are provided.
- Relevant waivers and appendices are provided.
- Recruitment materials are provided.
- Research materials (e.g. surveys, interview guides, etc.) are provided.
- Any relevant letters of support are provided.

Instructions on the non-exempt review process and guidance to submitting applications, can be found on the OPRS [website](#). You may also contact OPRS by email at [irb@illinois.edu](mailto:irb@illinois.edu) or phone at 217-333-2670.

**Submit completed applications via email to:** [irb@illinois.edu](mailto:irb@illinois.edu).



Office for the Protection  
of Research Subjects

# Protocol Form

## Section 1: PRINCIPAL INVESTIGATOR (PI)

<p>The Illinois <a href="#">Campus Administrative Manual</a> allows assistant, associate, and full professors to act as PI. Other individuals may serve as PI after obtaining approval from the necessary party.</p>			
Last Name: Hasegawa-Johnson	First Name: Mark	Degree(s): Ph.D.	
Dept. or Unit: Beckman Institute	Office Address: 2011 Beckman		
Street Address: 405 N Mathews Av	City: Urbana	State: IL	Zip Code: 61801
Phone: 217-333-0925	E-mail: <a href="mailto:jhasegaw@illinois.edu">jhasegaw@illinois.edu</a>		
Urbana-Champaign Campus Status: Non-visiting member of (Mark One) <input checked="" type="checkbox"/> Faculty <input type="checkbox"/> Academic Professional/Staff <i>(Student Investigators cannot serve as PI)</i>			
Training <input checked="" type="checkbox"/> Required CITI Training, Date of Completion (valid within the last 3 years), 3/19/22 <input type="checkbox"/> Additional training, Date of Completion,			

## Section 2. RESEARCH TEAM

<b>2A. Are there other investigators engaged in the research?</b> <input checked="" type="checkbox"/> Yes (include a <a href="#">Research Team Form</a> ) <input type="checkbox"/> No
<b>2B. If yes, are any of the researchers not affiliated with Illinois?</b> <input checked="" type="checkbox"/> Yes <input type="checkbox"/> No

## Section 3. PROTOCOL TITLE

NOVA HABILITAS (NOVA H) Consortium
------------------------------------

## Section 4. FUNDING SOURCE

<b>4A. Is the research funded?</b> <input type="checkbox"/> Research is <b>not funded</b> and is <b>not pending</b> a funding decision (Proceed to Section 5). <input checked="" type="checkbox"/> Research is <b>funded</b> (funding decision has been made). <input type="checkbox"/> Funding decision is <b>pending</b> . Funding proposal submission date:
<b>4B. Indicate the source of the funding.</b> <input type="checkbox"/> University of Illinois Department, College or Campus, <i>please specify</i> : <input type="checkbox"/> Federal, <i>please specify</i> : <input checked="" type="checkbox"/> Commercial Sponsorship & Industry <sup>1,2</sup> , <i>please specify</i> : Industry consortium MOU.

<sup>1</sup> Clarify whether or not sponsor requires specific language in the contractual agreement that impacts human subjects research

<sup>2</sup> Clarify whether or not the sponsor requires the protocol adhere to ICH GCP (E6) standards



# Protocol Form

<input type="checkbox"/> State of Illinois Department or Agency, <i>please specify</i> :
<input type="checkbox"/> Other, <i>please specify</i> :
<b>4C. Sponsor-assigned grant number, if known:</b>
<b>4D. A complete copy of the funding proposal or contract is attached.</b>
<input checked="" type="checkbox"/> Attached, <i>please specify title</i> : Membership Agreement for Artificial Intelligence Accessibility Coalition
<b>4E. Funding Agency Official To Be Notified of IRB Approval (if Applicable)</b>
Name:
Agency:
E-mail:
Phone:

## Section 5. CONFLICTS OF INTEREST

Please indicate below whether any investigators or members of their immediate families have any of the following. If the answer to any of the following items is yes, please submit the University of Illinois approved conflict management plan. If you have any questions about conflicts of interest, contact [coi@illinois.edu](mailto:coi@illinois.edu).

<b>5A. Financial interest or fiduciary relationship with the research sponsor (e.g. investigator is a consultant for the research sponsor).</b> <input type="checkbox"/> Yes <input checked="" type="checkbox"/> No
<b>5B. Financial interest or fiduciary relationship that is related to the research (e.g. investigator owns a startup company, and the intellectual property developed in this protocol may be useful to the company).</b> <input type="checkbox"/> Yes <input checked="" type="checkbox"/> No
<b>5C. Two or more members of the same family are acting as research team members on this protocol.</b> <input type="checkbox"/> Yes <input checked="" type="checkbox"/> No

## Section 6. SUMMARY & PURPOSE OF RESEARCH

**6A. In lay language, briefly summarize the objective and significance of the research.**

The goal of the Speech Accessibility Project at the UIUC Beckman Institute (<https://speechaccessibilityproject.beckman.illinois.edu>) is to collect, annotate, and curate a shared database of speech samples from people with atypical speech, and share this data set with researchers at other organizations. This two-year project plans to collect 1,200,000 speech samples from 2,000 people, each of whom will provide 600 samples. In Year 1, the initial focus will be people with Parkinson's. In Year 2, four more etiologies of interest will be recruited (such as ALS patients, stroke patients, etc.). UIUC will build an open-source software infrastructure to collect annotated speech samples and share these data in an appropriately secure fashion with researchers from our partner technology companies (and eventually, other organizations as well) so that they can use these data to improve their automatic speech recognition algorithms. This project promotes diversity, equity, and inclusion by helping technology companies to fully support all types of speech, and it is also more efficient and less burdensome for these specialized patient populations to have one centralized "collector" of speech samples.

**6B. Indicate if your research includes any of the following:**

- Secondary data (use of data collected for purposes other than the current research project)
- Data collected internationally (include [International Research Form](#))
- Translated documents (include [Certificate of Translation Form](#) and translated documents)
- Research activities will take place at Carle (include documentation (email or letter) from Carle stating that the review of your [Research Services Request Form](#) is complete)

**6C. Letters of support from outside institutions or entities that are allowing recruitment, research, or record access at their site(s) are attached.**  Yes  Not Applicable

## Section 7. PROCEDURES

**7A. Select all research methods and/or data sources that apply.**

- Surveys or questionnaires, *select all that apply:*  Paper  Telephone  Online
- Interviews
- Focus groups
- Field work or ethnography
- Standardized written, oral, or visual tests
- Taste or smell testing
- Intervention or experimental manipulation
- Exercise and muscular strength testing
- Noninvasive procedures to collect biological specimens (e.g., hair and nail clippings, saliva, etc.)
- Noninvasive procedures to collect physiological data (e.g., physical sensors, electrocardiography, etc.)
- Procedures involving radiation
- Recording audio and/or video and/or taking photographs: Yes
- Recording other imaging
- Materials that have already been collected or already exist, *specify source of data:*
- [HIPAA-protected data](#)
- [FERPA-protected data](#)
- [GDPR-protected data](#)
- Other, *please specify:*

**7B. List all testing instruments, surveys, interview guides, etc. that will be used in this research.**

Drafts or final copies of all research materials are attached.  Yes

**7C. List approximate study dates.** February 2023 – August 2024**7D. What is the duration of participants' involvement?** Up to 5 hours, which is the time estimated to be trained on the system and then use it to provide 600 speech samples, where each sample is approximately 5 to 10 seconds of speech.**7E. How many times will participants engage in research activities?** Multiple times as needed at their own pace and convenience to record speech samples.**7F. Narratively describe the research procedures in the order in which they will be conducted.**



# Protocol Form

The goal of our project is to collect 1,200,000 speech samples from 2,000 people with dysarthria, where we expect to collect data from 400 people each from five different patient populations. Each person would provide 600 speech samples.

(600 samples/person x 400 persons/etiology x 5 etiologies = 1,200,000 samples)

Our schedule of research procedures is:

1. February or March-August 2023: data collection of speech samples from 400 people with Parkinson's. Parkinson's Disease is "etiology 1".
2. August 2023-August 2024: data collection of speech samples from 1,600 people with other etiologies to be determined (i.e., voted on by our Advisory Committee). One of these etiologies will be ALS. The other three etiologies are to be determined.

In Year 1, we focus on people with Parkinson's and therefore, recruitment will be managed by our partner the Davis Phinney Foundation, a well-known non-profit advocacy group for people with Parkinson's. DPF will distribute our recruitment flyer (see **Attachment 3, recruitment for speech**) and engage in informational sessions with potential participants. (In Year 2, we will do the same procedure for other populations of interest, such as ALS patients. Therefore, in Year 2, we will add additional patient advocacy organizations to assist in recruitment, such as Team Gleason to recruit patients with ALS.)

Data collection of speech samples in Year 1 will be a collaboration of UIUC and LSVT Global team members. Potential participants will be screened both with a questionnaire and by providing a short set of "quality control" speech samples (**Attachment 4**). If the participant does not pass screening, they will be thanked for their interest. Otherwise, the participant is eligible for the study and can do the informed consent process (**Attachment 5**) and then engage in contributing speech samples (**Attachment 7**).

Participants can do as many recordings as they wish at whatever time of day is convenient for them. Participants will be able to login to the system at any time, 24/7.

In Year 2, this procedure will be performed with patients from other etiologies with Team Gleason Foundation, Lingraphica, and Adapt Community Network as partners.

Participants also have the option to provide additional data about themselves, such as their age, race and ethnicity, year of their diagnosis, and other information as shown in **Attachment 9**. These "metadata tags" are completely optional but are helpful for analysis.

The collected speech samples will be stored securely in a custom database built by the UIUC Beckman Institute. All samples are stored with a unique participant ID code along with any of the optional



metadata tags. All samples are annotated by our UIUC research team with technical information about the acoustic waveform and other information.

The entire database of speech samples will be shared with our coalition partners (Amazon, Apple, Google, Meta, and Microsoft). Each partner has signed a data use agreement with UIUC that allows these deidentified data to be used for improvements in speech recognition technology and assures the privacy of participants and confidentiality of data. In the future, other partners may join this coalition, subject to a vote by the current members, and those future partners would also sign a data use agreement with the same terms and conditions.

## Section 8. PERFORMANCE SITES TO INCLUDE INTERNATIONAL, SCHOOL, AND COLLABORATIVE STUDIES

**8A. List all research sites for the protocol. For non-University of Illinois at Urbana-Champaign sites, describe their status of approval and provide contact information for the site. If the site has an IRB, note whether the IRB has approved the research or plans to defer review to the University of Illinois at Urbana-Champaign.**

### Performances Sites

- #1 University of Illinois Urbana-Champaign
- #2 Davis Phinney Foundation, Colorado
- #3 LSVT Global, Colorado and Arizona
- #4 Team Gleason Foundation, Louisiana

If there are additional performance sites, include them on an attachment and check here:

**8B. Is this a multi-center study in which the Illinois investigator is the lead investigator, or the University of Illinois at Urbana-Champaign is the lead site?  Yes  No**

If yes, answer 8C and 8D. If no, proceed to Section 8E.

**8C. Who is the prime recipient of funding, if funded? UIUC**

**8D. What is the management and communication plan for information that might be relevant to the protection of research subjects (e.g. unanticipated problems involving risks to subjects, interim results, and protocol modifications)?** UIUC will issue subawards to DPF, LSVT Global and Team Gleason Foundation. UIUC, DPF, LSVT Global and Team Gleason will establish a regular series of Zoom meetings with shared documents on Box or Teams. Any issues with participant engagement will be discussed regularly and remediated as appropriate. UIUC is responsible for monthly reports out to our founding partners, the Artificial Intelligence Accessibility Consortium.

**8E. If subjects will be recruited from Illinois public or private elementary or secondary schools, additional deadlines and procedures may apply. Criminal background clearances might be required. Special consideration must be given to the exclusion of protected populations. Please contact the School University Research Relations ([researchplacements@education.illinois.edu](mailto:researchplacements@education.illinois.edu)) for more information. Select one:**  Illinois schools **will** be used  Illinois schools **will not** be used



# Protocol Form

## Section 9. SUBJECT ENROLLMENT GOAL & EQUITABLE SELECTION OF SUBJECTS

9A. For each performance site, indicate the estimated total number of participants.			
Performance Site	# Male	# Female	Total
#1 UIUC (20 for focus groups, 2000 for speech samples)	1010	1010	2020
#2 Davis Phinney Foundation collaborates with UIUC on recruitment	1000	1000	2000
#3 LVST Global collaborates with UIUC on focus groups, intake and collection of speech samples	1010	1010	2020
#4 Team Gleason collaborates with UIUC on focus groups, intake and collection of speech samples	1010	1010	2020
#5 Lingraphica collaborates with UIUC on focus groups, intake, and collection of speech samples	1010	1010	2020
#6 Adapt Community Network collaborates with UIUC on recruitment	1010	1010	2020
<b>TOTALS</b>	1010	1010	2020
If additional performance sites are included on an attachment, check here: <input type="checkbox"/>			
9B. Select all participant populations that will be recruited.			
<b>Age:</b>			
<input checked="" type="checkbox"/> Adults (18+ years old) <input type="checkbox"/> Minors ( $\leq 17$ years old) <input type="checkbox"/> Specific age range, <i>please specify:</i>			
<b>Gender:</b>			
<input checked="" type="checkbox"/> No targeted gender (both men and women will be recruited/included) <input type="checkbox"/> Targeted gender, <i>please indicate:</i> <input type="checkbox"/> Men/boys <input type="checkbox"/> Women/girls <input type="checkbox"/> Other, <i>please specify:</i>			
<b>Race/Ethnicity:</b>			
<input checked="" type="checkbox"/> No targeted race or ethnicity (all races and ethnicities will be recruited/included) <input type="checkbox"/> Targeted race or ethnicity, <i>please specify:</i>			
<b>College Students:</b>			
<input checked="" type="checkbox"/> No targeted college population <input type="checkbox"/> UIUC general student body <input type="checkbox"/> Targeted UIUC student population, <i>provide the instructor or course information, name of the departmental subject pool, or other specific characteristics:</i> <input type="checkbox"/> Students at institution(s) other than UIUC, <i>please specify:</i>			
<i>Any research with students on UIUC's campus needs to be registered with the </i> <a href="#"><i>Office of the Dean of Students.</i></a>			



**Other:**

- Inpatients
- Outpatients
- People who are illiterate or educationally disadvantaged
- People who are low-income or economically disadvantaged
- People with mental or cognitive disabilities or otherwise impaired decision-making capacities
- Adults with legal guardians
- People who are non-English speaking
- People with physical disabilities
- Pregnant or lactating women, human fetuses, and/or neonates
- Prisoners or people with otherwise limited civil freedoms
- Other, *please specify*: Adults with caregivers to help participants use computers

**9C. Describe additional safeguards included in the protocol to protect the rights and welfare of the populations selected above.**

We have designed the remuneration plan to include support for caregivers.

Our subawards with DPF, LSVT Global, Team Gleason, Lingraphica and Adapt Community Network are meant to provide extra support for potential participants. DPF's, Team Gleason's, Lingraphica's and Adapt Community Network's recruitment activities include providing information and support as needed to potential participants to answer their questions. If a potential participant does not seem able to provide informed consent, either on their own or without the assistance of caretaker or legal guardian, they will not be recruited. LSVT Global, Team Gleason and Lingraphica will provide an intake coordinator and a number of mentors to support data collection. We anticipate that many participants will need advice and support during their attempts at data collection online, and so they will be able to schedule time with a mentor to get help.

## Section 10. INCLUSION/EXCLUSION

**10A. List specific criteria for inclusion and exclusion of subjects in the study, including treatment and control groups.**

**For the main study to collect speech samples, total n=2000 where Cohort 1 (n=400) etiology is Parkinson's Disease, and Cohort 2 etiologies are Down syndrome, ALS, stroke or aphasia and cerebral palsy (n=1600) for Year 2.**

Inclusion criteria:

- Adult (age  $\geq$  18 years)
- Self-reported diagnosis of Parkinson's Disease or other etiology of interest
- Reads or speaks English
- Has a valid email address



- Ability to access web browser to participate in study

Exclusion criteria:

- Is a resident of the State of Washington, Texas, or Illinois (because these states have privacy laws that would not allow us to collect ‘voice prints’)
- If quality control screening of initial speech samples “fails” because of poor data quality (e.g., poor quality recording environment, or person’s speech is “too typical” and not sufficiently interesting to continue collecting)

**10B. Explain how the inclusion/exclusion criteria will be assessed and by whom. If special expertise is required to evaluate screening responses or data, list who will make this evaluation and describe their training and experience.**

Assessing the inclusion/exclusion criteria will be in collaboration with patient advocacy organizations. In Year 1, that will be Davis Phinney Foundation and LSVT Global. In Year, that will be Team Gleason, Lingraphica and Adapt Community Network.

**10C. Drafts or final copies of all screening materials are attached.**  Yes  Not Applicable

**10D. Describe procedures to assure equitable selection of subjects. Justify the use of the groups marked in Section 9B. Selection criteria that target one sex, race, or ethnic group require a clear scientific rationale.**

We will offer participants the optional ability to provide demographic data, so that we can assess the diversity and equitable representation of our sample.

## Section 11. RECRUITMENT

**11A. Select all recruitment procedures that will be used.**

- Student subject pool, *please specify*:
- Email distribution
- MTurk, Qualtrics Panel, or similar online population, *please specify*:
- US Mail
- Flyers/brochures
- Website ad, online announcement (e.g. eWeek), or other online recruitment, *please specify*: newsletter announcements via organizations that advocate for and support people with disabilities.
- Newspaper ad
- Verbal announcement
- Other, *please specify*:
- Not applicable (secondary data only)

**11B. Drafts or final copies of all recruitment materials (including verbal scripts) are attached.**

Yes  Not Applicable

**11C. For each group of participants, describe the details of the recruitment process.**



In Years 1 and 2, Davis Phinney Foundation, Team Gleason and Lingraphica will include our recruitment flyer in their distribution channels, will also host optional information and Q&A sessions with potential participants, and will answer individual questions as needed from potential participants. The UIUC Speech Accessibility Project's website will remain as a public source of information, including an updated FAQ list that may be updated based on feedback from advocacy organizations.

## Section 12. REMUNERATION AND PLAN FOR DISTRIBUTION

Refer to the University [Business and Financial Policies and Procedures](#) for further guidance on the compensation process and reporting requirements.

**12A. Will subjects receive inducements or rewards before, during, or after participation?**

Yes  No

If yes, complete the rest of Section 12. If no, proceed to Section 13.

**12B. Select all forms of remuneration that apply.**

- Cash, please specify amount:  
 Check, please specify amount:  
 Gift Certificate, please specify amount: Amazon ecodes adding up to \$180 per data contributor and up to \$90 per caregiver.  
 Lottery, please specify amount: and odds:  
 Course Credit, please specify amount: and specify equivalent alternative activity:  
 Other, please specify:

**12C. Will payment be prorated before, during, or after participation?**

- Yes, please specify how: prorated by number of recorded sentences uploaded to our system  
 No

**12D. For each group of participants, describe the details of the remuneration plan, including how, when and by whom they will be notified.**

Our project plans to obtain a total of 600 recorded sentences per participant. Each participant will be compensated with up to \$180: \$60 after recording the 200<sup>th</sup> sentence, an additional \$60 after recording their 400<sup>th</sup> sentence, and an additional \$60 after recording their 600<sup>th</sup> sentence.

In addition to the participant compensation described above, participants will also be able to name a caregiver who is assisting them with performing the computer usage tasks in this study, and to specify additional compensation to the caregiver. Named caregivers will be compensated with up to \$90: \$30 after recording the 200<sup>th</sup> sentence, an additional \$30 after recording their 400<sup>th</sup> sentence, and an additional \$30 after recording their 600<sup>th</sup> sentence.

The NOVA H customized database system will track number of sentences per participant that are uploaded. The database will have a regular schedule of quality control checks and a built-in alert to remind the NOVA H project manager that payments are due as appropriate. The NOVA H project manager will then use the usual UIUC procedure to compensate the participants via automated email messages containing an appropriate Amazon ecode.



The Beckman Institute is fully aware of OBFS guidance on potential IRS tax implications of human subject payments. The NOVA H project manager will assure that the team follows OBFS guidance and will work closely with the Beckman business affairs office to ensure that tax implications are addressed.

**12E. The information listed above is provided on the relevant consent forms.**

Yes

## Section 13. RISKS & BENEFITS

**13A. Describe all known risks to the participants for the activities proposed, such as risks to the participants' physical well-being, privacy, dignity, self-respect, psyche, emotions, reputation, employability, and criminal and legal status. Risks must be described on consent forms.**

Risks include (1) the risk of loss of privacy in case any research institution that has been granted data access violates their data use agreement by releasing participant speech samples to people not covered by the agreement, and (2) the risk that subjects may get tired while recording sentences.

**13B. Describe the steps that will be taken to minimize the risks listed above.**

(1) The risk of loss of privacy will be minimized in two ways. First, we recognize that speech audio recordings are considered to be personally identifiable information, but no other personally identifiable information will be stored in the same database with the speech audio recordings. Personally identifiable information other than the speech audio recordings will be stored securely in a database at the Beckman Institute, which will only be accessible to personnel named in this IRB application. Second, before any research institution is permitted to receive any speech audio recordings, they will be required to sign a data use agreement. Key terms of the data use agreement will include commitments that the research institution (a) is an organization or individual with the legal status necessary to sign a contract, (b) will not make the speech audio recordings available to any individual who is not bound by the member's signature on the data use agreement, (c) will store the data in a secure fashion to prevent data theft, and (d) will not seek to identify any of the participants.

(2) The risk of tiredness will be minimized by designing our user interface so it will allow participants to log in to the system at their convenient time and location. They will be allowed to record at their own pace. They can log out at any time, then log in again to continue recording as they wish.

**13C. Indicate the risk level.**

**No more than minimal risk**

(The probability and magnitude of harm or discomfort anticipated for participation in the proposed research are not greater in and of themselves than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests).

**More than minimal risk (answer 13D)**



**13D. If you checked that the research is more than minimal risk, describe the provisions for monitoring the data to ensure the safety of subjects, such as who will monitor data and how often, what criteria will be used to stop the research, etc.**

**13E. Describe the expected benefits of the research to the subjects and/or to society.**

There is no immediate benefit of the proposed research to the participants, but participants may eventually benefit from improved communication tools that will be developed based on the outcome of the proposed research. Potential benefits to society and people with speech disorders include automatic speech recognition (ASR) technologies of which accuracy will outperform any existing ASR tools applied to disordered speech.

**13F. Weigh the risks with regard to the benefits. Provide evidence that benefits outweigh risks.**

The benefit to society of the proposed research is the development of improved communication tools for people with speech disorders, while no more than minimal risk is anticipated for participation in the proposed research. Thus, the potential benefits outweigh the minimal risks.

## Section 14. INFORMED CONSENT PROCESS TO INCLUDE: WAIVERS, ASSENTS, ALTERATIONS, ETC.

**14A. Indicate all that apply for the consent/assent/parental permission process.**

- Written informed consent (assent) with a document signed by
  - adult subjects
  - parent(s) or guardian(s)
  - adolescents aged 8–17 years
- Waiver of Documentation (signature) of Informed Consent (*include the relevant [Waiver Form](#)*)
  - adult subjects
  - parent(s) or guardian(s)
  - adolescents aged 8–17 years
- Waiver of Informed Consent (*include the relevant [Waiver Form](#)*)
  - adult subjects
  - parent(s) or guardian(s)
  - adolescents aged 8–17 years
- Alteration of Informed Consent (*include the relevant [Alteration Form](#)*)
  - adult subjects
  - parent(s) or guardian(s)
  - adolescents aged 8–17 years

**14B. List all researchers who will obtain consent/assent/parental permission from participants.**

DPF, Team Gleason, and Lingraphica personnel will publicize the opportunity to their clients and offer support as needed, such as group information sessions and one-on-one dialog.

**14C. Describe the method for obtaining consent/assent/parental permission.**

The consent procedure will be conducted online with optional support as needed. Advocacy organizations for the project include DPF, LSVT, Team Gleason and Lingraphica. Advocacy organizations will generate many referrals of potential participants who self-report that they meet initial inclusion criteria. These referrals will be kept as a list of names and email addresses that are shared with the advocacy organizations and UIUC. The intake coordinator for each advocacy organization will contact each potential participant to verify their information and ensure that they provide the quality control samples of their speech in the online UIUC system. UIUC personnel will assess these two pre-screening data sources (initial questionnaire and quality control samples) and decide whether or not the participant is eligible. If not, then the participant is thanked for their interest. If the person is eligible, then the person's information is confirmed with the advocacy organization and the person is directed to the online informed consent document with the option to get help. While it is possible for the person to



read and sign the online informed consent document and immediately begin data collection through the UIUC online system all by themselves, it is more likely that the intake coordinator for the advocacy organization will schedule time with their mentor or UIUC research team member to talk the person through the informed consent document and help them with the first few sessions of data collection.

**14D. Describe when consent/assent/parental permission will be obtained.** A pre-screening procedure will happen first, and once a participant is deemed eligible, then they will be offered the opportunity to review and sign the informed consent document.

**14E. Will participants receive a copy of the consent form for their records?**

Yes  No, if no, explain:

**14F. Indicate factors that may interfere or influence the collection of voluntary informed consent/assent/parental permission.**

- No known factors  
 Research will involve students enrolled in a course or program taught by a member of the research team  
 Research will involve employees whose supervisor(s) is/are recruiting participants  
 Participants have a close relationship to the research team  
 Other, specify any relationship that exists between the research team and participants:

If applicable, describe the procedures to mitigate the above factors.

**14G. Copies of the consent form(s) are attached.**  Yes  Not applicable

**14H. Will this project be registered as a clinical trial?**  Yes  No

If yes, effective January 21, 2019, an informed consent form must be posted on the Federal Web site after the clinical trial is closed to recruitment, and no later than 60 days after the last study visit.

LSVT Global has told us that in order for the Michael J. Fox Foundation to become an additional future recruitment partner for this project, that we must register our study as a clinical trial. Therefore, we intend to do this registration after we receive UIUC IRB approval.

## Section 15. DEVICES & DRUGS

**Indicate if your research includes any of the following.**

- Equipment [Researchers collecting physiological data, not testing the device]  
*(include Appendix A, the [Research Equipment Form](#))*  
 Devices [Researchers planning to test devices on human subjects]  
*(include Appendix B, the [Device Form](#))*  
 Materials of Human Origin  
*(include Appendix C, the [Biological Materials Form](#))*  
 Drugs and Biologics  
*(include Appendix D, the [Drug and Chemical Usage Form](#))*



MRI AT BIC To use the [Beckman Institute Biomedical Imaging Center \(BIC\)](#) in human subject's research, you must obtain prior approval from the BIC (217.244.0446; [ryambert@illinois.edu](mailto:ryambert@illinois.edu)) and use BIC-approved screening and consent forms. Attach:

BIC approval    BIC screening form    BIC consent form

## Section 16. CONFIDENTIALITY OF DATA & PRIVACY OF PARTICIPATION

**16A. How is participant data, records, or specimens identified when received or collected by researchers? Identifiers include, but are not limited to, name, date of birth, email address, street address, phone number, audio or video recordings, and SSN.**

- No identifiers are collected  
 Direct identifiers are collected  
 Indirect identifiers (e.g. a code or pseudonym used to track participants);  
Does the research team have access to the identity key?  Yes  No

**16B. Select all methods used to safeguard research records during storage:**

- Written consent, assent, or parental permission forms are stored separately from the data  
 Data is collected or given to research team without identifiers  
 Data is recorded by research team without identifiers  
 Direct identifiers are removed from collected data as soon as possible  
 Direct identifiers are deleted and no identity key exists as soon as possible  
 Participant codes or pseudonyms are used on all data and the existing identity key is stored separately from the data  
 Electronic data is stored in a secure, [UIUC-approved location](#), please specify Beckman Institute Dell/EMC Isilon storage system that is fully encrypted  
 Hard-copy data is stored in a secure location on UIUC's campus, please specify  
 Other, please specify:

**16C. How long will identifiable data be kept?** In perpetuity according to the terms of the industry consortium agreement. The identity key linking participant data to their identity will be destroyed no less than three years after the date of the participant's last speech recording, by deleting the corresponding entry in the participant identity database.

**16D. Describe provisions to protect the privacy interests of subjects.** Separation of identifiable information from speech samples in the database. Participant's recorded speech samples are stored on a virtual machine that contains no other information contributed by the participant; on this virtual machine, the participant's identity is coded using a participant code that consists of a randomly generated sequence of letters and numbers. Participant contact information, and any other information contributed by the participant, are stored in a separate database, on a separate virtual machine, accessible only by University of Illinois employees with administrative access to the server. Both virtual machines are implemented on the Beckman Institute Dell EMC Isilon/PowerScale encrypted storage array. The Isilon system currently has storage to accommodate approximately 1 PB of data, with over 300 TB of open space available. We can



add additional storage, up to a maximum of 64 PB, by adding additional nodes to the array. The primary system is in a secure server room in the Beckman Institute building, and the disaster recovery (DR) system is in a secure data center on the Illinois State University campus. Both systems are fully encrypted and connected on the University of Illinois WAN. The data is synchronized from the production site to the DR site 3 times a day. Snapshots of the data are taken 3 times a day, and each is retained for 4 weeks. In addition, a monthly snapshot is taken and retained for 12 months. Participant speech data, and text transcripts created by University of Illinois annotators, will be distributed to coalition partners approximately once per month during the period of data collection and annotation. Each data distribution will consist of a single archive file, compressed, and encrypted using a secure key technology such as RSA or DSA, so that it can only be decompressed by individual researchers who have been authorized by the coalition partner company under the terms of the data use agreement. Data distributions to coalition partners will never contain any information contributed by a participant other than their recorded speech samples.

**16E. Describe the training and experience of all persons who will collect or have access to the data.** The team members are up to date on CITI modules and Research IT trainings on data privacy and security. The IT system developers are all professional full-time employees of the Beckman Institute's IT services department.

## Section 17. DISSEMINATION OF RESULTS

**17A. List proposed forms of dissemination (e.g. journal articles, thesis, academic paper, conference presentation, sharing within industry, etc.).**

The de-identified data repository is accessible only to institutions who sign a data use agreement with UIUC. In addition, we expect this project to lead to the usual forms of research dissemination (journal articles, conference presentations and papers).

**17B. Will any identifiers be published, shared, or otherwise disseminated?**  Yes  No  
**If yes, does the consent form explicitly ask consent for such dissemination, or otherwise inform participants that it is required in order to participate in the study?**  Yes, optionally, participants may also consent to allow the use of their voice recordings in research settings such as professional conference presentations or journal publications. This is completely optional and not required for participation.

**17C. Do you intend to put de-identified data in a data repository?**  Yes  No  
**If yes, explain how data will be de-identified.** We will assign a unique universal ID number to each participant. The Beckman Institute will host this private data repository and only share its contents with partner organizations who have signed the data use agreement.

## Section 18. INVESTIGATOR & DEPARTMENTAL ASSURANCES

- I certify that the information provided in this application is complete and correct.
- I certify that I will follow my IRB Approved Protocol.



Office for the Protection  
of Research Subjects

# Protocol Form

- I accept ultimate responsibility for the conduct of this study, the ethical performance of the project, and the protection of the rights and welfare of the human subjects who are directly or indirectly involved in this project.
- I will comply with all applicable federal, state and local laws regarding the protection of human subjects in research.
- I will ensure that the personnel performing this study are qualified and adhere to the provisions of this IRB-certified protocol.

2024 April 5

Principal Investigator

Date

If the PI is not eligible to serve as PI under the [Campus Administrative Manual](#), the applicable academic dean, institute director, or campus administrative officer indicates their approval of the researcher to act as Principal Investigator. Please note that departmental assurance only needs to be provided in the initial application.

Name of Authorizing Individual

Signature of Authorizing Individual

Date

## **NOVA H IRB PROTOCOL ATTACHMENT 2:**

This is EXHIBIT C from the MOU signed by the five founding partners of the coalition. "Organization" here refers to the University of Illinois Urbana-Champaign.

### **LICENSED SPEECH SIGNAL DATA DATA USE TERMS AND CONDITIONS**

The licenses to Licensed Speech Signal Data granted to each Member pursuant to Section 9.b. of the Agreement are subject to the following Terms and Conditions. Each Member receiving such a license is referred to herein as the "**Data User**". Any capitalized terms used but not otherwise defined in this Exhibit C shall have the meanings given to them elsewhere in the Agreement.

#### **1. Provision and Use of Data**

**1.1** Effective as of the Effective Date, Data User may make, have made, use, load, access, store, copy, reproduce, destroy, modify, transmit, display, make derivative works of and otherwise use the Data made available to it by Organization under the Agreement on a non-exclusive, non-assignable, non-sublicensable (except as provided in Section 2.1 below) basis solely for the Permitted Purpose subject to the terms of this Exhibit C and the Agreement. Any such modifications or derivative works made by Data User are "**Modification(s)**".

**1.2** This Agreement does not restrict Data User's use or modification of any portions of the Data that Organization makes publicly available under a more permissive license, if applicable, to the extent Data User uses or modifies the Data under the terms of such public license, or that otherwise become public.

**1.3** The rights and restrictions set forth in this Exhibit C as it relates to the use, modification, distribution, disclosure, and privacy of the Data may only be modified by a written agreement between Data User and Organization specifically indicating it is amending these Terms and Conditions, e.g., these Terms and Conditions may not be amended by or superseded by any general terms and conditions that are presented or included as part of access to or use of data storage, data processing, platform, or computing resources.

#### **2. Restrictions**

**2.1** Data User may not sell, rent, lease, lend, transfer or otherwise redistribute the Data or any Modification of the Data, in whole or in part, except that Data User may disclose or provide access to the Data (including any Modifications) to its employees, Affiliates, contractors, and consultants working on behalf of Data User for the Permitted Purpose (collectively "**Representatives**").

**2.2** This Agreement (including without limitation Section 2.1) does not impose any restriction with respect to the development, use, modification, or distribution of Results.

**2.3** All costs and expenses incurred by Data User to exercise its rights and comply with its obligations under this Exhibit C shall be paid by Data User, and Data User shall not be entitled to reimbursement from Organization for such costs and expenses.

### **3. Confidentiality of Data, Privacy**

**3.1** Data User may not disclose or provide access to the Data (including any Modifications) to any third party, except that Data User may disclose or provide access to the Data (including any Modifications) to its Representatives, and may allow such Representatives to use and modify the Data (including any Modifications) solely to the extent needed for the Permitted Purpose, provided that such Representatives are bound by the terms and conditions of this Agreement (including Sections 3.2 and 3.3) or have entered into a written agreement with the Data User that (i) requires the Representative to process the Data on behalf of Data User and only in accordance with Data User's documented instruction, and (ii) contains data protection obligations not less protective than those imposed on Data User in this Agreement with respect to the protection of the Data processed by the Representative on behalf of Data User for the Permitted Purpose. Data User is responsible for its Representatives' compliance with the applicable terms of this Exhibit C.

**3.2** Data User agrees to implement and maintain reasonable physical, administrative, and technical safeguards to protect the Data from inadvertent or unauthorized access, disclosure, or use, and shall comply with applicable privacy laws relating to its possession and use of the Data or any Modification of the Data, including any biometric laws.

**3.3** Data User will not attempt to identify any natural person from any anonymized, pseudonymized, or otherwise de-identified personal data included in the Data.

### **4. Representations and Certifications**

**4.1** Organization represents and certifies that (i) it will provide Data composed solely of anonymized, pseudonymized, or otherwise de-identified data; (ii) it will provide Data to the Data User that does not violate any laws, rules, standards, or regulations governing the use of personal data, including but not limited to United States privacy laws and biometric laws; (iii) it will obtain all consents, waivers and clearances required to provide the Data to the Data User and for the Data User to process the Data for the Permitted Purpose, subject to the terms of this Exhibit C and the Agreement; and (iv) it will promptly delete any Data from the centralized repository following a data contributor's withdrawal of consent and/or request for data deletion and promptly notify Data User of the same.

**4.2 EXCEPT AS PROVIDED IN SECTION 4.1 ABOVE, THE DATA IS PROVIDED ON AN "AS IS" BASIS, WITHOUT WARRANTIES, CERTIFICATIONS, OR CONDITIONS OF ANY KIND, EITHER EXPRESS OR IMPLIED INCLUDING, WITHOUT LIMITATION, ANY WARRANTIES, CERTIFICATIONS, OR CONDITIONS OF TITLE, NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.**

## **5. Termination, Terms and Conditions, Notices**

**5.1** Organization may terminate Data User's license to the Data upon written notice if Data User does not follow the terms of this Exhibit C, including the limitations set forth in Section 2, and does not fully cure such failure within thirty (30) days of receiving written notice from Organization. In the event Data User's license is terminated under this Section 5.1, Data User shall cease use of the Data and any Modifications and use reasonable efforts to delete any copies of the Data, but shall not be restricted from continued use of any Results.

**5.2** Data User may terminate its license to the Data upon written notice to Organization. In the event Data User's license is terminated under this Section 5.2, Data User shall not receive access to additional Data following termination, but shall not be restricted from continued use of the Data to which Data User was previously granted access pursuant to this Exhibit C or to any Modifications.

**5.3** The terms and conditions of the Agreement (including but not limited to Sections 11, 13.d., and 13.f. of the Agreement) shall continue to apply with respect to Data User's exercise of its license rights pursuant to Section 9.b and this Exhibit C; provided, however, in the event of a conflict between the terms and conditions of this Exhibit C and the terms and conditions elsewhere in the Agreement, the terms and conditions of this Exhibit C shall prevail.

**5.4** Any notices provided under this Exhibit C shall be provided pursuant to Section 13.e. of the Agreement using the applicable party contact information listed in the Agreement.

## **6. Definitions**

**6.1 "Data"** means the Licensed Speech Signal Data (as defined in the Agreement) that Data User receives under the Agreement during the Term while such Data User is a party to the Agreement, in modified or unmodified form. Data does not include Results.

**6.2 "Permitted Purpose"** means research by or on behalf of Data User to advance understanding of machine learning-related data variances associated with atypical speech variations, and the development of products, technologies, and services by or on behalf of Data User to improve communications and increase access to technology, including computational use, processing of Biometric Data, and the creation and training of machine learning solutions and artificial intelligence models including for research and for commercial purposes.

**6.3 "Result"** means anything that Data User develops or improves for the Permitted Purpose that does not include more than a de minimis portion of the Data (e.g., summary statistics such as the number of speakers, and the gender, age range, etiology, ethnicity, and similar characteristics of speakers) on which the use is based. Results may include de minimis portions of the Data necessary to report on or explain use that has been conducted with the Data, such as figures in scientific papers, but do not include more. For example, machine learning and artificial intelligence models trained on Data (and which do not include more than a

de minimis portion of Data) are Results to the extent they are developed by or on behalf of Data User pursuant to the terms of the Agreement and this Exhibit C.

**6.4 “Biometric Data”** means data relating to an individual’s physical and/or behavioral characteristics that could be used for identification, including without limitation information relating to an individual’s DNA, fingerprints, face, hand, retina, gestures, voiceprint, typing rhythm, and gait.



IRB Number: 23183  
IRB Approval Date: 05/08/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

### **Speech Accessibility Project: Individuals with disabilities helping researchers to improve technology**

You are being asked to participate in a voluntary research study that is called the “Speech Accessibility Project.” The purpose of this study is to help researchers at universities and companies to develop spoken-language user interfaces that work for people with atypical speech. Software programs that understand speech are developed using machine learning. Machine learning is a software development method that imitates the way humans learn, by testing the software repeatedly, and changing its behavior whenever it makes a mistake. In this way, the software learns to reduce the risk of similar mistakes in the future. By providing examples of the sounds of atypical speech, this study will make it possible for the software to learn how to make fewer mistakes when it encounters atypical speech in the future. Participating in this study is all online and will involve recording sentences and uploading your audio recordings to a secure encrypted storage system. We ask you to provide up to 600 speech samples or about 5 hours of your time which you can do any time online, at your own pace and convenience. Risks associated with this study are no greater than those in everyday life. There is no expected benefit of this research to you personally; the expected benefit to society is that companies and organizations using your speech recordings will be able to train better speech technology for people with speech and motor disabilities, which will make technology more accessible for everyone.

**Project Name:** The Speech Accessibility Project, which has the website:  
<https://speechaccessibilityproject.beckman.illinois.edu/>

**Principal Investigator Name and Title:** Professor Mark Hasegawa-Johnson, University of Illinois Urbana-Champaign (“UIUC”), Beckman Institute and Department of Electrical and Computer Engineering.

**Contact Information:** [speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu) is the team email address.

**Organizations that will perform this study:** University of Illinois Urbana-Champaign (UIUC) is the lead organization that will perform this study and has partnered with the Davis Phinney Foundation and LSVT Global, two organizations that advocate for people with Parkinson's.

**Sponsor:** The Speech Accessibility Project is funded by a consortium that currently includes five companies: Amazon, Apple, Google, Microsoft, and Meta (parent company of Facebook). In the future, other companies, universities, and organizations may join this consortium as well.

### **Why am I being asked?**

You are being asked to participate in this study because you are an adult who is part of a community of people who may have atypical or non-standard speech, and you have already qualified to be in the study by successfully answering our pre-screening questions and providing a set of speech samples.

We plan to recruit more than 1,000 people to be in this study. Because the purpose of this study is to help organizations develop better speech interfaces for technologies, our team needs to gather many different types of speech samples from many types of voices.

Your participation in this research is voluntary. Your decision whether or not to participate will not affect your current or future dealings with the University of Illinois Urbana-Champaign, its partners, or its sponsors of this project. If you decide to participate, you are free to withdraw at any time without affecting that relationship.



IRB Number: 23183  
IRB Approval Date: 05/08/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

### What procedures are involved?

We will provide you with access to a website where you can record up to 600 speech samples by either reading the text prompt aloud, or by speaking the answer to a question. You can do these recordings at any time, 24/7, at whatever pace you wish. We encourage you to record at different times of day so that we can learn how your speech might change throughout the day. And don't worry about giving us "bad" samples – all data are really helpful. The more diverse our data set is, the better!

You have the option to work with our partner organization, LSVT Global, to get help. LSVT Global can provide you with a mentor to help you do your recordings, give you advice about your microphone settings, and give you tips for when and where to do your recordings.

You may take up to three months to finish recording the 600 speech samples. If you will need more than three months, please contact the research team at [speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu).

We will ask you whether you consider U.S. English to be your first language. People who speak English as a second language are sometimes better recognized by devices that also know multiple languages. If you provide this information, it will be shared with researchers using the data.

We will also ask you to provide optional data about yourself, such as your age, racial and ethnic identity, and questions about how your speech ability may have impacted your everyday life. You are not required to answer any of these questions. This information will only be used to confirm diverse representation in this data collection, and, if provided, will not be shared outside of the UIUC research team.

The UIUC research team will review all of your speech samples, and we will edit them to protect your privacy by removing any references to personally identifiable information about you or others that you might have said in your answers to our question prompts. We will also annotate your samples with "tags" that refer to technical aspects of the acoustic signals in your samples. We will not attach your name or email address to these samples; instead, we will assign a unique participant identification code to you and will annotate your samples with that code. Only the UIUC research team will know the "identity key" that relates your email address to your participant identification code.

The UIUC Speech Accessibility Project research team has established a secure, fully encrypted storage system at the UIUC Beckman Institute to store your samples. We will separately store your identifiable information in a separate secure database. We have also established a secure data sharing mechanism with the partners in our consortium. Right now, our partners are the five company sponsors of this project: Amazon, Apple, Google, Meta, and Microsoft. UIUC researchers and researchers at each of the partners will use the collected speech samples to train machine learning algorithms to perform better on speech recognition tasks. These partners will only have access to your speech samples and to no other data about you.

The five corporate sponsors of this research at UIUC have signed a legal agreement with the University that sets forth the permitted purposes of the speech data collected from this research project. These permitted purposes are only for research and technology development as stated here:

- Your samples will be kept forever, to assist in research and technology development in perpetuity.
- Your samples are not identified with your name or email address, but only with a unique participant code that is only known to the UIUC research team.



## University of Illinois Urbana-Champaign Online Consent Form

- Your samples will be analyzed by the UIUC research team to verify that the study is being done properly and to provide advice to other researchers about how to design or improve future studies.
- Your samples will be analyzed and combined with other people's samples to create aggregate results that will be published in the research literature.
- Your samples will be used to advance understanding of machine learning-related data variances associated with atypical speech variations, and eventually to develop products, technologies, and services by or on behalf of our sponsors to improve communications and increase access to technology, including the creation and training of machine learning solutions and artificial intelligence models for both research and commercial purposes.

In the future, any additional partner organizations in this consortium will sign data use agreements with the same terms and conditions. For these future uses of your data, we will not ask you to consent again. In other words, the future use of your de-identified data could happen without additional informed consent.

In summary, your speech samples will help researchers at UIUC and its partner organizations to develop better speech recognition technology algorithms and may eventually help commercial products perform more effectively for people with atypical speech.

### What are the potential risks and discomforts?

In general, we believe that the risks of your participation in this research study are no more than you would encounter in everyday life. However, we do recognize two issues that may be risks in this study, and we also have designed measures to minimize these risks.

First, you may get tired. To minimize your tiredness, we have arranged to let you record your speech samples over multiple sessions. You can log in to the system at a time and location convenient to you, and record at your own pace. You can also log out at any time, then log in again to continue recording as you wish. We hope that this flexibility and convenience will minimize any fatigue that you may experience.

Second, as is true for almost any online data collection research study, there is a risk that your speech samples could be released for purposes and to other individuals, companies, and organizations beyond the purposes to which you have given consent. Please know that the UIUC team has designed measures to minimize the possibility of unauthorized data release include the following:

- Participant code numbers will be used to annotate all speech recordings; your name and email address are not attached to any speech samples directly.
- All organizations using your speech recordings are required to sign a data use agreement, in which they agree that they will store the data securely, that they will prevent unauthorized access, and that they will not attempt to identify you or obtain any other identifying information about you.
- Researchers from outside UIUC will not know your name, e-mail address, state of residence, or any other identifying information about you. Identifying information will be stored separately from your speech recordings, in a database accessible only to UIUC research team members. In addition, as stated above, the UIUC team will edit and annotate your speech samples to protect your privacy.
- You are allowed to withdraw your voluntary consent at any time and request that all of your data be destroyed. When you notify the UIUC team that you wish to have all of your speech samples destroyed, UIUC will ensure that your speech recordings are removed from the UIUC secure system



IRB Number: 23183  
IRB Approval Date: 05/08/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

and will direct our partners that you wish your data to be removed from any copies they have as well.

### **Are there benefits to participating in the research?**

There is no immediate benefit of this research to you personally.

Information obtained from you in this research may help in the development of scientific research, open-source products, and commercial products by UIUC, its research partners, or members of the consortium. There are no plans to provide financial compensation to you should this occur.

The expected benefit to society is that companies and organizations will be able to develop better spoken-language user interfaces so that people with atypical speech can be as easily understood as other people with typical speech.

### **What other options are there?**

You have the option to not participate in this study. This study is entirely voluntary.

### **Will my study-related information be kept confidential?**

The UIUC research team has made every effort to protect the audio records of your speech samples. The UIUC team will edit out identifying information from your speech samples (such as your name) and will only store your audio files with a participant code number. This participant code number is stored with your identifiable information on a secure server at the UIUC Beckman Institute, and this "identity key" information is only accessible by the UIUC research team. This information is necessary in order to pay you for your participation in this project.

As described previously, the UIUC team will also listen to all of your speech samples and edit out any personally identifiable information.

Faculty, staff, students, and others with permission or authority to see your study information (such as University auditors) will maintain its confidentiality to the extent permitted and required by laws and university policies.

Speech recordings will be distributed only to companies and organizations that have signed a data use agreement, agreeing to confidentiality terms specified above. No other personally identifiable information will be published or presented.

Because this is a research project, we do expect to publish research papers as a result of this study. These papers may be in the public domain. Research papers will never state your name or any identifying information about you. Optionally, you may choose to allow the university to share anonymized samples of your voice as part of research publications or presentations, but this is not required.

### **Will I be reimbursed for any expenses or paid for my participation in this research?**

You will be compensated up to \$180 as follows: \$60 after recording the 200<sup>th</sup> sample, an additional \$60 after recording the 400<sup>th</sup> sample, and an additional \$60 after recording the 600<sup>th</sup> sample.



## University of Illinois Urbana-Champaign Online Consent Form



IRB Number: 23183  
IRB Approval Date: 05/08/2023  
IRB Expiration Date: 03/02/2028

In addition to your own payment, if you require the help of a caregiver to use a web browser, please contact the UIUC research team at [speechaccessibilityproject@beckman.illinois.edu](mailto:speechaccessibilityproject@beckman.illinois.edu) to validate that your caregiver may also receive compensation for helping you. Named caregivers will be compensated with up to \$90 as follows: \$30 after recording the 200<sup>th</sup> sentence, an additional \$30 after recording the 400<sup>th</sup> sentence, and an additional \$30 after recording the 600<sup>th</sup> sentence.

Payments will be made through email using Amazon ecodes.

### **Can I withdraw or be removed from the study?**

Yes. You are free to withdraw your consent and discontinue participation at any time by notifying the UIUC research team by email or phone. The UIUC research team also have the right to stop your participation in this study without your consent if they believe it is in your best interests. If you withdraw consent, the UIUC research team will remove your speech recordings from the UIUC database and will notify the sponsors and partners that they should also remove your data from their own databases using only your anonymized participant code number. If you withdraw consent, your data will not be used for any future research. If the data you previously provided has been used to inform or create computer code or machine learning algorithms it may not be possible for your data to be removed from them.

### **Will data collected from me be used for any other research?**

As stated above, UIUC and its partners in the Speech Accessibility Project consortium will use your speech samples to advance understanding of machine learning-related data variances associated with atypical speech variations, and to develop products, technologies, and services to improve communications and increase access to technology, including computational use, and the creation and training of machine learning solutions and artificial intelligence models including for research and for commercial purposes.

It is possible that, in the future, other organizations will be granted access to the data, in addition to the founding consortium partners that are Amazon, Apple, Google, Meta, and Microsoft. These other organizations may be other universities or other companies. All organizations seeking access to the data will be required to sign data use agreements that specify the same terms as above: that your speech samples will only be used for research and technology development, that your data will be stored securely, and that you may withdraw consent at any time. You may not be asked to provide additional consent for these future uses of your data.

### **Who should I contact if I have questions?**

Contact the UIUC research team leader, Professor Mark Hasegawa-Johnson ([jhasegaw@illinois.edu](mailto:jhasegaw@illinois.edu)), if you have any questions about this study or your part in it, or if you have concerns or complaints about the research.

### **What are my rights as a research subject?**

If you have any questions about your rights as a research subject, including concerns, complaints, or to offer input, you may call the UIUC Office for the Protection of Research Subjects (OPRS) at 217-333-2670 or e-mail OPRS at [irb@illinois.edu](mailto:irb@illinois.edu). If you would like to complete a brief survey to provide OPRS feedback about your experiences as a research participant, please follow the link [here](#) or through a link on the OPRS website: <https://oprs.research.illinois.edu/>. You will have the option to provide feedback or concerns anonymously or you may provide your name and contact information for follow-up purposes.

**Please print this consent form if you would like to retain a copy for your records.**



IRB Number: 23183  
IRB Approval Date: 05/08/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

**Please read each item carefully and check each box to indicate your agreement.**

- I have read and understand the above consent form.
- I understand that my speech samples are collected for the purpose of improving speech recognition technology and are protected by a legal agreement between UIUC and the current members of the Speech Accessibility Project coalition, and that future members of the data coalition and any future data licensees will also be bound by similar legal agreements.
- I certify that I am 18 years old or older.
- I understand that I must maintain a valid email address in order to participate in this study and to be paid for my participation.
- I certify that I and I alone will use my login to record speech samples. I will not allow other people to record speech samples on my behalf.
- I understand that the only financial compensation that I will receive for participating in this study is the payment for contributing speech samples, which will be \$60 for every batch of 200 samples I provide.
- I understand that my goal is to provide 600 samples total, and that therefore my maximum payment will be \$180.
- I understand that if I need the help of a caregiver to contribute speech samples, that I must notify the UIUC team and have my caregiver approved to be paid in addition to myself.
- I understand that at any time, I can end my participation and withdraw my consent. When I do so, the UIUC research team will remove all of my speech samples from the UIUC secure system. I also understand that it is UIUC's responsibility to ask all sponsors and partners to remove any copies of my speech samples from their own databases as well.

**Optional checkbox:** By checking the box below, you may also specifically grant your consent for the research team to play or provide audio files of your voice in a public research setting or publication.

- I consent to have my voice samples shared in public at research conferences and similar professional meetings, and to have digital audio files of my voice be shared as part of professional research publications.

**Optional checkbox:** By checking the box below, you may also specifically grant your consent to share your contact information with other research teams, so that they may contact you to be a potential participant in their own research studies.



IRB Number: 23183  
IRB Approval Date: 05/08/2023  
IRB Expiration Date: 03/02/2028

University of Illinois Urbana-Champaign  
Online Consent Form

*I consent to have my email address and phone number be shared with other research teams, who may contact me to participate in their studies.*

By clicking the “I CONSENT” button below, I indicate my willingness to voluntarily take part in this study.

**I CONSENT**

Speech Accessibility Project  
Documentation of Assent

You are being asked to join a voluntary research study called the “Speech Accessibility Project.” This study will help computers learn to understand people with different types of speech. A computer learns to understand speech in almost the same way humans learn, that is, by listening to speech, guessing what it says, and then guessing again whenever it makes a mistake. If the computer tries many times, then after a while it learns to guess correctly. By giving computers some examples of various types of speech, we can help them to better understand different types of speech in the future.

Participating in this study is all online. You will record your speech online. We ask you to record up to 600 sentences. It should take you about 5 hours. You don't have to do it all at once; you can stop and start at any time, at your own pace and whenever you want. Risks associated with this study are no greater than those in everyday life. You will not get any personal benefit from this research. Society may benefit, because computers might learn how to better understand all types of speech.

**Project Name:** The Speech Accessibility Project, which has the website:  
<https://speechaccessibilityproject.beckman.illinois.edu/>

**Principal Investigator Name and Title:** Professor Mark Hasegawa-Johnson, University of Illinois Urbana-Champaign (“UIUC”), Beckman Institute and Department of Electrical and Computer Engineering.

**Contact Information:** [speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu) is the team email address.

**Organizations that will perform this study:** University of Illinois Urbana-Champaign (UIUC) is the lead organization that will perform this study and has partnered with the Davis Phinney Foundation and LSVT Global, two organizations that advocate for people with Parkinson's.

**Sponsor:** The Speech Accessibility Project is funded by a consortium that currently includes five companies: Amazon, Apple, Google, Microsoft, and Meta (parent company of Facebook). In the future, other companies, universities, and organizations may join this consortium as well.

### **Why am I being asked?**

You are being asked to participate in this study because you are an adult, and because we think that computers should learn to understand speech like yours.

We plan to recruit more than 1,000 people to be in this study. Because the purpose of this study is to help computers learn to understand lots of different types of speech, we need to gather many different types of speech from many types of people.

Your participation in this research is voluntary. Your decision whether or not to participate will not affect your current or future dealings with the University of Illinois, its partners, or its sponsors of this project. Even if you decide to participate, you can withdraw later at any time without affecting that relationship.

### **What will you do for this research?**

We will give you access to a website where you can record up to 600 speech recordings. For each recording, you will either read a sentence, or repeat a sentence that somebody reads to you, or answer a question out loud. You can do these recordings any time, whenever you wish.

We will also give you access to a mentor. Your mentor will help you do your recordings, give you advice about your microphone settings, and give you tips for when and where to do your recordings.



You may take up to three months to finish your recordings. If you will need more than three months, please tell us so by e-mail at [speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu).

We will ask you whether you consider U.S. English to be your first language. People who speak English as a second language are sometimes better recognized by computers that also know multiple languages. If you provide this information, it will be shared with researchers using the data.

We will also ask you to provide optional data about yourself, such as your age, race and ethnicity. We will also ask questions about how your speech ability affects your everyday life. You are not required to answer any of these questions. If you tell us these things, it will help us to know when we have collected enough different types of speech. Other people who use your speech samples will not know your answers to these questions.

Right now, the people who will get copies of your speech are the five company sponsors of this project: Amazon, Apple, Google, Meta, and Microsoft. Those five companies signed a contract with the University saying that they will only use your recordings to help computers better understand speech. Some of the terms of the contract include:

- Your recordings will be kept forever, so that computers in the future can also learn to understand speech like yours.
- Your recordings will be deleted if you tell us to delete them.
- Your recordings are not identified with your name or email address, but only with an ID code. Only UIUC will know your name or your email address.
- UIUC will listen to your recordings before we give them to anybody else.
- Scientists might study your speech, and write papers about it, but they will not publish your speech or your identity.
- Your speech might be used to develop new computer programs, including both free programs and non-free programs that help computers to understand people with speech like yours.

In the future we might let other people use your recordings. We will not ask for your permission every time we give the speech to somebody new, as long as they agree to all of the things listed above. In other words, you agree that other people can use your speech as long as they agree to the things listed above.

In summary, your speech samples will help researchers at UIUC and other places to make computers that can understand speech like yours.

### **What are the potential risks and discomforts?**

The risks are no more than you would encounter in everyday life. We can think of two things that you might not like about this study, so here is how we will try to avoid these risks:

First, you may get tired. To minimize your tiredness, we have arranged to let you record your speech samples over multiple sessions. You can log in at any time, from anywhere, and record at your own pace. You can also log out at any time, then log in again later. We hope that this flexibility will help you to avoid getting tired.

Second, even though we are working hard to keep your data safe, there is a risk that somebody might steal your data from us. In order to keep that from happening, here is what we have done:

- We will not attach your name or email address to your speech. Instead, we will assign an ID code



## Speech Accessibility Project Documentation of Assent



to you, and we will label your samples with that code. Only the UIUC research team will know the “identity key” that relates your email address to your ID code.

- We will listen to all of your speech. If you say your own name or the names of anybody close to you, we will cut those sections out of your speech recording to protect your privacy.
- We will store your speech in a secure computer system at UIUC. We will store your name and e-mail address on a different computer, also secure.
- Before anybody from outside UIUC can use your speech, they need to sign a contract. In the contract, they promise to store your speech securely. They agree that only approved people can use your speech. They agree that they won't try to find out who you are. They also agree that, if we tell them to destroy all of your speech recordings, they will.
- Nobody from outside UIUC will know your name, your e-mail address, or what state you live in. That information will only be stored at UIUC, on a computer that only UIUC team members can use.
- You can change your mind at any time. If you change your mind, and want us to destroy your speech recordings, we will.

### **Are there benefits to participating in the research?**

There is no benefit to you personally.

Your recordings might help companies to make computer software that they can sell. You will not receive any extra payment if that happens.

The whole world may benefit if computers learn how to better understand people with speech like yours.

### **What other options are there?**

You have the option to not participate in this study. This study is entirely voluntary.

### **Will my study-related information be kept confidential?**

Only people from UIUC will know your e-mail address or your name. We only need to know those things so that we can pay you for your participation in this project.

As described previously, the UIUC team will also listen to all of your speech samples and edit out any personal names.

Faculty, staff, students, and other people from UIUC will follow the law and UIUC policies in order to protect your confidentiality.

Speech recordings will be given only to people who have signed a contract, as described above.

Because this is a research project, we do expect to publish research papers as a result of this study.

Research papers will never give your name or identity. If you wish, you may choose to allow the university to share anonymized samples of your voice in research presentations, but you don't need to.

### **Will I be reimbursed for any expenses or paid for my participation in this research?**

You will be paid up to \$180. You will be paid \$60 after recording your 200<sup>th</sup> sample. You will be paid another \$60 after recording your 400<sup>th</sup> sample, and you will get a final \$60 after recording your 600<sup>th</sup> sample.

**Speech Accessibility Project  
Documentation of Assent**

In addition to your own payment, if somebody helps you to use your web browser for this project, they can be paid up to \$90. They will be paid \$30 after you record your 200<sup>th</sup> sentence, an additional \$30 after you record your 400<sup>th</sup> sentence, and a final \$30 after you record your 600<sup>th</sup> sentence.

Payments will be made through email using Amazon ecodes.

**Can I withdraw or be removed from the study?**

Yes. You can withdraw at any time, and you can tell us to delete your recordings. We also have the right to stop your participation if we think it is in your best interests. If you want us to delete your recordings, you only need to tell us that you want them deleted. If you tell us to delete your recordings, we will delete them. We will also tell your participant ID number to anybody else who has a copy of your recordings, and say that they need to delete your recordings. Nobody will ever use your recordings after that for any future research. If some computers have already learned from your recordings, it may not be possible to force the computer to unlearn what it has learned.

**Will data collected from me be used for any other research?**

UIUC, Amazon, Apple, Google, Meta and Microsoft will use your speech recordings to make computers that better understand speech like yours. In the future, we may also let other people use your recordings, if they are willing to sign a data use agreement similar to the data use agreement that Amazon, Apple, Google, Meta, and Microsoft have already signed. As long as new users of the data agree to the same terms, we will assume that you have already given permission.

**Who should I contact if I have questions?**

Contact the UIUC research team leader, Professor Mark Hasegawa-Johnson ([jhasegaw@illinois.edu](mailto:jhasegaw@illinois.edu)), if you have any questions about this study or your part in it. You can also contact him if you have any worries or complaints about the research.

**What are my rights as a research subject?**

If you have any questions about your rights as a research subject, including worries, complaints, or ideas, you may call the UIUC Office for the Protection of Research Subjects (OPRS) at 217-333-2670 or e-mail OPRS at [irb@illinois.edu](mailto:irb@illinois.edu).

If you would like to complete a brief survey to provide OPRS feedback about your experiences as a research participant, please follow the link [here](#) or through a link on the OPRS website:

<https://oprs.research.illinois.edu/>. You will be able to provide feedback or concerns without giving your name, or you can give us your name and contact information.

**Please print this consent form if you would like a copy for your records.**

**Please confirm, verbally, the participant's assent to each of the following points. If the participant gives assent to each point, please check the corresponding box.**

- I understand the assent form as it has been described to me.
- I understand that my speech samples will be used to help computers better understand speech like mine.

Speech Accessibility Project  
Documentation of Assent

- I understand that my speech samples are protected by a contract between UIUC and the current members of the Speech Accessibility Project. I understand that future users of the data will have to sign a similar contract before they can use the data.
- I am 18 years old or older.
- I won't let other people record their speech in place of me.
- The only payment I will get is the payment for contributing speech samples, which will be \$60 for every batch of 200 samples I provide.
- My goal is to provide 600 samples total. My maximum payment will be \$180.
- I can stop participating at any time. Any time I want to, I can tell UIUC to delete all of my recordings. If I tell UIUC to delete my recordings, they will delete them. They will also tell anybody else who has copies to delete them.

**Optional:** If the participant assents to let us play his or her speech as part of a presentation or an online demo of the Speech Accessibility Project, they can agree to this. If they agree verbally, please check this box. This is not required.

- I agree to let my speech be played during public presentations about the Speech Accessibility Project or as part of an online demo of the project.

**Optional:** If the participant assents to let us give their name and your e-mail address or phone number to people who are planning other research studies, so that they can be a participant in other studies too, then please check the following box. This is not required.

- I agree to that my email address and phone number can be shared with other research teams, who may contact me to participate in their studies.

By clicking the “I ASSENT” button below, I, the participant's legal guardian, indicate that the participant has expressed positive assent to voluntarily take part in this study.

**I ASSENT**

## Assent to Participate in a Research Study

### Who are we and what are we doing?

*We are from the University of Illinois Urbana-Champaign. We would like to ask if you would be in a research study. A research study is a way to find out new information about something. This is the way we try to find out how to help computers learn to understand people with different types of speech.*

### Why are we asking you to be in this research study?

*We are asking you to be in this research study because we want to learn more about how you speak. We will use this information to help computers learn to recognize speech from different types of people like you. We want you to be in this study because we think that computers should learn to understand speech like yours.*

### What happens in the research study?

*If you decide to be in this research study and your legally authorized representative agrees, this is what will happen. We will give you and your caregiver access to a website where you can record up to 600 speech recordings. For each recording, you will be asked to either read a sentence or repeat a sentence that the caregiver says, or you will be asked a question. Your caregiver will help you learn what to do. You will also have a mentor who will help you learn how to make the recordings. You and your caregiver will do these recordings together at any time, whenever you want. You can take up to three months to finish your recordings. If you need more time, please let us know.*

### Will any part of the research study hurt you?

*There is a chance that during this research study you could get tired. To help with that, you can record your speech at your own pace, whenever is best for you. There is also a chance that somebody could steal your speech recordings from us. To keep that from happening, we will not put your name on your speech recordings, and we will follow the rules to keep your information safe.*

### Will the research study help you or anyone else?

*We do not know for sure if being in this research study will help you. It is possible that we could learn something to help other people like you some day.*

### Who will see the information about you?

*Only the researchers or others who are doing their jobs will be able to see the information about you from this research study. We will not tell anyone else that you are in the study.*

### What if you have any questions about the research study?

*It is okay to ask questions. If you don't understand something, you can ask us. We want you to ask questions now and anytime you think of them. If you have a question later that you didn't think of now, you can contact Professor Hasegawa-Johnson ([jhasegaw@illinois.edu](mailto:jhasegaw@illinois.edu)) or ask us the next time we see you.*

**Do you have to be in the research study?**

*You do not have to be in this study if you don't want to. Being in this study is up to you. No one will be upset if you don't want to do it. Even if you say yes now, you can change your mind later and tell us you want to stop. You can take your time to decide. You can talk to your caregiver before you decide. We will also ask your legally authorized representative to give their permission for you to be in this study. But even if your legally authorized representative says "yes" you can still decide not to be in the research study.*

**Agreeing to be in the study**

*I was able to ask questions about this study. Signing my name at the bottom means that I agree to be in this study. My legally authorized representative and I will be given a copy of this form after I have signed it.*

---

Printed Name

---

Sign your name on this line

---

Date

---

Printed Name of Person Obtaining Assent

---

Signature of Person Obtaining Assent

---

Date

The following should be completed by the study member conducting the assent process if the participant agrees to be in the study. Initial the appropriate selection:

The participant is capable of reading the assent form and has signed above as documentation of assent to take part in this study.

The participant is not capable of reading the assent form, but the information was verbally explained to him/her. The participant signed above as documentation of assent to take part in this study.



IRB Number: 23183  
IRB Approval Date: 11/01/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

### **Speech Accessibility Project: Individuals with disabilities helping researchers to improve technology**

#### **Consent Form for Recorded Caregivers**

You are being asked to participate in a voluntary research study that is called the “Speech Accessibility Project.” The purpose of this study is to help researchers at universities and companies to develop spoken-language user interfaces that work for people with atypical speech. Software programs that understand speech are developed using machine learning. Machine learning is a software development method that imitates the way humans learn, by testing the software repeatedly, and changing its behavior whenever it makes a mistake. In this way, the software learns to reduce the risk of similar mistakes in the future. By providing examples of the sounds of atypical speech, this study will make it possible for the software to learn how to make fewer mistakes when it encounters atypical speech in the future. Participating in this study is all online and will involve helping your friend or family member with a disability to record sentences and to upload the audio recordings to a secure encrypted storage system. We ask you to help your friend or family member with a disability to provide up to 600 speech samples or about 5 hours of your time which you can do any time online, at your own pace and convenience. Risks associated with this study are no greater than those in everyday life. There is no expected benefit of this research to you personally; the expected benefit to society is that companies and organizations using your speech recordings will be able to train better speech technology for people with speech and motor disabilities, which will make technology more accessible for everyone.

**Project Name:** The Speech Accessibility Project, which has the website:  
<https://speechaccessibilityproject.beckman.illinois.edu/>

**Principal Investigator Name and Title:** Professor Mark Hasegawa-Johnson, University of Illinois Urbana-Champaign (“UIUC”), Beckman Institute and Department of Electrical and Computer Engineering.

**Contact Information:** [speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu) is the team email address.

**Organizations that will perform this study:** University of Illinois Urbana-Champaign (UIUC) is the lead organization that will perform this study.

**Sponsor:** The Speech Accessibility Project is funded by a consortium that currently includes five companies: Amazon, Apple, Google, Microsoft, and Meta (parent company of Facebook). In the future, other companies, universities, and organizations may join this consortium as well.

#### **Why am I being asked?**

You are being asked to participate in this study because you are an adult who assists a person with a disability to interact with computers, and that person has already qualified to be in the study by successfully answering our pre-screening questions and providing a set of speech samples, and that person either is not able to read prompts from the screen, or does not wish to do so.

We plan to recruit more than 1,000 people to be in this study. Because the purpose of this study is to help organizations develop better speech interfaces for technologies, our team needs to gather many different types of speech samples from many types of voices.

Your participation in this research is voluntary. Your decision whether or not to participate will not affect your current or future dealings with the University of Illinois Urbana-Champaign, its partners, or its



IRB Number: 23183  
IRB Approval Date: 11/01/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

sponsors of this project. If you decide to participate, you are free to withdraw at any time without affecting that relationship.

### What procedures are involved?

We will provide you with access to a website where you and the person that you are assisting can record up to 600 speech samples. Some of the prompts will instruct you to read a sentence out loud so that the person you're assisting can repeat it out loud. Some of the prompts will instruct you to ask the person you're assisting a question, so that they can speak their answer to the question. You and the person that you are assisting can do these recordings at any time, 24/7, at whatever pace you wish. We encourage you to record at different times of day so that we can learn how your speech might change throughout the day. And don't worry about giving us "bad" samples – all data are really helpful. The more diverse our data set is, the better!

You and the person that you are assisting have the option to work with UIUC, to get help. UIUC can provide you with a mentor to help you do your recordings, give you advice about your microphone settings, and give you tips for when and where to do your recordings.

You may take up to three months to finish recording the 600 speech samples. If you will need more than three months, please contact the research team at [speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu).

We will ask the person you're assisting whether they consider U.S. English to be their first language. People who speak English as a second language are sometimes better recognized by devices that also know multiple languages. If you provide this information, it will be shared with researchers using the data.

We will also ask the person you're assisting to provide optional data about themselves, such as age, racial and ethnic identity, and questions about how their speech ability may have impacted their everyday life. They are not required to answer any of these questions. This information will only be used to confirm diverse representation in this data collection, and, if provided, will not be shared outside of the UIUC research team.

The UIUC research team will review all of your speech samples, and we will edit them to protect your privacy by removing any references to personally identifiable information about you or others that you might have said in your answers to our question prompts. We will also annotate your samples with "tags" that refer to technical aspects of the acoustic signals in your samples. We will not attach your name or email address to these samples; instead, we will assign a unique participant identification code to you and will annotate your samples with that code. Only the UIUC research team will know the "identity key" that relates your email address to your participant identification code.

The UIUC Speech Accessibility Project research team has established a secure, fully encrypted storage system at the UIUC Beckman Institute to store your samples. We will separately store your identifiable information in a separate secure database. We have also established a secure data sharing mechanism with the partners in our consortium. Right now, our partners are the five company sponsors of this project: Amazon, Apple, Google, Meta, and Microsoft. UIUC researchers and researchers at each of the partners will use the collected speech samples to train machine learning algorithms to perform better on speech recognition tasks. These partners will only have access to your speech samples and to no other data about you.



## University of Illinois Urbana-Champaign Online Consent Form



IRB Number: 23183  
IRB Approval Date: 11/01/2023  
IRB Expiration Date: 03/02/2028

The five corporate sponsors of this research at UIUC have signed a legal agreement with the University that sets forth the permitted purposes of the speech data collected from this research project. These permitted purposes are only for research and technology development as stated here:

- Your samples will be kept forever, to assist in research and technology development in perpetuity.
- Your samples are not identified with your name or email address, but only with a unique participant code that is only known to the UIUC research team.
- Your samples will be analyzed by the UIUC research team to verify that the study is being done properly and to provide advice to other researchers about how to design or improve future studies.
- Your samples will be analyzed and combined with other people's samples to create aggregate results that will be published in the research literature.
- Your samples will be used to advance understanding of machine learning-related data variances associated with atypical speech variations, and eventually to develop products, technologies, and services by or on behalf of our sponsors to improve communications and increase access to technology, including the creation and training of machine learning solutions and artificial intelligence models for both research and commercial purposes.

In the future, any additional partner organizations in this consortium will sign data use agreements with the same terms and conditions. For these future uses of your data, we will not ask you to consent again. In other words, the future use of your de-identified data could happen without additional informed consent.

In summary, your speech samples will help researchers at UIUC and its partner organizations to develop better speech recognition technology algorithms and may eventually help commercial products perform more effectively for people with atypical speech.

### **What are the potential risks and discomforts?**

In general, we believe that the risks of your participation in this research study are no more than you would encounter in everyday life. However, we do recognize two issues that may be risks in this study, and we also have designed measures to minimize these risks.

First, you may get tired. To minimize your tiredness, we have arranged to let you record your speech samples over multiple sessions. You can log in to the system at a time and location convenient to you, and record at your own pace. You can also log out at any time, then log in again to continue recording as you wish. We hope that this flexibility and convenience will minimize any fatigue that you may experience.

Second, as is true for almost any online data collection research study, there is a risk that your speech samples could be released for purposes and to other individuals, companies, and organizations beyond the purposes to which you have given consent. Please know that the UIUC team has designed measures to minimize the possibility of unauthorized data release include the following:

- Participant code numbers will be used to annotate all speech recordings; your name and email address are not attached to any speech samples directly.
- All organizations using your speech recordings are required to sign a data use agreement, in which they agree that they will store the data securely, that they will prevent unauthorized access, and that they will not attempt to identify you or obtain any other identifying information about you.



## University of Illinois Urbana-Champaign Online Consent Form



IRB Number: 23183  
IRB Approval Date: 11/01/2023  
IRB Expiration Date: 03/02/2028

- Researchers from outside UIUC will not know your name, e-mail address, state of residence, or any other identifying information about you. Identifying information will be stored separately from your speech recordings, in a database accessible only to UIUC research team members. In addition, as stated above, the UIUC team will edit and annotate your speech samples to protect your privacy.
- You are allowed to withdraw your voluntary consent at any time and request that all of your data be destroyed. When you notify the UIUC team that you wish to have all of your speech samples destroyed, UIUC will ensure that your speech recordings are removed from the UIUC secure system and will direct our partners that you wish your data to be removed from any copies they have as well.

### **Are there benefits to participating in the research?**

There is no immediate benefit of this research to you personally.

Information obtained from you in this research may help in the development of commercial products by UIUC, its research partners, or members of the consortium. There are no plans to provide financial compensation to you should this occur.

The expected benefit to society is that companies and organizations will be able to develop better spoken-language user interfaces so that people with atypical speech can be as easily understood as other people with typical speech.

### **What other options are there?**

You have the option to not participate in this study. This study is entirely voluntary.

### **Will my study-related information be kept confidential?**

The UIUC research team has made every effort to protect the audio records of your speech samples. The UIUC team will edit out identifying information from your speech samples (such as your name) and will only store your audio files with a participant code number. This participant code number is stored with your identifiable information on a secure server at the UIUC Beckman Institute, and this "identity key" information is only accessible by the UIUC research team. This information is necessary in order to pay you for your participation in this project.

As described previously, the UIUC team will also listen to all of your speech samples and edit out any personally identifiable information.

Faculty, staff, students, and others with permission or authority to see your study information (such as University auditors) will maintain its confidentiality to the extent permitted and required by laws and university policies.

Speech recordings will be distributed only to companies and organizations that have signed a data use agreement, agreeing to confidentiality terms specified above. No other personally identifiable information will be published or presented.

Because this is a research project, we do expect to publish research papers as a result of this study. These papers may be in the public domain. Research papers will never state your name or any identifying



IRB Number: 23183  
IRB Approval Date: 11/01/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

information about you. Optionally, you may choose to allow the university to share anonymized samples of your voice as part of research publications or presentations, but this is not required.

### **Will I be reimbursed for any expenses or paid for my participation in this research?**

The person that you are assisting will be compensated up to \$180 as follows: \$60 after recording the 200<sup>th</sup> sample, an additional \$60 after recording the 400<sup>th</sup> sample, and an additional \$60 after recording the 600<sup>th</sup> sample.

In addition, since you are helping them to record their speech samples, you will also receive compensation for helping them. As a named caregiver you will be compensated with up to \$90 as follows: \$30 after they record the 200<sup>th</sup> sentence, an additional \$30 after they record the 400<sup>th</sup> sentence, and an additional \$30 after they record the 600<sup>th</sup> sentence.

Payments will be made through email using Amazon ecodes.

### **Can I withdraw or be removed from the study?**

Yes. You are free to withdraw your consent and discontinue participation at any time by notifying the UIUC research team by email or phone. The UIUC research team also have the right to stop your participation in this study without your consent if they believe it is in your best interests. If you withdraw consent, the UIUC research team will remove your speech recordings from the UIUC database and will notify the sponsors and partners that they should also remove your data from their own databases using only your anonymized participant code number. If you withdraw consent, your data will not be used for any future research. If the data you previously provided has been used to inform or create computer code or machine learning algorithms it may not be possible for your data to be removed from them.

### **Will data collected from me be used for any other research?**

As stated above, UIUC and its partners in the Speech Accessibility Project consortium will use your speech samples to advance understanding of machine learning-related data variances associated with atypical speech variations, and to develop products, technologies, and services to improve communications and increase access to technology, including computational use, and the creation and training of machine learning solutions and artificial intelligence models including for research and for commercial purposes.

It is possible that, in the future, other organizations will be granted access to the data, in addition to the founding consortium partners that are Amazon, Apple, Google, Meta, and Microsoft. These other organizations may be other universities or other companies. All organizations seeking access to the data will be required to sign data use agreements that specify the same terms as above: that your speech samples will only be used for research and technology development, that your data will be stored securely, and that you may withdraw consent at any time. You may not be asked to provide additional consent for these future uses of your data.

### **Who should I contact if I have questions?**

Contact the UIUC research team leader, Professor Mark Hasegawa-Johnson ([jhasegaw@illinois.edu](mailto:jhasegaw@illinois.edu)), if you have any questions about this study or your part in it, or if you have concerns or complaints about the research.

### **What are my rights as a research subject?**



IRB Number: 23183  
IRB Approval Date: 11/01/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

If you have any questions about your rights as a research subject, including concerns, complaints, or to offer input, you may call the UIUC Office for the Protection of Research Subjects (OPRS) at 217-333-2670 or e-mail OPRS at [irb@illinois.edu](mailto:irb@illinois.edu). If you would like to complete a brief survey to provide OPRS feedback about your experiences as a research participant, please follow the link [here](#) or through a link on the OPRS website: <https://oprs.research.illinois.edu/>. You will have the option to provide feedback or concerns anonymously or you may provide your name and contact information for follow-up purposes.

**Please print this consent form if you would like to retain a copy for your records.**

**Please read each item carefully and check each box to indicate your agreement.**

- I have read and understand the above consent form.
- I understand that my speech samples are collected for the purpose of improving speech recognition technology and are protected by a legal agreement between UIUC and the current members of the Speech Accessibility Project coalition, and that future members of the data coalition and any future data licensees will also be bound by similar legal agreements.
- I certify that I am 18 years old or older.
- I understand that I must maintain a valid email address in order to participate in this study and to be paid for my participation.
- I certify that only I and the person that I am assisting will use our login to record speech samples. I will not allow other people to record speech samples on our behalf.
- I understand that the only financial compensation that I will receive for participating in this study is the payment for contributing speech samples, which will be \$30 for every batch of 200 samples that the person I am assisting provides.
- I understand that my goal is for the person I am assisting to provide 600 samples total, and that therefore my maximum payment will be \$90.
- I understand that at any time, I can end my participation and withdraw my consent. When I do so, the UIUC research team will remove all of my speech samples from the UIUC secure system. I also understand that it is UIUC's responsibility to ask all sponsors and partners to remove any copies of my speech samples from their own databases as well.

**Optional checkbox:** By checking the box below, you may also specifically grant your consent for the research team to play or provide audio files of your voice in a public research setting or publication.



IRB Number: 23183  
IRB Approval Date: 11/01/2023  
IRB Expiration Date: 03/02/2028

University of Illinois Urbana-Champaign  
Online Consent Form

*I consent to have my voice samples shared in public at research conferences and similar professional meetings, and to have digital audio files of my voice be shared as part of professional research publications.*

**Optional checkbox:** By checking the box below, you may also specifically grant your consent to share your contact information with other research teams, so that they may contact you to be a potential participant in their own research studies.

*I consent to have my email address and phone number be shared with other research teams, who may contact me to participate in their studies.*

By clicking the “I CONSENT” button below, I indicate my willingness to voluntarily take part in this study.

**I CONSENT**



IRB Number: 23183  
IRB Approval Date: 10/17/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

### **Speech Accessibility Project: Individuals with disabilities helping researchers to improve technology**

#### **Consent Form for the Contribution of Previously Recorded Speech**

You are being asked to participate in a voluntary research study that is called the “Speech Accessibility Project.” The purpose of this study is to help researchers at universities and companies to develop spoken-language user interfaces that work for people with atypical speech. Software programs that understand speech are developed using machine learning. Machine learning is a software development method that imitates the way humans learn, by testing the software repeatedly, and changing its behavior whenever it makes a mistake. In this way, the software learns to reduce the risk of similar mistakes in the future. By providing examples of the sounds of atypical speech, this study will make it possible for the software to learn how to make fewer mistakes when it encounters atypical speech in the future. Participating in this study is all online and will involve uploading speech samples that you have previously recorded to a secure encrypted storage system. Risks associated with this study are no greater than those in everyday life. There is no expected benefit of this research to you personally; the expected benefit to society is that companies and organizations using your speech recordings will be able to train better speech technology for people with speech and motor disabilities, which will make technology more accessible for everyone.

**Project Name:** The Speech Accessibility Project, which has the website:  
<https://speechaccessibilityproject.beckman.illinois.edu/>

**Principal Investigator Name and Title:** Professor Mark Hasegawa-Johnson, University of Illinois Urbana-Champaign (“UIUC”), Beckman Institute and Department of Electrical and Computer Engineering.

**Contact Information:** [speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu) is the team email address.

**Organizations that will perform this study:** University of Illinois Urbana-Champaign (UIUC) is the lead organization that will perform this study.

**Sponsor:** The Speech Accessibility Project is funded by a consortium that currently includes five companies: Amazon, Apple, Google, Microsoft, and Meta (parent company of Facebook). In the future, other companies, universities, and organizations may join this consortium as well.

#### **Why am I being asked?**

You are being asked to participate in this study because you are an adult who is part of a community of people who may have atypical or non-standard speech, and you have already qualified to be in the study by successfully answering our pre-screening questions and providing a set of speech samples.

We plan to recruit more than 1,000 people to be in this study. Because the purpose of this study is to help organizations develop better speech interfaces for technologies, our team needs to gather many different types of speech samples from many types of voices.

Your participation in this research is voluntary. Your decision whether or not to participate will not affect your current or future dealings with the University of Illinois Urbana-Champaign, its partners, or its sponsors of this project. If you decide to participate, you are free to withdraw at any time without affecting that relationship.



## University of Illinois Urbana-Champaign Online Consent Form



IRB Number: 23183  
IRB Approval Date: 10/17/2023  
IRB Expiration Date: 03/02/2028

### What procedures are involved?

We will provide you with access to a website where you can upload speech samples that you have already recorded. You can upload your recordings at any time. All data are really helpful. The more diverse our data set is, the better!

You have the option to work with one of our partner organizations to get help, or to ask for help through our tech support e-mail. Our partner organizations can provide you with a mentor to help you upload your recordings.

We will ask you whether you consider U.S. English to be your first language. People who speak English as a second language are sometimes better recognized by devices that also know multiple languages. If you provide this information, it will be shared with researchers using the data.

We will also ask you to provide optional data about yourself, such as your age, racial and ethnic identity, and questions about how your speech ability may have impacted your everyday life. You are not required to answer any of these questions. This information will only be used to confirm diverse representation in this data collection, and, if provided, will not be shared outside of the UIUC research team.

The UIUC research team will review all of your speech samples, and we will edit them to protect your privacy by removing any references to personally identifiable information about you or others that you might have said. We will not attach your name or email address to these samples; instead, we will assign a unique participant identification code to you and will annotate your samples with that code. Only the UIUC research team will know the “identity key” that relates your email address to your participant identification code.

The UIUC Speech Accessibility Project research team has established a secure, fully encrypted storage system at the UIUC Beckman Institute to store your samples. We will separately store your identifiable information in a separate secure database. We have also established a secure data sharing mechanism with the partners in our consortium. Right now, our partners are the five company sponsors of this project: Amazon, Apple, Google, Meta, and Microsoft. UIUC researchers and researchers at each of the partners will use the collected speech samples to train machine learning algorithms to perform better on speech recognition tasks. These partners will only have access to your speech samples and to no other data about you.

The five corporate sponsors of this research at UIUC have signed a legal agreement with the University that sets forth the permitted purposes of the speech data collected from this research project. These permitted purposes are only for research and technology development as stated here:

- Your samples will be kept forever, to assist in research and technology development in perpetuity.
- Your samples are not identified with your name or email address, but only with a unique participant code that is only known to the UIUC research team.
- Your samples will be analyzed by the UIUC research team to verify that the study is being done properly and to provide advice to other researchers about how to design or improve future studies.
- Your samples will be analyzed and combined with other people's samples to create aggregate results that will be published in the research literature.



## University of Illinois Urbana-Champaign Online Consent Form



IRB Number: 23183  
IRB Approval Date: 10/17/2023  
IRB Expiration Date: 03/02/2028

- Your samples will be used to advance understanding of machine learning-related data variances associated with atypical speech variations, and eventually to develop products, technologies, and services by or on behalf of our sponsors to improve communications and increase access to technology, including the creation and training of machine learning solutions and artificial intelligence models for both research and commercial purposes.

In the future, any additional partner organizations in this consortium will sign data use agreements with the same terms and conditions. For these future uses of your data, we will not ask you to consent again. In other words, the future use of your de-identified data could happen without additional informed consent.

In summary, your speech samples will help researchers at UIUC and its partner organizations to develop better speech recognition technology algorithms and may eventually help commercial products perform more effectively for people with atypical speech.

### What are the potential risks and discomforts?

In general, we believe that the risks of your participation in this research study are no more than you would encounter in everyday life. However, we do recognize one issue that may be a risk in this study, and we also have designed measures to minimize this risk.

As is true for almost any online data collection research study, there is a risk that your speech samples could be released for purposes and to other individuals, companies, and organizations beyond the purposes to which you have given consent. Please know that the UIUC team has designed measures to minimize the possibility of unauthorized data release include the following:

- Participant code numbers will be used to annotate all speech recordings; your name and email address are not attached to any speech samples directly.
- All organizations using your speech recordings are required to sign a data use agreement, in which they agree that they will store the data securely, that they will prevent unauthorized access, and that they will not attempt to identify you or obtain any other identifying information about you.
- Researchers from outside UIUC will not know your name, e-mail address, state of residence, or any other identifying information about you. Identifying information will be stored separately from your speech recordings, in a database accessible only to UIUC research team members. In addition, as stated above, the UIUC team will edit and annotate your speech samples to protect your privacy.
- You are allowed to withdraw your voluntary consent at any time and request that all of your data be destroyed. When you notify the UIUC team that you wish to have all of your speech samples destroyed, UIUC will ensure that your speech recordings are removed from the UIUC secure system and will direct our partners that you wish your data to be removed from any copies they have as well.

### Are there benefits to participating in the research?

There is no immediate benefit of this research to you personally.

Information obtained from you in this research may help in the development of scientific research, open-source products and commercial products by UIUC, its research partners, or members of the consortium. There are no plans to provide financial compensation to you should this occur.



IRB Number: 23183  
IRB Approval Date: 10/17/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

The expected benefit to society is that companies and organizations will be able to develop better spoken-language user interfaces so that people with atypical speech can be as easily understood as other people with typical speech.

### **What other options are there?**

You have the option to not participate in this study. This study is entirely voluntary.

### **Will my study-related information be kept confidential?**

The UIUC research team has made every effort to protect the audio records of your speech samples. The UIUC team will edit out identifying information from your speech samples (such as your name) and will only store your audio files with a participant code number. This participant code number is stored with your identifiable information on a secure server at the UIUC Beckman Institute, and this "identity key" information is only accessible by the UIUC research team. This information is necessary in order to pay you for your participation in this project.

As described previously, the UIUC team will also listen to all of your speech samples and edit out any personally identifiable information.

Faculty, staff, students, and others with permission or authority to see your study information (such as University auditors) will maintain its confidentiality to the extent permitted and required by laws and university policies.

Speech recordings will be distributed only to companies and organizations that have signed a data use agreement, agreeing to confidentiality terms specified above. No other personally identifiable information will be published or presented.

Because this is a research project, we do expect to publish research papers as a result of this study. These papers may be in the public domain. Research papers will never state your name or any identifying information about you. Optionally, you may choose to allow the university to share anonymized samples of your voice as part of research publications or presentations, but this is not required.

### **Will I be reimbursed for any expenses or paid for my participation in this research?**

You will be compensated up to \$30 if the data that you upload includes at least 100 sentences of speech, and the associated text.

Payments will be made through email using Amazon ecodes.

### **Can I withdraw or be removed from the study?**

Yes. You are free to withdraw your consent and discontinue participation at any time by notifying the UIUC research team by email or phone. The UIUC research team also have the right to stop your participation in this study without your consent if they believe it is in your best interests. If you withdraw consent, the UIUC research team will remove your speech recordings from the UIUC database and will notify the sponsors and partners that they should also remove your data from their own databases using only your anonymized participant code number. If you withdraw consent, your data will not be used for any future research. If the



## University of Illinois Urbana-Champaign Online Consent Form



IRB Number: 23183  
IRB Approval Date: 10/17/2023  
IRB Expiration Date: 03/02/2028

data you previously provided has been used to inform or create computer code or machine learning algorithms it may not be possible for your data to be removed from them.

### **Will data collected from me be used for any other research?**

As stated above, UIUC and its partners in the Speech Accessibility Project consortium will use your speech samples to advance understanding of machine learning-related data variances associated with atypical speech variations, and to develop products, technologies, and services to improve communications and increase access to technology, including computational use, and the creation and training of machine learning solutions and artificial intelligence models including for research and for commercial purposes.

It is possible that, in the future, other organizations will be granted access to the data, in addition to the founding consortium partners that are Amazon, Apple, Google, Meta, and Microsoft. These other organizations may be other universities or other companies. All organizations seeking access to the data will be required to sign data use agreements that specify the same terms as above: that your speech samples will only be used for research and technology development, that your data will be stored securely, and that you may withdraw consent at any time. You may not be asked to provide additional consent for these future uses of your data.

### **Who should I contact if I have questions?**

Contact the UIUC research team leader, Professor Mark Hasegawa-Johnson ([jhasegaw@illinois.edu](mailto:jhasegaw@illinois.edu)), if you have any questions about this study or your part in it, or if you have concerns or complaints about the research.

### **What are my rights as a research subject?**

If you have any questions about your rights as a research subject, including concerns, complaints, or to offer input, you may call the UIUC Office for the Protection of Research Subjects (OPRS) at 217-333-2670 or e-mail OPRS at [irb@illinois.edu](mailto:irb@illinois.edu). If you would like to complete a brief survey to provide OPRS feedback about your experiences as a research participant, please follow the link [here](#) or through a link on the OPRS website: <https://oprs.research.illinois.edu/>. You will have the option to provide feedback or concerns anonymously or you may provide your name and contact information for follow-up purposes.

**Please print this consent form if you would like to retain a copy for your records.**

**Please read each item carefully and check each box to indicate your agreement.**

- I have read and understand the above consent form.
- I understand that my speech samples are collected for the purpose of improving speech recognition technology and are protected by a legal agreement between UIUC and the current members of the Speech Accessibility Project coalition, and that future members of the data coalition and any future data licensees will also be bound by similar legal agreements.
- I certify that I am 18 years old or older.



IRB Number: 23183  
IRB Approval Date: 10/17/2023  
IRB Expiration Date: 03/02/2028

University of Illinois Urbana-Champaign  
Online Consent Form

- I understand that I must maintain a valid email address in order to participate in this study and to be paid for my participation.
- I certify that the recordings I upload will be recordings of me and me alone. I will not upload recordings of any other people.
- I understand that the only financial compensation that I will receive for participating in this study is the payment for contributing speech samples, which will be \$30 if I upload at least 100 samples.
- I understand that at any time, I can end my participation and withdraw my consent. When I do so, the UIUC research team will remove all of my speech samples from the UIUC secure system. I also understand that it is UIUC's responsibility to ask all sponsors and partners to remove any copies of my speech samples from their own databases as well.

**Optional checkbox:** By checking the box below, you may also specifically grant your consent for the research team to play or provide audio files of your voice in a public research setting or publication.

- I consent to have my voice samples shared in public at research conferences and similar professional meetings, and to have digital audio files of my voice be shared as part of professional research publications.

**Optional checkbox:** By checking the box below, you may also specifically grant your consent to share your contact information with other research teams, so that they may contact you to be a potential participant in their own research studies.

- I consent to have my email address and phone number be shared with other research teams, who may contact me to participate in their studies.

By clicking the "I CONSENT" button below, I indicate my willingness to voluntarily take part in this study.

I CONSENT

## Assent to Participate in a Research Study

### Who are we and what are we doing?

*We are from the University of Illinois Urbana-Champaign. We would like to ask if you would be in a research study. A research study is a way to find out new information about something. This is the way we try to find out how to help computers learn to understand people with different types of speech.*

### Why are we asking you to be in this research study?

*We are asking you to be in this research study because we want to learn more about how you speak. We will use this information to help computers learn to recognize speech from different types of people like you. We want you to be in this study because we think that computers should learn to understand speech like yours.*

### What happens in the research study?

*If you decide to be in this research study and your legally authorized representative agrees, this is what will happen. We will give you and your caregiver access to a website where you can record up to 600 speech recordings. For each recording, you will be asked to either read a sentence or repeat a sentence that the caregiver says, or you will be asked a question. Your caregiver will help you learn what to do. You will also have a mentor who will help you learn how to make the recordings. You and your caregiver will do these recordings together at any time, whenever you want. You can take up to three months to finish your recordings. If you need more time, please let us know.*

### Will any part of the research study hurt you?

*There is a chance that during this research study you could get tired. To help with that, you can record your speech at your own pace, whenever is best for you. There is also a chance that somebody could steal your speech recordings from us. To keep that from happening, we will not put your name on your speech recordings, and we will follow the rules to keep your information safe.*

### Will the research study help you or anyone else?

*We do not know for sure if being in this research study will help you. It is possible that we could learn something to help other people like you some day.*

### Who will see the information about you?

*Only the researchers or others who are doing their jobs will be able to see the information about you from this research study. We will not tell anyone else that you are in the study.*

### What if you have any questions about the research study?

*It is okay to ask questions. If you don't understand something, you can ask us. We want you to ask questions now and anytime you think of them. If you have a question later that you didn't think of now, you can contact Professor Hasegawa-Johnson ([jhasegaw@illinois.edu](mailto:jhasegaw@illinois.edu)) or ask us the next time we see you.*

**Do you have to be in the research study?**

*You do not have to be in this study if you don't want to. Being in this study is up to you. No one will be upset if you don't want to do it. Even if you say yes now, you can change your mind later and tell us you want to stop. You can take your time to decide. You can talk to your caregiver before you decide. We will also ask your legally authorized representative to give their permission for you to be in this study. But even if your legally authorized representative says "yes" you can still decide not to be in the research study.*

**Agreeing to be in the study**

*I was able to ask questions about this study. Signing my name at the bottom means that I agree to be in this study. My legally authorized representative and I will be given a copy of this form after I have signed it.*

---

Printed Name

---

Sign your name on this line

---

Date

---

Printed Name of Person Obtaining Assent

---

Signature of Person Obtaining Assent

---

Date

The following should be completed by the study member conducting the assent process if the participant agrees to be in the study. Initial the appropriate selection:

The participant is capable of reading the assent form and has signed above as documentation of assent to take part in this study.

The participant is not capable of reading the assent form, but the information was verbally explained to him/her. The participant signed above as documentation of assent to take part in this study.



Número de IRB: 23183

Fecha de aprobación de IRB: 05/08/2023

Fecha de vencimiento de IRB: 03/02/2028

## Universidad de Illinois Urbana-Champaign Formulario de consentimiento en línea

### Proyecto de accesibilidad del habla: personas con discapacidad que ayudan a los investigadores a mejorar la tecnología

Se le pide que participe de manera voluntaria en un estudio de investigación llamado "Proyecto de accesibilidad del habla". El objetivo de este estudio es ayudar a los investigadores de universidades y empresas a desarrollar interfaces de usuario en lenguaje hablado para que funcionen de manera correcta para personas con habla atípica. Los programas informáticos (de computadoras) que comprenden el habla se desarrollan mediante aprendizaje automático (del inglés machine learning). El aprendizaje automático es un método de desarrollo de software que imita la forma en que aprenden los humanos, probando el software repetidamente y cambiando su comportamiento cada vez que comete un error. De este modo, el software aprende a reducir el riesgo de cometer errores similares en el futuro. Al proporcionar ejemplos de sonidos de habla atípica, este estudio hará posible que el software aprenda a cometer menos errores cuando se encuentre con habla atípica en el futuro. La participación en este estudio es totalmente en línea (del inglés online) y consistirá en grabar frases y subir sus grabaciones de audio a un sistema de almacenamiento cifrado seguro. Le pedimos que proporcione hasta 600 muestras de habla o unas 5 horas de su tiempo, y lo puede hacer en cualquier momento en línea, a su propio ritmo y conveniencia. Los riesgos asociados a este estudio no son mayores que los de la vida cotidiana. No hay expectativa de beneficio para usted de esta investigación; el beneficio esperado para la sociedad es que las empresas y organizaciones que utilicen sus grabaciones del habla podrán entrenar una mejor tecnología del habla para las personas con discapacidades motoras y del habla, lo que hará que la tecnología sea más accesible para todos.

**Nombre del proyecto:** Proyecto de accesibilidad del habla, que tiene el sitio web:  
<https://speechaccessibilityproject.beckman.illinois.edu/>

**Nombre y cargo del investigador principal:** El Profesor Mark Hasegawa-Johnson, de la Universidad de Illinois Urbana- Champaign ("UIUC"), Instituto Beckman y facultad de Ingeniería Eléctrica e Informática.

**Datos de contacto:** La dirección de correo electrónico del equipo es:  
[speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu).

**Organizaciones que realizarán este estudio:** La Universidad de Illinois Urbana-Champaign (UIUC) es la organización principal que llevará a cabo este estudio y se ha asociado con la Fundación Davis Phinney y LSVT Global, dos organizaciones que abogan por las personas con Parkinson.

**Patrocinador:** El Proyecto de accesibilidad del habla está financiado por un consorcio que actualmente incluye a cinco empresas: Amazon, Apple, Google, Microsoft y Meta (empresa matriz de Facebook). En el futuro, otras empresas, universidades y organizaciones podrían unirse también a este consorcio.

#### ¿Por qué se me pidió participar?

Se le pide que participe en este estudio porque es un adulto que forma parte de una comunidad de personas que pueden tener un habla atípica, y ha cumplido los requisitos para participar en el estudio al responder satisfactoriamente a nuestras preguntas de preselección y proporcionar un conjunto de muestras de habla.

Tenemos previsto reclutar a más de 1.000 personas para participar en este estudio. Dado que el objetivo de este estudio es ayudar a estas empresas a desarrollar mejores interfaces de habla para las tecnologías, nuestro equipo necesita reunir muchos tipos diferentes de muestras de habla de muchos tipos de voces.

Su participación en esta investigación es voluntaria. Su decisión de participar no afectará a sus relaciones



Número de IRB: 23183

Fecha de aprobación de IRB: 05/08/2023

Fecha de vencimiento de IRB: 03/02/2028

## Universidad de Illinois Urbana-Champaign Formulario de consentimiento en línea

actuales o futuras con la Universidad de Illinois Urbana-Champaign, sus socios o los patrocinadores de este proyecto. Si decide participar, es libre de retirarse en cualquier momento sin que ello afecte a esa relación.

### ¿Qué procedimientos hay que seguir?

Le proporcionaremos acceso a un sitio web en el que podrá grabar hasta 600 muestras de voz, bien sea leyendo el texto en voz alta o respondiendo a una pregunta. Puede hacer estas grabaciones en cualquier momento del día a través de la semana y al ritmo que desee. Le animamos a que grabe a distintas horas del día para que podamos saber cómo cambia su habla a lo largo del día. Y no se preocupe por darnos muestras "malas": todos los datos son realmente útiles. Cuanto más diverso sea nuestro conjunto de muestras de habla, ¡mejor!

Tiene la opción de trabajar con nuestra organización asociada, LSVT Global, para obtener ayuda. LSVT Global puede proporcionarle un mentor que le ayude a hacer sus grabaciones, le aconseje sobre la configuración de su micrófono y le dé consejos sobre cuándo y dónde hacer sus grabaciones.

Puede tardar hasta tres meses en terminar de grabar las 600 muestras de voz. Si va a necesitar más de tres meses, póngase en contacto con el equipo de investigación de la UIUC a través de [speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu).

Le preguntaremos si considera que el inglés de EE. UU. es su primer idioma. Las personas que hablan inglés como segundo idioma a veces son mejor reconocidas por los dispositivos que también conocen varios idiomas. Si proporciona esta información, se compartirá con los investigadores que utilicen las muestras.

También le pediremos que proporcione datos opcionales sobre usted, como su edad, identidad racial y étnica, y preguntas sobre cómo puede haber influido su capacidad de habla en su vida cotidiana. No está obligado a responder a ninguna de estas preguntas. Esta información solo se utilizará para confirmar la representación diversa en esta recopilación de datos y, si se facilita, no se compartirá fuera del equipo de investigación de la UIUC.

El equipo de investigación de la UIUC revisará todas sus muestras de habla y las editaremos para proteger su privacidad eliminando cualquier referencia a información personal identifiable sobre usted o sobre otras personas que pueda haber mencionado en sus respuestas a nuestras preguntas. También anotaremos sus muestras con "etiquetas" que se refieren a aspectos técnicos de las señales acústicas de sus muestras. No adjuntaremos su nombre ni su dirección de correo electrónico a estas muestras; en su lugar, le asignaremos un código único de identificación de participante y anotaremos sus muestras con ese código. Solo el equipo de investigación de la UIUC conocerá la "clave de identidad" que relaciona su dirección de correo electrónico con su código de identificación de participante.

El equipo de investigación del Proyecto de accesibilidad del habla de la UIUC ha establecido un sistema de almacenamiento seguro y totalmente cifrado o encriptado en el Instituto Beckman de la UIUC para guardar sus muestras. Almacenaremos por separado su información identifiable en una base de datos segura e independiente. También hemos establecido un mecanismo seguro de intercambio de datos con los socios de nuestro consorcio. En este momento, nuestros socios son las cinco empresas patrocinadoras de este proyecto: Amazon, Apple, Google, Meta y Microsoft. Los



## Universidad de Illinois Urbana-Champaign Formulario de consentimiento en línea

investigadores de la UIUC y los de cada uno de las empresas utilizarán las muestras de voz recogidas para capacitar a los algoritmos de aprendizaje automático para que funcionen mejor en las tareas de reconocimiento de voz. Estos colaboradores de investigación solo tendrán acceso a sus muestras de habla y a ningún otro dato sobre usted.

Las cinco empresas patrocinadoras de esta investigación en la UIUC han firmado un acuerdo legal con la universidad que establece los fines permitidos de las muestras de voz recogidas en este proyecto de investigación. Estos fines permitidos son solo para la investigación y el desarrollo tecnológico, como se indica aquí:

- Sus muestras se conservarán para siempre, para ayudar a la investigación y al desarrollo tecnológico en perpetuidad.
- Sus muestras no se identificarán con su nombre o dirección de correo electrónico, sino solo con un código único de participante que solo conoce el equipo de investigación de la UIUC.
- Sus muestras serán analizadas por el equipo de investigación de la UIUC para verificar que el estudio se realiza correctamente y para asesorar a otros investigadores sobre cómo diseñar o mejorar futuros estudios.
- Sus muestras se analizarán y combinará con las muestras de otras personas para crear resultados agregados que se publicarán en la literatura de investigación.
- Sus muestras se utilizarán para mejorar la comprensión de las variaciones de los datos relacionadas con el aprendizaje automático que están asociadas a las variaciones atípicas del habla. Eventualmente conducirá al desarrollo de productos, tecnologías y servicios por o en nombre de nuestros patrocinadores para mejorar las comunicaciones y aumentar el acceso a la tecnología, incluyendo la creación y capacitación de soluciones de aprendizaje automático y modelos de inteligencia artificial tanto para fines de investigación como comerciales.

En el futuro, cualquier organización asociada adicional de este consorcio firmará acuerdos de uso de datos con los mismos términos y condiciones. Para tales usos futuros de sus datos, no pediremos su consentimiento nuevamente. En otras palabras, el uso futuro de sus datos des-identificados podría producirse sin consentimiento informado adicional.

En resumen, sus muestras de habla ayudarán a los investigadores de la UIUC y de sus organizaciones asociadas a desarrollar mejores algoritmos de tecnología de reconocimiento del habla y, con el tiempo, podrán ayudar a que los productos comerciales funcionen con mayor eficacia para las personas con habla atípica.

### ¿Cuáles son los posibles riesgos y molestias?

En general, creemos que los riesgos de su participación en este estudio de investigación no son mayores de los que encontraría en la vida cotidiana. Sin embargo, reconocemos dos asuntos que pueden suponer riesgos en este estudio, y también hemos diseñado medidas para minimizarlos.

En primer lugar, podría cansarse. Para minimizar su cansancio, hemos dispuesto que pueda grabar sus muestras de habla en varias sesiones. Puede conectarse al sistema a la hora y en el lugar que le convenga y grabar a su propio ritmo. También puede cerrar la sesión en cualquier momento y volver a conectarse para seguir grabando como deseé. Esperamos que esta flexibilidad y conveniencia reduzcan al mínimo la fatiga que pueda experimentar.

En segundo lugar, como ocurre en casi todos los estudios de investigación de recopilación de datos en línea,



## Universidad de Illinois Urbana-Champaign Formulario de consentimiento en línea

existe el riesgo que sus muestras de habla se divulguen con otros fines y a personas, empresas y organizaciones más allá de los que ha dado su consentimiento. Debe saber que el equipo de la UIUC ha diseñado medidas para minimizar la posibilidad de divulgación no autorizada de datos, entre las que se incluyen las siguientes:

- Los números de código de los participantes se utilizarán para hacer anotaciones en todas las grabaciones de habla; su nombre y dirección de correo electrónico no se adjuntarán directamente a ninguna muestra de habla.
- Todas las organizaciones que utilicen sus grabaciones de habla deberán firmar un acuerdo de uso de estos datos, en el que se comprometen a almacenar los datos de forma segura, a impedir el acceso no autorizado y a no intentar identificarle ni obtener ninguna otra información que le identifique.
- Los investigadores ajenos a la UIUC no conocerán su nombre, dirección de correo electrónico, estado de residencia ni ninguna otra información que le identifique. La información de identificación se almacenará separada de sus grabaciones del habla, en una base de datos accesible solo a los miembros del equipo de investigación de la UIUC. Además, como indicamos anteriormente, el equipo de la UIUC editaría y anotaría sus muestras de habla para proteger su privacidad.
- Puede retirar su consentimiento voluntario en cualquier momento y solicitar que se destruyan todos sus datos. Cuando notifique al equipo de la UIUC que desea que se destruyan todas sus muestras de habla, la UIUC se asegurará de que sus grabaciones de voz se eliminen del sistema seguro de la UIUC y les dará instrucciones a nuestros colaboradores que usted desea que sus datos sean borrados de todas las copias que posean también.

### ¿Hay beneficios por participar en esta investigación?

No existe un beneficio inmediato para usted por participar en esta investigación.

La información que obtendremos de usted en esta investigación puede ayudar en el desarrollo de investigación científica, productos de código abierto, y productos comerciales por parte de UIUC, sus socios de investigación o miembros del consorcio. No existen planes de proveerle compensación financiera adicional de esto ocurrir.

El beneficio esperado para la sociedad es que las empresas y organizaciones podrán desarrollar mejores interfaces que requieren el uso del lenguaje hablado para que las personas con un habla atípica puedan ser entendidas tan fácilmente como otras personas con habla típica.

### ¿Qué otras opciones existen?

Usted tiene la opción de no participar en este estudio. Este estudio es completamente voluntario.

### ¿Se mantendrá confidencial mi información relacionada con el estudio?

El equipo de investigación de UIUC ha hecho todo lo posible para proteger las audiorecordinaciones de sus



Número de IRB: 23183

Fecha de aprobación de IRB: 05/08/2023

Fecha de vencimiento de IRB: 03/02/2028



## Universidad de Illinois Urbana-Champaign Formulario de consentimiento en línea

muestras de habla. El equipo de UIUC editará la información de identificación de sus muestras de habla (como su nombre) y solo almacenará sus archivos de audio con un número de código de participante. Este número de código de participante se almacenará con su información de identificación en un servidor seguro en el Instituto UIUC Beckman, y esta "clave de identidad" es solo accesible por el equipo de investigación de UIUC. Esta información es necesaria para poder pagarle por su participación en este proyecto.

Como se describió anteriormente, el equipo de UIUC también escuchará todas sus muestras de habla y las editará para eliminar cualquier información de identificación personal.

Los miembros de facultad, personal, estudiantes y otras personas con permiso o autoridad para ver su información de este estudio (como auditores universitarios) mantendrán su confidencialidad en la medida permitida y requerida por las leyes y políticas universitarias.

Las grabaciones de habla se distribuirán únicamente a empresas y organizaciones que hayan firmado un acuerdo de uso de datos aceptando los términos de confidencialidad especificados anteriormente. Ninguna otra información de identificación personal será publicada o presentada.

Debido a que se trata de un proyecto de investigación, esperamos publicar artículos de investigación como resultado de este estudio. Estos artículos pueden ser de dominio público. Las publicaciones sobre esta investigación nunca incluirán su nombre ni ninguna información que le identifique. De manera opcional, puede decidir conceder permiso para que la universidad comparta muestras anónimas de su voz como parte de las publicaciones o presentaciones de investigación, pero esto no es obligatorio.

### **¿Seré reembolsado por gastos que incurra o pagado por mi participación en esta investigación?**

Usted será compensado hasta \$180 como sigue: \$60 luego de concluir la grabación de la frase número 200, \$60 adicionales luego de la grabación de la frase número 400, y \$60 adicionales luego la grabación de la frase número 600.

Además de su propio pago, si necesita la ayuda de un cuidador para utilizar un navegador web, comuníquese con el equipo de investigación de la UIUC a través de [speechaccessibilityproject@beckman.illinois.edu](mailto:speechaccessibilityproject@beckman.illinois.edu) para validar que su cuidador también pueda recibir una compensación por ayudarle. Los cuidadores aprobados recibirán una compensación de hasta \$90, de la siguiente manera \$30 tras usted grabar la frase número 200, otros \$30 tras usted grabar la frase número 400 y otros \$30 tras usted grabar la frase número 600.

Los pagos se realizarán por correo electrónico utilizando *eCodes* de Amazon.

### **¿Puedo retirarme o pueden terminar mi participación en el estudio?**

Sí. Usted es libre de retirar su consentimiento e interrumpir su participación en cualquier momento notificándolo al equipo de investigación de la UIUC por correo electrónico o por teléfono. El equipo de investigación de la UIUC también tiene derecho a interrumpir su participación en este estudio sin su



## Universidad de Illinois Urbana-Champaign Formulario de consentimiento en línea

consentimiento si considera que es lo mejor para usted. Si retira su consentimiento, el equipo de investigación de la UIUC eliminará sus grabaciones de habla de la base de datos de la UIUC y notificará a los patrocinadores y socios que también deben eliminar sus datos de sus propias bases de datos utilizando solo su número de código de participante anónimo. Si retira su consentimiento, sus datos no serán utilizados en futuras investigaciones. Si los datos que proporcionó anteriormente se han utilizado para informar o crear códigos informáticos o algoritmos de aprendizaje automático, puede que no sea posible eliminar sus datos de ellos.

### ¿Utilizarán mis datos para otras investigaciones?

Como se ha indicado anteriormente, la UIUC y sus socios en el consorcio del Proyecto de accesibilidad del habla utilizarán sus muestras de habla para avanzar la comprensión de las variaciones de los datos relacionados con el aprendizaje automático y asociados a variaciones atípicas del habla. Además serán utilizados para desarrollar productos, tecnologías y servicios que mejoren las comunicaciones y aumenten el acceso a la tecnología, incluido el uso computacional, y la creación y capacitación de soluciones de aprendizaje automático y modelos de inteligencia artificial, tanto con fines de investigación como comerciales.

Es posible que en el futuro se conceda acceso a los datos a otras organizaciones, además de los socios fundadores del consorcio que son Amazon, Apple, Google, Meta y Microsoft. Estas otras organizaciones pueden ser otras universidades u otras empresas. Todas las organizaciones que deseen acceder a los datos deberán firmar acuerdos de uso de datos que especifiquen los mismos términos que los anteriores: que sus muestras de habla solo se utilizarán para investigación y desarrollo tecnológico, que sus datos se almacenarán de forma segura y que podrá retirar su consentimiento en cualquier momento. No se le pedirá que de consentimiento adicional para estos usos futuros de sus datos.

### ¿Con quién debo ponerme en contacto si tengo preguntas?

Puede comunicarse con el jefe del equipo de investigación de la UIUC, el profesor Mark Hasegawa-Johnson ([jhasegaw@illinois.edu](mailto:jhasegaw@illinois.edu)), si tiene alguna pregunta sobre este estudio o sobre su participación en él, o si tiene alguna inquietud o queja sobre la investigación.

### ¿Cuáles son mis derechos como sujeto de investigación?

Si tiene alguna pregunta sobre sus derechos como sujeto de investigación, que incluyan inquietudes, quejas o para ofrecer su opinión, puede llamar a la Oficina para la Protección de Sujetos de Investigación (OPRS) de la UIUC al 217-333-2670 o enviar un correo electrónico a la OPRS a [irb@illinois.edu](mailto:irb@illinois.edu). Si quiere completar una breve encuesta para dar su opinión a la OPRS sobre sus experiencias como participante en una investigación, siga el enlace que aparece aquí o a través de un enlace en el sitio web de la OPRS: <https://oprs.research.illinois.edu/>. Tendrá la opción de dar su opinión o compartir sus inquietudes de forma anónima, o puede proveer su nombre e información de contacto para recibir seguimiento.

**Lee atentamente cada punto y marque cada casilla para indicar que está de acuerdo.**

- He leído y comprendo el formulario de consentimiento anterior.
- Entiendo que mis muestras de habla se recogen con el fin de mejorar la tecnología de reconocimiento del habla y están protegidas por un acuerdo legal entre la UIUC y los miembros actuales de la coalición del Proyecto de accesibilidad del habla, y que los futuros miembros de la coalición de datos y cualquier futuro licenciatario de datos también estarán sujetos a acuerdos legales similares.



Universidad de Illinois Urbana-Champaign  
Formulario de consentimiento en línea

- Certifico que tengo 18 años o más.
- Entiendo que debo mantener una dirección de correo electrónico válida para participar en este estudio y para que se me pague por mi participación.
- Certifico que yo y solo yo utilizaré mi nombre de usuario para grabar muestras de habla. No permitiré que otras personas graben muestras de habla en mi nombre.
- Entiendo que la única compensación económica que recibiré por participar en este estudio es el pago por contribuir con muestras de habla, que será de \$60 por cada lote de 200 muestras que proporcione.
- Entiendo que mi objetivo es proporcionar 600 muestras en total, y que por lo tanto mi pago máximo será de \$180 .
- Entiendo que si necesito la ayuda de un cuidador para contribuir mis muestras de habla, debo notificarlo al equipo de la UIUC para que aprueben el que se pague a mi cuidador además de a mí mismo.
- Entiendo que en cualquier momento puedo terminar mi participación y retirar mi consentimiento. Cuando lo haga, el equipo de investigación de la UIUC eliminará todas mis muestras de habla del sistema seguro de la UIUC. También entiendo que es responsabilidad de la UIUC pedir a todos los patrocinadores y socios que eliminan cualquier copia de mis muestras de habla de sus propias bases de datos.

**Casilla opcional:** Al marcar la siguiente casilla, también puede dar específicamente su consentimiento para que el equipo de investigación reproduzca o proporcione archivos de audio de su voz en un entorno público de investigación o publicación.

Doy mi consentimiento para que se compartan muestras de mi voz en público en conferencias de investigación y reuniones profesionales similares, y para que se compartan archivos de audio digitales de mi voz como parte de publicaciones de investigación profesional.

**Casilla opcional:** Al marcar la casilla de abajo, también puede dar específicamente su consentimiento para compartir su información de contacto con otros equipos de investigación, de modo que puedan ponerse en contacto con usted para ser un participante potencial en sus propios estudios de investigación.

Doy mi consentimiento para que mi dirección de correo electrónico y mi número de teléfono sean compartidos con otros equipos de investigación que podrán ponerse en contacto conmigo para participar en sus estudios.

Al hacer clic en el botón "ACEPTO" que aparece a continuación, indico mi deseo de participar voluntariamente en este estudio.

**ACEPTO**



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

### **Speech Accessibility Project: Individuals with disabilities helping researchers to improve technology**

#### **Consent Form for Recorded Caregivers**

You are being asked to participate in a voluntary research study that is called the “Speech Accessibility Project.” The purpose of this study is to help researchers at universities and companies to develop spoken-language user interfaces that work for people with atypical speech. Software programs that understand speech are developed using machine learning. Machine learning is a software development method that imitates the way humans learn, by testing the software repeatedly, and changing its behavior whenever it makes a mistake. In this way, the software learns to reduce the risk of similar mistakes in the future. By providing examples of the sounds of atypical speech, this study will make it possible for the software to learn how to make fewer mistakes when it encounters atypical speech in the future. Participating in this study is all online and will involve helping your friend or family member with a disability to record sentences and to upload the audio recordings to a secure encrypted storage system. We ask you to help your friend or family member with a disability to provide up to 600 speech samples or about 5 hours of your time which you can do any time online, at your own pace and convenience. Risks associated with this study are no greater than those in everyday life. There is no expected benefit of this research to you personally; the expected benefit to society is that companies and organizations using your speech recordings will be able to train better speech technology for people with speech and motor disabilities, which will make technology more accessible for everyone.

**Project Name:** The Speech Accessibility Project, which has the website:  
<https://speechaccessibilityproject.beckman.illinois.edu/>

**Principal Investigator Name and Title:** Professor Mark Hasegawa-Johnson, University of Illinois Urbana-Champaign (“UIUC”), Beckman Institute and Department of Electrical and Computer Engineering.

**Contact Information:** [speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu) is the team email address.

**Organizations that will perform this study:** University of Illinois Urbana-Champaign (UIUC) is the lead organization that will perform this study and has partnered with Team Gleason Foundation, a non-profit organization that advocates for people living with ALS.

**Sponsor:** The Speech Accessibility Project is funded by a consortium that currently includes five companies: Amazon, Apple, Google, Microsoft, and Meta (parent company of Facebook). In the future, other companies, universities, and organizations may join this consortium as well.

#### **Why am I being asked?**

You are being asked to participate in this study because you are an adult who assists a person with a disability to interact with computers, and that person has already qualified to be in the study by successfully answering our pre-screening questions and providing a set of speech samples, and that person either is not able to read prompts from the screen, or does not wish to do so.

We plan to recruit more than 1,000 people to be in this study. Because the purpose of this study is to help organizations develop better speech interfaces for technologies, our team needs to gather many different types of speech samples from many types of voices.

Your participation in this research is voluntary. Your decision whether or not to participate will not affect



## University of Illinois Urbana-Champaign Online Consent Form

your current or future dealings with the University of Illinois Urbana-Champaign, its partners, or its sponsors of this project. If you decide to participate, you are free to withdraw at any time without affecting that relationship.

### What procedures are involved?

We will provide you with access to a website where you and the person that you are assisting can record up to 600 speech samples. Some of the prompts will instruct you to read a sentence out loud so that the person you're assisting can repeat it out loud. Some of the prompts will instruct you to ask the person you're assisting a question, so that they can speak their answer to the question. You and the person that you are assisting can do these recordings at any time, 24/7, at whatever pace you wish. We encourage you to record at different times of day so that we can learn how your speech might change throughout the day. And don't worry about giving us "bad" samples – all data are really helpful. The more diverse our data set is, the better!

You and the person that you are assisting have the option to work with our partner organization, Team Gleason, to get help. Team Gleason can provide you with a mentor to help you do your recordings, give you advice about your microphone settings, and give you tips for when and where to do your recordings.

You may take up to three months to finish recording the 600 speech samples. If you will need more than three months, please contact the research team at [speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu).

We will ask the person you're assisting whether they consider U.S. English to be their first language. People who speak English as a second language are sometimes better recognized by devices that also know multiple languages. If you provide this information, it will be shared with researchers using the data.

We will also ask the person you're assisting to provide optional data about themselves, such as age, racial and ethnic identity, and questions about how their speech ability may have impacted their everyday life. They are not required to answer any of these questions. This information will only be used to confirm diverse representation in this data collection, and, if provided, will not be shared outside of the UIUC research team.

The UIUC research team will review all of your speech samples, and we will edit them to protect your privacy by removing any references to personally identifiable information about you or others that you might have said in your answers to our question prompts. We will also annotate your samples with "tags" that refer to technical aspects of the acoustic signals in your samples. We will not attach your name or email address to these samples; instead, we will assign a unique participant identification code to you and will annotate your samples with that code. Only the UIUC research team will know the "identity key" that relates your email address to your participant identification code.

The UIUC Speech Accessibility Project research team has established a secure, fully encrypted storage system at the UIUC Beckman Institute to store your samples. We will separately store your identifiable information in a separate secure database. We have also established a secure data sharing mechanism with the partners in our consortium. Right now, our partners are the five company sponsors of this project: Amazon, Apple, Google, Meta, and Microsoft. UIUC researchers and researchers at each of the partners will use the collected speech samples to train machine learning algorithms to perform better on speech recognition tasks. These partners will only have access to your speech samples and to no other data about you.



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

The five corporate sponsors of this research at UIUC have signed a legal agreement with the University that sets forth the permitted purposes of the speech data collected from this research project. These permitted purposes are only for research and technology development as stated here:

- Your samples will be kept forever, to assist in research and technology development in perpetuity.
- Your samples are not identified with your name or email address, but only with a unique participant code that is only known to the UIUC research team.
- Your samples will be analyzed by the UIUC research team to verify that the study is being done properly and to provide advice to other researchers about how to design or improve future studies.
- Your samples will be analyzed and combined with other people's samples to create aggregate results that will be published in the research literature.
- Your samples will be used to advance understanding of machine learning-related data variances associated with atypical speech variations, and eventually to develop products, technologies, and services by or on behalf of our sponsors to improve communications and increase access to technology, including the creation and training of machine learning solutions and artificial intelligence models for both research and commercial purposes.

In the future, any additional partner organizations in this consortium will sign data use agreements with the same terms and conditions. For these future uses of your data, we will not ask you to consent again. In other words, the future use of your de-identified data could happen without additional informed consent.

In summary, your speech samples will help researchers at UIUC and its partner organizations to develop better speech recognition technology algorithms and may eventually help commercial products perform more effectively for people with atypical speech.

### What are the potential risks and discomforts?

In general, we believe that the risks of your participation in this research study are no more than you would encounter in everyday life. However, we do recognize two issues that may be risks in this study, and we also have designed measures to minimize these risks.

First, you may get tired. To minimize your tiredness, we have arranged to let the participant record their speech samples over multiple sessions. They can log in to the system at a time and location convenient to them, and record at their own pace. They can also log out at any time, then log in again to continue recording as you wish. We hope that this flexibility and convenience will minimize any fatigue that you may experience.

Second, as is true for almost any online data collection research study, there is a risk that your speech samples could be released for purposes and to other individuals, companies, and organizations beyond the purposes to which you have given consent. Please know that the UIUC team has designed measures to minimize the possibility of unauthorized data release include the following:

- Participant code numbers will be used to annotate all speech recordings; your name and email address are not attached to any speech samples directly.
- All organizations using your speech recordings are required to sign a data use agreement, in which



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

they agree that they will store the data securely, that they will prevent unauthorized access, and that they will not attempt to identify you or obtain any other identifying information about you.

- Researchers from outside UIUC will not know your name, e-mail address, state of residence, or any other identifying information about you. Identifying information will be stored separately from your speech recordings, in a database accessible only to UIUC research team members. In addition, as stated above, the UIUC team will edit and annotate your speech samples to protect your privacy.
- You are allowed to withdraw your voluntary consent at any time and request that all of your data be destroyed. When you notify the UIUC team that you wish to have all of your speech samples destroyed, UIUC will ensure that your speech recordings are removed from the UIUC secure system and will direct our partners that you wish your data to be removed from any copies they have as well.

### **Are there benefits to participating in the research?**

There is no immediate benefit of this research to you personally.

Information obtained from you in this research may help in the development of commercial products by UIUC, its research partners, or members of the consortium. There are no plans to provide financial compensation to you should this occur.

The expected benefit to society is that companies and organizations will be able to develop better spoken-language user interfaces so that people with atypical speech can be as easily understood as other people with typical speech.

### **What other options are there?**

You have the option to not participate in this study. This study is entirely voluntary.

### **Will my study-related information be kept confidential?**

The UIUC research team has made every effort to protect the audio records of your speech samples. The UIUC team will edit out identifying information from your speech samples (such as your name) and will only store your audio files with a participant code number. This participant code number is stored with your identifiable information on a secure server at the UIUC Beckman Institute, and this “identity key” information is only accessible by the UIUC research team. This information is necessary in order to pay you for your participation in this project.

As described previously, the UIUC team will also listen to all of your speech samples and edit out any personally identifiable information.

Faculty, staff, students, and others with permission or authority to see your study information (such as University auditors) will maintain its confidentiality to the extent permitted and required by laws and university policies.

Speech recordings will be distributed only to companies and organizations that have signed a data use agreement, agreeing to confidentiality terms specified above. No other personally identifiable information will be published or presented.



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

Because this is a research project, we do expect to publish research papers as a result of this study. These papers may be in the public domain. Research papers will never state your name or any identifying information about you. Optionally, you may choose to allow the university to share anonymized samples of your voice as part of research publications or presentations, but this is not required.

### **Will I be reimbursed for any expenses or paid for my participation in this research?**

The person that you are assisting will be compensated up to \$180 as follows: \$60 after recording the 200<sup>th</sup> sample, an additional \$60 after recording the 400<sup>th</sup> sample, and an additional \$60 after recording the 600<sup>th</sup> sample.

In addition, since you are helping them to record their speech samples, you will also receive compensation for helping them. As a named caregiver you will be compensated with up to \$90 as follows: \$30 after they record the 200<sup>th</sup> sentence, an additional \$30 after they record the 400<sup>th</sup> sentence, and an additional \$30 after they record the 600<sup>th</sup> sentence.

Payments will be made through email using Amazon ecodes.

### **Can I withdraw or be removed from the study?**

Yes. You are free to withdraw your consent and discontinue participation at any time by notifying the UIUC research team by email or phone. The UIUC research team also have the right to stop your participation in this study without your consent if they believe it is in your best interests. If you withdraw consent, the UIUC research team will remove your speech recordings from the UIUC database and will notify the sponsors and partners that they should also remove your data from their own databases using only your anonymized participant code number. If you withdraw consent, your data will not be used for any future research. If the data you previously provided has been used to inform or create computer code or machine learning algorithms it may not be possible for your data to be removed from them.

### **Will data collected from me be used for any other research?**

As stated above, UIUC and its partners in the Speech Accessibility Project consortium will use your speech samples to advance understanding of machine learning-related data variances associated with atypical speech variations, and to develop products, technologies, and services to improve communications and increase access to technology, including computational use, and the creation and training of machine learning solutions and artificial intelligence models including for research and for commercial purposes.

It is possible that, in the future, other organizations will be granted access to the data, in addition to the founding consortium partners that are Amazon, Apple, Google, Meta, and Microsoft. These other organizations may be other universities or other companies. All organizations seeking access to the data will be required to sign data use agreements that specify the same terms as above: that your speech samples will only be used for research and technology development, that your data will be stored securely, and that you may withdraw consent at any time. You may not be asked to provide additional consent for these future uses of your data.

### **Who should I contact if I have questions?**

Contact the UIUC research team leader, Professor Mark Hasegawa-Johnson ([jhasegaw@illinois.edu](mailto:jhasegaw@illinois.edu)), if you have any questions about this study or your part in it, or if you have concerns or complaints about the research.



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

### What are my rights as a research subject?

If you have any questions about your rights as a research subject, including concerns, complaints, or to offer input, you may call the UIUC Office for the Protection of Research Subjects (OPRS) at 217-333-2670 or e-mail OPRS at [irb@illinois.edu](mailto:irb@illinois.edu). If you would like to complete a brief survey to provide OPRS feedback about your experiences as a research participant, please follow the link [here](#) or through a link on the OPRS website: <https://oprs.research.illinois.edu/>. You will have the option to provide feedback or concerns anonymously or you may provide your name and contact information for follow-up purposes.

**Please print this consent form if you would like to retain a copy for your records.**

**Please read each item carefully and check each box to indicate your agreement.**

- I have read and understand the above consent form.
- I understand that my speech samples are collected for the purpose of improving speech recognition technology and are protected by a legal agreement between UIUC and the current members of the Speech Accessibility Project coalition, and that future members of the data coalition and any future data licensees will also be bound by similar legal agreements.
- I certify that I am 18 years old or older.
- I understand that I must maintain a valid email address in order to participate in this study and to be paid for my participation.
- I certify that only I and the person that I am assisting will use our login to record speech samples. I will not allow other people to record speech samples on our behalf.
- I understand that the only financial compensation that I will receive for participating in this study is the payment for contributing speech samples, which will be \$30 for every batch of 200 samples that the person I am assisting provides.
- I understand that my goal is for the person I am assisting to provide 600 samples total, and that therefore my maximum payment will be \$90.
- I understand that at any time, I can end my participation and withdraw my consent. When I do so, the UIUC research team will remove all of my speech samples from the UIUC secure system. I also understand that it is UIUC's responsibility to ask all sponsors and partners to remove any copies of my speech samples from their own databases as well.



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

**Optional checkbox:** By checking the box below, you may also specifically grant your consent for the research team to play or provide audio files of your voice in a public research setting or publication.

I consent to have my voice samples shared in public at research conferences and similar professional meetings, and to have digital audio files of my voice be shared as part of professional research publications.

**Optional checkbox:** By checking the box below, you may also specifically grant your consent to share your contact information with other research teams, so that they may contact you to be a potential participant in their own research studies.

I consent to have my email address and phone number be shared with other research teams, who may contact me to participate in their studies.

By clicking the "I CONSENT" button below, I indicate my willingness to voluntarily take part in this study.

**I CONSENT**



## University of Illinois Urbana-Champaign Online Consent Form



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

### **Speech Accessibility Project: Individuals with disabilities helping researchers to improve technology**

You are being asked to join a voluntary research study called the “Speech Accessibility Project.” This study will help computers learn to understand people with different types of speech. A computer learns to understand speech in almost the same way humans learn. It listens to speech, guesses what it says, and then guesses again whenever it makes a mistake. If the computer tries many times, then after a while it learns to guess correctly. By giving computers some examples of different speech, we can help them to better understand speech differences in the future.

Participating in this study is all online. You will record your speech online. We ask you to provide up to 600 speech samples. It should take you about 5 hours. You don't have to do it all at once. You can stop and start at any time, at your own pace and whenever you want. Risks associated with this study are no more than those in everyday life. You will not get any personal benefit from this research. Society may benefit, because computers might learn how to better understand speech differences.

**Project Name:** The Speech Accessibility Project, which has the website:  
<https://speechaccessibilityproject.beckman.illinois.edu/>

**Principal Investigator Name and Title:** Professor Mark Hasegawa-Johnson, University of Illinois Urbana-Champaign (“UIUC”), Beckman Institute and Department of Electrical and Computer Engineering.

**Contact Information:** [speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu) is the team email address.

**Organizations that will perform this study:** University of Illinois Urbana-Champaign (UIUC) is the lead organization that will perform this study and has partnered with Team Gleason Foundation, a non-profit organization that advocates for people living with ALS.

**Sponsor:** The Speech Accessibility Project is funded by a consortium that currently includes five companies: Amazon, Apple, Google, Microsoft, and Meta (parent company of Facebook). In the future, other companies, universities, and organizations may join this consortium as well.

### **Why am I being asked?**

You are being asked to participate in this study because you are an adult, and because we think that computers should learn to understand speech like yours.

We plan to recruit more than 1,000 people to be in this study. Because the purpose of this study is to help computers learn to understand lots of different types of speech, we need to gather many different speech recordings from many types of people.

Your participation in this research is voluntary. Your decision whether or not to participate will not affect your current or future relationship with the University of Illinois, its partners, or its sponsors of this project. Even if you decide to participate, you can choose to stop at any time without affecting that relationship.

### **What will you do for this research?**

We will give you access to a website where you can record up to 600 speech recordings. For each recording, you will either read a sentence, or repeat a sentence that you hear somebody else saying, or answer a question



## University of Illinois Urbana-Champaign Online Consent Form

out loud. You can do these recordings any time, whenever you want. We encourage you to record at different times of day so that we can learn how your speech might change throughout the day. And don't worry about giving us "bad" samples – all data are really helpful.

We will also give you access to a mentor. Your mentor will help you do your recordings, give you advice about your microphone settings, and give you tips for when and where to do your recordings.

You may take up to three months to finish your recordings. If you will need more than three months, please tell us so by e-mail at [speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu).

We will ask you whether you consider U.S. English to be your first language. People who speak English as a second language are sometimes better recognized by computers that also know different languages. If you provide this information, it will be shared with researchers using the data.

We will also ask you to provide optional data about yourself, such as your age, race and ethnicity. We will also ask questions about how your speech ability affects your everyday life. You are not required to answer any of these questions. If you tell us these things, it will help us to know when we have collected enough different types of speech. Other people who use your speech samples will not know your answers to these questions.

We will listen to all of your speech. If you say your own name or the names of anybody close to you, we will cut those sections out of your speech recording to protect your privacy. We will also write a short list of things we think a computer should know about your speech. We will not attach your name or email address to these samples. Instead, we will assign an ID code to you, and we will label your samples with that code. Only the UIUC research team will know the "identity key" that relates your email address to your ID code.

We will store your speech in a special secure computer system at UIUC. We will store your ID on a different computer, also secure. When we give your speech to other people, we will require them to know the password before they can listen to it. They will only have access to your speech samples and they will not know your ID or anything else about you.

Right now, the people who will get copies of your speech are the five company sponsors of this project: Amazon, Apple, Google, Meta, and Microsoft. Those five companies signed a contract with the University saying that they will only use your recordings to help computers better understand speech. Some of the terms of the contract include:

- Your recordings will be kept forever, so that computers in the future can also learn to understand speech like yours.
- Your recordings are not identified with your name or email address, but only with an ID code. Only UIUC will know your name or your email address.
- UIUC will listen to your recordings before we give them to anybody else.
- Scientists might study your speech, and write papers about it, but they will not publish your speech or your identity.
- Your speech might be used to develop new computer programs, including both free programs and non-free programs that help computers to understand people with speech like yours.



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

In the future we might let other people use your recordings. We will not ask for your permission every time we give the speech to somebody new, as long as they agree to all of the things listed above. In other words, you agree that other people can use your speech as long as they agree to the things listed above.

In summary, your speech samples will help researchers at UIUC and other places to make computers that can understand speech like yours.

### **What are the potential risks and discomforts?**

The risks are no more than what happens in everyday life. We can think of two things that you might not like about this study, so here is how we will try to avoid these risks:

First, you may get tired. To help, we will let you record your speech samples over multiple sessions. You can log in at any time, from anywhere, and record at your own pace. You can also log out at any time, then log in again later. We hope that this will help you to avoid getting tired.

Second, even though we are working hard to keep your data safe, there is a risk that somebody might steal your data from us. Here is what we have done to keep that from happening:

- Your name and e-mail are not stored with your speech. Instead, your speech is labeled using an ID number.
- Before anybody from outside UIUC can use your speech, they need to sign a contract. In the contract, they promise to store your speech securely. They agree that only approved people can use your speech. They agree that they won't try to find out who you are. They also agree that, if we tell them to destroy all of your speech recordings, they will.
- Nobody from outside UIUC will know your name, your e-mail address, or what state you live in. That information will only be stored at UIUC, on a computer that only UIUC team members can use.
- Before your speech is given to anybody else, we will listen to your speech. If you say your name or the name of anybody you know, we will cut out those parts.
- You can change your mind at any time. If you change your mind, and ask us to destroy your speech recordings, we will.

### **Are there benefits to participating in the research?**

There is no benefit to you personally.

Your recordings might help companies to make computer software that they can sell. You will not receive any extra payment if that happens.

The whole world may benefit if computers learn how to better understand people with speech like yours.

### **What other options are there?**

You have the option to not participate in this study. This study is voluntary.

### **Will my study-related information be kept confidential?**

Only people from UIUC will know your e-mail address or your name. We only need to know those things so that we can pay you for your participation in this project.

The UIUC team will also listen to all of your speech samples and edit out any personal names.



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

Faculty, staff, students, and other people from UIUC will follow the law and UIUC policies in order to protect your confidentiality.

Speech recordings will be given only to people who have signed a contract, as described above.

Because this is a research project, we plan to publish research papers about this study. Research papers will never give your name or identity. If you wish, you may choose to allow the university to share anonymous samples of your voice in research presentations, but you don't need to.

### **Will I be reimbursed for any expenses or paid for my participation in this research?**

You will be paid up to \$180. You will be paid \$60 after recording your 200<sup>th</sup> speech sample. You will be paid another \$60 after recording your 400<sup>th</sup> sample. You will get a final \$60 after recording your 600<sup>th</sup> sample.

In addition to your own payment, if somebody helps you to use your web browser for this project, they can be paid up to \$90. They will be paid \$30 after you record your 200<sup>th</sup> sample, an additional \$30 after you record your 400<sup>th</sup> sample, and a final \$30 after you record your 600<sup>th</sup> sample.

Payments will be made through email using Amazon ecodes.

### **Can I withdraw or be removed from the study?**

Yes. You can stop participating at any time, and you can tell us to delete your recordings. We also have the right to stop your participation if we think it is in your best interests. If you want us to delete your recordings, you only need to tell us that you want them deleted. If you tell us to delete your recordings, we will delete them. We will also tell your participant ID number to anybody else who has a copy of your recordings, and say that they need to delete your recordings. Nobody will ever use your recordings after that for any future research. If some computers have already learned from your recordings, it may not be possible to force the computer to unlearn what it has learned.

### **Will data collected from me be used for any other research?**

UIUC, Amazon, Apple, Google, Meta and Microsoft will use your speech recordings to make computers that better understand speech like yours.

In the future, we may let other people use your recordings, if they are willing to sign a contract. The contract will promise to use your speech only to help make better computers. The contract will also promise to protect your data. The contract will also promise that they will delete your recordings if we say that your recordings should be deleted. We will not ask you for your permission again. As long as new users of the data agree to the same terms, we will assume that you have already given permission.

### **Who should I contact if I have questions?**

Contact the UIUC research team leader, Professor Mark Hasegawa-Johnson ([jhasegaw@illinois.edu](mailto:jhasegaw@illinois.edu)), if you have any questions about this study or your part in it. You can also contact him if you have any worries or complaints about the research.

### **What are my rights as a research subject?**



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

If you have any questions about your rights as a research subject, including worries, complaints, or ideas, you may call the UIUC Office for the Protection of Research Subjects (OPRS) at 217-333-2670 or e-mail OPRS at [irb@illinois.edu](mailto:irb@illinois.edu).

If you would like to complete a brief survey to provide OPRS feedback about your experiences as a research participant, please follow the link [here](#) or through a link on the OPRS website: <https://oprs.research.illinois.edu/>. You will be able to provide feedback or concerns without giving your name, or you can give us your name and contact information.

**Please print this consent form if you would like a copy for your records.**

**Please read each item carefully and check each box to indicate your agreement.**

- I understand this consent form.
- I understand that my speech samples will be used to help computers better understand speech like mine. I understand that my speech samples are protected by a contract between UIUC and the current members of the Speech Accessibility Project. I understand that future users of the data will have to sign a similar contract before they can use the data.
- I am 18 years old or older.
- I have a valid email address that can use to send me my payments.
- I won't let other people record their speech in place of me.
- The only payment I will get is the payment for contributing speech samples, which will be \$60 for every batch of 200 samples I provide.
- My goal is to provide 600 samples total. My maximum payment will be \$180.
- If I need a caregiver to help me make recordings, I will tell UIUC, so that my caregiver can also get paid for participating.
- I can stop participating at any time. Any time I want to, I can tell UIUC to delete all of my recordings. If I tell UIUC to delete my recordings, they will delete them. They will also tell anybody else who has copies to delete them.

**Optional:** If you want to let us play your speech as part of a presentation or an online demo of the Speech Accessibility Project, you can check this box. This is not required.



University of Illinois Urbana-Champaign  
Online Consent Form



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

*I agree to let my speech be played during public presentations about the Speech Accessibility Project or as part of an online demo of the project.*

**Optional:** If you wish, we can give your name and your e-mail address or phone number to people who are planning other research studies, so you can be a participant in their studies too. This is not required.

*I agree to that my email address and phone number can be shared with other research teams, who may contact me to participate in their studies.*

By clicking the “I CONSENT” button below, I indicate my willingness to voluntarily take part in this study.

**I CONSENT**



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

### **Speech Accessibility Project: Individuals with disabilities helping researchers to improve technology**

You are being asked to join a voluntary research study called the “Speech Accessibility Project.” This study will help computers learn to understand people with different types of speech. A computer learns to understand speech in almost the same way humans learn. It listens to speech, guesses what it says, and then guesses again whenever it makes a mistake. If the computer tries many times, then after a while it learns to guess correctly. By giving computers some examples of different speech, we can help them to better understand speech differences in the future.

Participating in this study is all online. You will record your speech online. We ask you to provide up to 600 speech samples. It should take you about 5 hours. You don't have to do it all at once. You can stop and start at any time, at your own pace and whenever you want. Risks associated with this study are no more than those in everyday life. You will not get any personal benefit from this research. Society may benefit, because computers might learn how to better understand speech differences.

**Project Name:** The Speech Accessibility Project, which has the website:  
<https://speechaccessibilityproject.beckman.illinois.edu/>

**Principal Investigator Name and Title:** Professor Mark Hasegawa-Johnson, University of Illinois Urbana-Champaign (“UIUC”), Beckman Institute and Department of Electrical and Computer Engineering.

**Contact Information:** [speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu) is the team email address.

**Organizations that will perform this study:** University of Illinois Urbana-Champaign (UIUC) is the lead organization that will perform this study and has partnered with Lingraphica.

**Sponsor:** The Speech Accessibility Project is funded by a consortium that currently includes five companies: Amazon, Apple, Google, Microsoft, and Meta (parent company of Facebook). In the future, other companies, universities, and organizations may join this consortium as well.

### **Why am I being asked?**

You are being asked to participate in this study because you are an adult, and because we think that computers should learn to understand speech like yours.

We plan to recruit more than 1,000 people to be in this study. Because the purpose of this study is to help computers learn to understand lots of different types of speech, we need to gather many different speech recordings from many types of people.

Your participation in this research is voluntary. Your decision whether or not to participate will not affect your current or future relationship with the University of Illinois, its partners, or its sponsors of this project. Even if you decide to participate, you can choose to stop at any time without affecting that relationship.

### **What will you do for this research?**

We will give you access to a website where you can record up to 600 speech recordings. For each recording, you will either read a sentence, or repeat a sentence that you hear somebody else saying, or answer a question out loud. You can do these recordings any time, whenever you want. We encourage you to record at different



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

times of day so that we can learn how your speech might change throughout the day. And don't worry about giving us "bad" samples – all data are really helpful.

We will also give you access to a mentor. Your mentor will help you do your recordings, give you advice about your microphone settings, and give you tips for when and where to do your recordings.

You may take up to three months to finish your recordings. If you will need more than three months, please tell us so by e-mail at [speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu).

We will ask you whether you consider U.S. English to be your first language. People who speak English as a second language are sometimes better recognized by computers that also know different languages. If you provide this information, it will be shared with researchers using the data.

We will also ask you to provide optional data about yourself, such as your age, race and ethnicity. We will also ask questions about how your speech ability affects your everyday life. You are not required to answer any of these questions. If you tell us these things, it will help us to know when we have collected enough different types of speech. Other people who use your speech samples will not know your answers to these questions.

We will listen to all of your speech. If you say your own name or the names of anybody close to you, we will cut those sections out of your speech recording to protect your privacy. We will also write a short list of things we think a computer should know about your speech. We will not attach your name or email address to these samples. Instead, we will assign an ID code to you, and we will label your samples with that code. Only the UIUC research team will know the "identity key" that relates your email address to your ID code.

We will store your speech in a special secure computer system at UIUC. We will store your ID on a different computer, also secure. When we give your speech to other people, we will require them to know the password before they can listen to it. They will only have access to your speech samples and they will not know your ID or anything else about you.

Right now, the people who will get copies of your speech are the five company sponsors of this project: Amazon, Apple, Google, Meta, and Microsoft. Those five companies signed a contract with the University saying that they will only use your recordings to help computers better understand speech. Some of the terms of the contract include:

- Your recordings will be kept forever, so that computers in the future can also learn to understand speech like yours.
- Your recordings are not identified with your name or email address, but only with an ID code. Only UIUC will know your name or your email address.
- UIUC will listen to your recordings before we give them to anybody else.
- Scientists might study your speech, and write papers about it, but they will not publish your speech or your identity.
- Your speech might be used to develop new computer programs, including both free programs and non-free programs that help computers to understand people with speech like yours.

In the future we might let other people use your recordings. We will not ask for your permission every



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

time we give the speech to somebody new, as long as they agree to all of the things listed above. In other words, you agree that other people can use your speech as long as they agree to the things listed above.

In summary, your speech samples will help researchers at UIUC and other places to make computers that can understand speech like yours.

### **What are the potential risks and discomforts?**

The risks are no more than what happens in everyday life. We can think of two things that you might not like about this study, so here is how we will try to avoid these risks:

First, you may get tired. To help, we will let you record your speech samples over multiple sessions. You can log in at any time, from anywhere, and record at your own pace. You can also log out at any time, then log in again later. We hope that this will help you to avoid getting tired.

Second, even though we are working hard to keep your data safe, there is a risk that somebody might steal your data from us. Here is what we have done to keep that from happening:

- Your name and e-mail are not stored with your speech. Instead, your speech is labeled using an ID number.
- Before anybody from outside UIUC can use your speech, they need to sign a contract. In the contract, they promise to store your speech securely. They agree that only approved people can use your speech. They agree that they won't try to find out who you are. They also agree that, if we tell them to destroy all of your speech recordings, they will.
- Nobody from outside UIUC will know your name, your e-mail address, or what state you live in. That information will only be stored at UIUC, on a computer that only UIUC team members can use.
- Before your speech is given to anybody else, we will listen to your speech. If you say your name or the name of anybody you know, we will cut out those parts.
- You can change your mind at any time. If you change your mind, and ask us to destroy your speech recordings, we will.

### **Are there benefits to participating in the research?**

There is no benefit to you personally.

Your recordings might help companies to make computer software that they can sell. You will not receive any extra payment if that happens.

The whole world may benefit if computers learn how to better understand people with speech like yours.

### **What other options are there?**

You have the option to not participate in this study. This study is voluntary.

### **Will my study-related information be kept confidential?**

Only people from UIUC will know your e-mail address or your name. We only need to know those things so that we can pay you for your participation in this project.

The UIUC team will also listen to all of your speech samples and edit out any personal names.



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

Faculty, staff, students, and other people from UIUC will follow the law and UIUC policies in order to protect your confidentiality.

Speech recordings will be given only to people who have signed a contract, as described above.

Because this is a research project, we plan to publish research papers about this study. Research papers will never give your name or identity. If you wish, you may choose to allow the university to share anonymous samples of your voice in research presentations, but you don't need to.

### **Will I be reimbursed for any expenses or paid for my participation in this research?**

You will be paid up to \$180. You will be paid \$60 after recording your 200<sup>th</sup> speech sample. You will be paid another \$60 after recording your 400<sup>th</sup> sample. You will get a final \$60 after recording your 600<sup>th</sup> sample.

In addition to your own payment, if somebody helps you to use your web browser for this project, they can be paid up to \$90. They will be paid \$30 after you record your 200<sup>th</sup> sample, an additional \$30 after you record your 400<sup>th</sup> sample, and a final \$30 after you record your 600<sup>th</sup> sample.

Payments will be made through email using Amazon ecodes.

### **Can I withdraw or be removed from the study?**

Yes. You can stop participating at any time, and you can tell us to delete your recordings. We also have the right to stop your participation if we think it is in your best interests. If you want us to delete your recordings, you only need to tell us that you want them deleted. If you tell us to delete your recordings, we will delete them. We will also tell your participant ID number to anybody else who has a copy of your recordings, and say that they need to delete your recordings. Nobody will ever use your recordings after that for any future research. If some computers have already learned from your recordings, it may not be possible to force the computer to unlearn what it has learned.

### **Will data collected from me be used for any other research?**

UIUC, Amazon, Apple, Google, Meta and Microsoft will use your speech recordings to make computers that better understand speech like yours.

In the future, we may let other people use your recordings, if they are willing to sign a contract. The contract will promise to use your speech only to help make better computers. The contract will also promise to protect your data. The contract will also promise that they will delete your recordings if we say that your recordings should be deleted. We will not ask you for your permission again. As long as new users of the data agree to the same terms, we will assume that you have already given permission.

### **Who should I contact if I have questions?**

Contact the UIUC research team leader, Professor Mark Hasegawa-Johnson ([jhasegaw@illinois.edu](mailto:jhasegaw@illinois.edu)), if you have any questions about this study or your part in it. You can also contact him if you have any worries or complaints about the research.

### **What are my rights as a research subject?**



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

If you have any questions about your rights as a research subject, including worries, complaints, or ideas, you may call the UIUC Office for the Protection of Research Subjects (OPRS) at 217-333-2670 or e-mail OPRS at [irb@illinois.edu](mailto:irb@illinois.edu).

If you would like to complete a brief survey to provide OPRS feedback about your experiences as a research participant, please follow the link [here](#) or through a link on the OPRS website: <https://oprs.research.illinois.edu/>. You will be able to provide feedback or concerns without giving your name, or you can give us your name and contact information.

**Please print this consent form if you would like a copy for your records.**

**Please read each item carefully and check each box to indicate your agreement.**

- I understand this consent form.
- I understand that my speech samples will be used to help computers better understand speech like mine. I understand that my speech samples are protected by a contract between UIUC and the current members of the Speech Accessibility Project. I understand that future users of the data will have to sign a similar contract before they can use the data.
- I am 18 years old or older.
- I have a valid email address that can use to send me my payments.
- I won't let other people record their speech in place of me.
- The only payment I will get is the payment for contributing speech samples, which will be \$60 for every batch of 200 samples I provide.
- My goal is to provide 600 samples total. My maximum payment will be \$180.
- If I need a caregiver to help me make recordings, I will tell UIUC, so that my caregiver can also get paid for participating.
- I can stop participating at any time. Any time I want to, I can tell UIUC to delete all of my recordings. If I tell UIUC to delete my recordings, they will delete them. They will also tell anybody else who has copies to delete them.

**Optional:** If you want to let us play your speech as part of a presentation or an online demo of the Speech Accessibility Project, you can check this box. This is not required.



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

University of Illinois Urbana-Champaign  
Online Consent Form

*I agree to let my speech be played during public presentations about the Speech Accessibility Project or as part of an online demo of the project.*

**Optional:** If you wish, we can give your name and your e-mail address or phone number to people who are planning other research studies, so you can be a participant in their studies too. This is not required.

*I agree to that my email address and phone number can be shared with other research teams, who may contact me to participate in their studies.*

By clicking the “I CONSENT” button below, I indicate my willingness to voluntarily take part in this study.

**I CONSENT**



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

### **Speech Accessibility Project: Individuals with disabilities helping researchers to improve technology**

#### **Consent Form for Recorded Caregivers**

You are being asked to participate in a voluntary research study that is called the “Speech Accessibility Project.” The purpose of this study is to help researchers at universities and companies to develop spoken-language user interfaces that work for people with atypical speech. Software programs that understand speech are developed using machine learning. Machine learning is a software development method that imitates the way humans learn, by testing the software repeatedly, and changing its behavior whenever it makes a mistake. In this way, the software learns to reduce the risk of similar mistakes in the future. By providing examples of the sounds of atypical speech, this study will make it possible for the software to learn how to make fewer mistakes when it encounters atypical speech in the future. Participating in this study is all online and will involve helping your friend or family member with a disability to record sentences and to upload the audio recordings to a secure encrypted storage system. We ask you to help your friend or family member with a disability to provide up to 600 speech samples or about 5 hours of your time which you can do any time online, at your own pace and convenience. Risks associated with this study are no greater than those in everyday life. There is no expected benefit of this research to you personally; the expected benefit to society is that companies and organizations using your speech recordings will be able to train better speech technology for people with speech and motor disabilities, which will make technology more accessible for everyone.

**Project Name:** The Speech Accessibility Project, which has the website:  
<https://speechaccessibilityproject.beckman.illinois.edu/>

**Principal Investigator Name and Title:** Professor Mark Hasegawa-Johnson, University of Illinois Urbana-Champaign (“UIUC”), Beckman Institute and Department of Electrical and Computer Engineering.

**Contact Information:** [speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu) is the team email address.

**Organizations that will perform this study:** University of Illinois Urbana-Champaign (UIUC) is the lead organization that will perform this study and has partnered with Lingraphica.

**Sponsor:** The Speech Accessibility Project is funded by a consortium that currently includes five companies: Amazon, Apple, Google, Microsoft, and Meta (parent company of Facebook). In the future, other companies, universities, and organizations may join this consortium as well.

#### **Why am I being asked?**

You are being asked to participate in this study because you are an adult who assists a person with a disability to interact with computers, and that person has already qualified to be in the study by successfully answering our pre-screening questions and providing a set of speech samples, and that person either is not able to read prompts from the screen, or does not wish to do so.

We plan to recruit more than 1,000 people to be in this study. Because the purpose of this study is to help organizations develop better speech interfaces for technologies, our team needs to gather many different types of speech samples from many types of voices.

Your participation in this research is voluntary. Your decision whether or not to participate will not affect your current or future dealings with the University of Illinois Urbana-Champaign, its partners, or its



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

sponsors of this project. If you decide to participate, you are free to withdraw at any time without affecting that relationship.

### What procedures are involved?

We will provide you with access to a website where you and the person that you are assisting can record up to 600 speech samples. Some of the prompts will instruct you to read a sentence out loud so that the person you're assisting can repeat it out loud. Some of the prompts will instruct you to ask the person you're assisting a question, so that they can speak their answer to the question. You and the person that you are assisting can do these recordings at any time, 24/7, at whatever pace you wish. We encourage you to record at different times of day so that we can learn how your speech might change throughout the day. And don't worry about giving us "bad" samples – all data are really helpful. The more diverse our data set is, the better!

You and the person that you are assisting have the option to work with our partner organization, Lingraphica, to get help. Lingraphica can provide you with a mentor to help you do your recordings, give you advice about your microphone settings, and give you tips for when and where to do your recordings.

You may take up to three months to finish recording the 600 speech samples. If you will need more than three months, please contact the research team at UIUC at [speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu).

We will ask the person you're assisting whether they consider U.S. English to be their first language. People who speak English as a second language are sometimes better recognized by devices that also know multiple languages. If you provide this information, it will be shared with researchers using the data.

We will also ask the person you're assisting to provide optional data about themselves, such as age, racial and ethnic identity, and questions about how their speech ability may have impacted their everyday life. They are not required to answer any of these questions. This information will only be used to confirm diverse representation in this data collection, and, if provided, will not be shared outside of the UIUC research team.

The UIUC research team will review all of your speech samples, and we will edit them to protect your privacy by removing any references to personally identifiable information about you or others that you might have said in your answers to our question prompts. We will also annotate your samples with "tags" that refer to technical aspects of the acoustic signals in your samples. We will not attach your name or email address to these samples; instead, we will assign a unique participant identification code to you and will annotate your samples with that code. Only the UIUC research team will know the "identity key" that relates your email address to your participant identification code.

The UIUC Speech Accessibility Project research team has established a secure, fully encrypted storage system at the UIUC Beckman Institute to store your samples. We will separately store your identifiable information in a separate secure database. We have also established a secure data sharing mechanism with the partners in our consortium. Right now, our partners are the five company sponsors of this project: Amazon, Apple, Google, Meta, and Microsoft. UIUC researchers and researchers at each of the partners will use the collected speech samples to train machine learning algorithms to perform better on speech recognition tasks. These partners will only have access to your speech samples and to no other data about you.



## University of Illinois Urbana-Champaign Online Consent Form



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

The five corporate sponsors of this research at UIUC have signed a legal agreement with the University that sets forth the permitted purposes of the speech data collected from this research project. These permitted purposes are only for research and technology development as stated here:

- Your samples will be kept forever, to assist in research and technology development in perpetuity.
- Your samples are not identified with your name or email address, but only with a unique participant code that is only known to the UIUC research team.
- Your samples will be analyzed by the UIUC research team to verify that the study is being done properly and to provide advice to other researchers about how to design or improve future studies.
- Your samples will be analyzed and combined with other people's samples to create aggregate results that will be published in the research literature.
- Your samples will be used to advance understanding of machine learning-related data variances associated with atypical speech variations, and eventually to develop products, technologies, and services by or on behalf of our sponsors to improve communications and increase access to technology, including the creation and training of machine learning solutions and artificial intelligence models for both research and commercial purposes.

In the future, any additional partner organizations in this consortium will sign data use agreements with the same terms and conditions. For these future uses of your data, we will not ask you to consent again. In other words, the future use of your de-identified data could happen without additional informed consent.

In summary, your speech samples will help researchers at UIUC and its partner organizations to develop better speech recognition technology algorithms and may eventually help commercial products perform more effectively for people with atypical speech.

### What are the potential risks and discomforts?

In general, we believe that the risks of your participation in this research study are no more than you would encounter in everyday life. However, we do recognize two issues that may be risks in this study, and we also have designed measures to minimize these risks.

First, you may get tired. To minimize your tiredness, we have arranged to let you record your speech samples over multiple sessions. You can log in to the system at a time and location convenient to you, and record at your own pace. You can also log out at any time, then log in again to continue recording as you wish. We hope that this flexibility and convenience will minimize any fatigue that you may experience.

Second, as is true for almost any online data collection research study, there is a risk that your speech samples could be released for purposes and to other individuals, companies, and organizations beyond the purposes to which you have given consent. Please know that the UIUC team has designed measures to minimize the possibility of unauthorized data release include the following:

- Participant code numbers will be used to annotate all speech recordings; your name and email address are not attached to any speech samples directly.
- All organizations using your speech recordings are required to sign a data use agreement, in which they agree that they will store the data securely, that they will prevent unauthorized access, and that they will not attempt to identify you or obtain any other identifying information about you.



## University of Illinois Urbana-Champaign Online Consent Form



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

- Researchers from outside UIUC will not know your name, e-mail address, state of residence, or any other identifying information about you. Identifying information will be stored separately from your speech recordings, in a database accessible only to UIUC research team members. In addition, as stated above, the UIUC team will edit and annotate your speech samples to protect your privacy.
- You are allowed to withdraw your voluntary consent at any time and request that all of your data be destroyed. When you notify the UIUC team that you wish to have all of your speech samples destroyed, UIUC will ensure that your speech recordings are removed from the UIUC secure system and will direct our partners that you wish your data to be removed from any copies they have as well.

### **Are there benefits to participating in the research?**

There is no immediate benefit of this research to you personally.

Information obtained from you in this research may help in the development of commercial products by UIUC, its research partners, or members of the consortium. There are no plans to provide financial compensation to you should this occur.

The expected benefit to society is that companies and organizations will be able to develop better spoken-language user interfaces so that people with atypical speech can be as easily understood as other people with typical speech.

### **What other options are there?**

You have the option to not participate in this study. This study is entirely voluntary.

### **Will my study-related information be kept confidential?**

The UIUC research team has made every effort to protect the audio records of your speech samples. The UIUC team will edit out identifying information from your speech samples (such as your name) and will only store your audio files with a participant code number. This participant code number is stored with your identifiable information on a secure server at the UIUC Beckman Institute, and this "identity key" information is only accessible by the UIUC research team. This information is necessary in order to pay you for your participation in this project.

As described previously, the UIUC team will also listen to all of your speech samples and edit out any personally identifiable information.

Faculty, staff, students, and others with permission or authority to see your study information (such as University auditors) will maintain its confidentiality to the extent permitted and required by laws and university policies.

Speech recordings will be distributed only to companies and organizations that have signed a data use agreement, agreeing to confidentiality terms specified above. No other personally identifiable information will be published or presented.

Because this is a research project, we do expect to publish research papers as a result of this study. These papers may be in the public domain. Research papers will never state your name or any identifying



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

information about you. Optionally, you may choose to allow the university to share anonymized samples of your voice as part of research publications or presentations, but this is not required.

### **Will I be reimbursed for any expenses or paid for my participation in this research?**

The person that you are assisting will be compensated up to \$180 as follows: \$60 after recording the 200<sup>th</sup> sample, an additional \$60 after recording the 400<sup>th</sup> sample, and an additional \$60 after recording the 600<sup>th</sup> sample.

In addition, since you are helping them to record their speech samples, you will also receive compensation for helping them. As a named caregiver you will be compensated with up to \$90 as follows: \$30 after they record the 200<sup>th</sup> sentence, an additional \$30 after they record the 400<sup>th</sup> sentence, and an additional \$30 after they record the 600<sup>th</sup> sentence.

Payments will be made through email using Amazon ecodes.

### **Can I withdraw or be removed from the study?**

Yes. You are free to withdraw your consent and discontinue participation at any time by notifying the UIUC research team by email or phone. The UIUC research team also have the right to stop your participation in this study without your consent if they believe it is in your best interests. If you withdraw consent, the UIUC research team will remove your speech recordings from the UIUC database and will notify the sponsors and partners that they should also remove your data from their own databases using only your anonymized participant code number. If you withdraw consent, your data will not be used for any future research. If the data you previously provided has been used to inform or create computer code or machine learning algorithms it may not be possible for your data to be removed from them.

### **Will data collected from me be used for any other research?**

As stated above, UIUC and its partners in the Speech Accessibility Project consortium will use your speech samples to advance understanding of machine learning-related data variances associated with atypical speech variations, and to develop products, technologies, and services to improve communications and increase access to technology, including computational use, and the creation and training of machine learning solutions and artificial intelligence models including for research and for commercial purposes.

It is possible that, in the future, other organizations will be granted access to the data, in addition to the founding consortium partners that are Amazon, Apple, Google, Meta, and Microsoft. These other organizations may be other universities or other companies. All organizations seeking access to the data will be required to sign data use agreements that specify the same terms as above: that your speech samples will only be used for research and technology development, that your data will be stored securely, and that you may withdraw consent at any time. You may not be asked to provide additional consent for these future uses of your data.

### **Who should I contact if I have questions?**

Contact the UIUC research team leader, Professor Mark Hasegawa-Johnson ([jhasegaw@illinois.edu](mailto:jhasegaw@illinois.edu)), if you have any questions about this study or your part in it, or if you have concerns or complaints about the research.

### **What are my rights as a research subject?**



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

If you have any questions about your rights as a research subject, including concerns, complaints, or to offer input, you may call the UIUC Office for the Protection of Research Subjects (OPRS) at 217-333-2670 or e-mail OPRS at [irb@illinois.edu](mailto:irb@illinois.edu). If you would like to complete a brief survey to provide OPRS feedback about your experiences as a research participant, please follow the link [here](#) or through a link on the OPRS website: <https://oprs.research.illinois.edu/>. You will have the option to provide feedback or concerns anonymously or you may provide your name and contact information for follow-up purposes.

**Please print this consent form if you would like to retain a copy for your records.**

**Please read each item carefully and check each box to indicate your agreement.**

- I have read and understand the above consent form.
- I understand that my speech samples are collected for the purpose of improving speech recognition technology and are protected by a legal agreement between UIUC and the current members of the Speech Accessibility Project coalition, and that future members of the data coalition and any future data licensees will also be bound by similar legal agreements.
- I certify that I am 18 years old or older.
- I understand that I must maintain a valid email address in order to participate in this study and to be paid for my participation.
- I certify that only I and the person that I am assisting will use our login to record speech samples. I will not allow other people to record speech samples on our behalf.
- I understand that the only financial compensation that I will receive for participating in this study is the payment for contributing speech samples, which will be \$30 for every batch of 200 samples that the person I am assisting provides.
- I understand that my goal is for the person I am assisting to provide 600 samples total, and that therefore my maximum payment will be \$90.
- I understand that at any time, I can end my participation and withdraw my consent. When I do so, the UIUC research team will remove all of my speech samples from the UIUC secure system. I also understand that it is UIUC's responsibility to ask all sponsors and partners to remove any copies of my speech samples from their own databases as well.

**Optional checkbox:** By checking the box below, you may also specifically grant your consent for the research team to play or provide audio files of your voice in a public research setting or publication.



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

University of Illinois Urbana-Champaign  
Online Consent Form

*I consent to have my voice samples shared in public at research conferences and similar professional meetings, and to have digital audio files of my voice be shared as part of professional research publications.*

**Optional checkbox:** By checking the box below, you may also specifically grant your consent to share your contact information with other research teams, so that they may contact you to be a potential participant in their own research studies.

*I consent to have my email address and phone number be shared with other research teams, who may contact me to participate in their studies.*

By clicking the “I CONSENT” button below, I indicate my willingness to voluntarily take part in this study.

**I CONSENT**



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

### **Speech Accessibility Project: Individuals with disabilities helping researchers to improve technology**

You are being asked to join a voluntary research study called the “Speech Accessibility Project.” This study will help computers learn to understand people with different types of speech. A computer learns to understand speech in almost the same way humans learn. It listens to speech, guesses what it says, and then guesses again whenever it makes a mistake. If the computer tries many times, then after a while it learns to guess correctly. By giving computers some examples of different speech, we can help them to better understand speech differences in the future.

Participating in this study is all online. You will record your speech online. We ask you to provide up to 600 speech samples. It should take you about 5 hours. You don't have to do it all at once. You can stop and start at any time, at your own pace and whenever you want. Risks associated with this study are no more than those in everyday life. You will not get any personal benefit from this research. Society may benefit, because computers might learn how to better understand speech differences.

**Project Name:** The Speech Accessibility Project, which has the website:  
<https://speechaccessibilityproject.beckman.illinois.edu/>

**Principal Investigator Name and Title:** Professor Mark Hasegawa-Johnson, University of Illinois Urbana-Champaign (“UIUC”), Beckman Institute and Department of Electrical and Computer Engineering.

**Contact Information:** [speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu) is the team email address.

**Organizations that will perform this study:** University of Illinois Urbana-Champaign (UIUC) is the lead organization that will perform this study.

**Sponsor:** The Speech Accessibility Project is funded by a consortium that currently includes five companies: Amazon, Apple, Google, Microsoft, and Meta (parent company of Facebook). In the future, other companies, universities, and organizations may join this consortium as well.

### **Why am I being asked?**

You are being asked to participate in this study because you are an adult, and because we think that computers should learn to understand speech like yours.

We plan to recruit more than 1,000 people to be in this study. Because the purpose of this study is to help computers learn to understand lots of different types of speech, we need to gather many different speech recordings from many types of people.

Your participation in this research is voluntary. Your decision whether or not to participate will not affect your current or future relationship with the University of Illinois, its partners, or its sponsors of this project. Even if you decide to participate, you can choose to stop at any time without affecting that relationship.

### **What will you do for this research?**

We will give you access to a website where you can record up to 600 speech recordings. For each recording, you will either read a sentence, or repeat a sentence that you hear somebody else saying, or answer a question out loud. You can do these recordings any time, whenever you want. We encourage you to record at different



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

times of day so that we can learn how your speech might change throughout the day. And don't worry about giving us "bad" samples – all data are really helpful.

We will also give you access to a mentor. Your mentor will help you do your recordings, give you advice about your microphone settings, and give you tips for when and where to do your recordings.

You may take up to three months to finish your recordings. If you will need more than three months, please tell us so by e-mail at [speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu).

We will ask you whether you consider U.S. English to be your first language. People who speak English as a second language are sometimes better recognized by computers that also know different languages. If you provide this information, it will be shared with researchers using the data.

We will also ask you to provide optional data about yourself, such as your age, race and ethnicity. We will also ask questions about how your speech ability affects your everyday life. You are not required to answer any of these questions. If you tell us these things, it will help us to know when we have collected enough different types of speech. Other people who use your speech samples will not know your answers to these questions.

We will listen to all of your speech. If you say your own name or the names of anybody close to you, we will cut those sections out of your speech recording to protect your privacy. We will also write a short list of things we think a computer should know about your speech. We will not attach your name or email address to these samples. Instead, we will assign an ID code to you, and we will label your samples with that code. Only the UIUC research team will know the "identity key" that relates your email address to your ID code.

We will store your speech in a special secure computer system at UIUC. We will store your ID on a different computer, also secure. When we give your speech to other people, we will require them to know the password before they can listen to it. They will only have access to your speech samples and they will not know your ID or anything else about you.

Right now, the people who will get copies of your speech are the five company sponsors of this project: Amazon, Apple, Google, Meta, and Microsoft. Those five companies signed a contract with the University saying that they will only use your recordings to help computers better understand speech. Some of the terms of the contract include:

- Your recordings will be kept forever, so that computers in the future can also learn to understand speech like yours.
- Your recordings are not identified with your name or email address, but only with an ID code. Only UIUC will know your name or your email address.
- UIUC will listen to your recordings before we give them to anybody else.
- Scientists might study your speech, and write papers about it, but they will not publish your speech or your identity.
- Your speech might be used to develop new computer programs, including both free programs and non-free programs that help computers to understand people with speech like yours.

In the future we might let other people use your recordings. We will not ask for your permission every



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

time we give the speech to somebody new, as long as they agree to all of the things listed above. In other words, you agree that other people can use your speech as long as they agree to the things listed above.

In summary, your speech samples will help researchers at UIUC and other places to make computers that can understand speech like yours.

### **What are the potential risks and discomforts?**

The risks are no more than what happens in everyday life. We can think of two things that you might not like about this study, so here is how we will try to avoid these risks:

First, you may get tired. To help, we will let you record your speech samples over multiple sessions. You can log in at any time, from anywhere, and record at your own pace. You can also log out at any time, then log in again later. We hope that this will help you to avoid getting tired.

Second, even though we are working hard to keep your data safe, there is a risk that somebody might steal your data from us. Here is what we have done to keep that from happening:

- Your name and e-mail are not stored with your speech. Instead, your speech is labeled using an ID number.
- Before anybody from outside UIUC can use your speech, they need to sign a contract. In the contract, they promise to store your speech securely. They agree that only approved people can use your speech. They agree that they won't try to find out who you are. They also agree that, if we tell them to destroy all of your speech recordings, they will.
- Nobody from outside UIUC will know your name, your e-mail address, or what state you live in. That information will only be stored at UIUC, on a computer that only UIUC team members can use.
- Before your speech is given to anybody else, we will listen to your speech. If you say your name or the name of anybody you know, we will cut out those parts.
- You can change your mind at any time. If you change your mind, and ask us to destroy your speech recordings, we will.

### **Are there benefits to participating in the research?**

There is no benefit to you personally.

Your recordings might help companies to make computer software that they can sell. You will not receive any extra payment if that happens.

The whole world may benefit if computers learn how to better understand people with speech like yours.

### **What other options are there?**

You have the option to not participate in this study. This study is voluntary.

### **Will my study-related information be kept confidential?**

Only people from UIUC will know your e-mail address or your name. We only need to know those things so that we can pay you for your participation in this project.

The UIUC team will also listen to all of your speech samples and edit out any personal names.



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

Faculty, staff, students, and other people from UIUC will follow the law and UIUC policies in order to protect your confidentiality.

Speech recordings will be given only to people who have signed a contract, as described above.

Because this is a research project, we plan to publish research papers about this study. Research papers will never give your name or identity. If you wish, you may choose to allow the university to share anonymous samples of your voice in research presentations, but you don't need to.

### **Will I be reimbursed for any expenses or paid for my participation in this research?**

You will be paid up to \$180. You will be paid \$60 after recording your 200<sup>th</sup> speech sample. You will be paid another \$60 after recording your 400<sup>th</sup> sample. You will get a final \$60 after recording your 600<sup>th</sup> sample.

In addition to your own payment, if somebody helps you to use your web browser for this project, they can be paid up to \$90. They will be paid \$30 after you record your 200<sup>th</sup> sample, an additional \$30 after you record your 400<sup>th</sup> sample, and a final \$30 after you record your 600<sup>th</sup> sample.

Payments will be made through email using Amazon ecodes.

### **Can I withdraw or be removed from the study?**

Yes. You can stop participating at any time, and you can tell us to delete your recordings. We also have the right to stop your participation if we think it is in your best interests. If you want us to delete your recordings, you only need to tell us that you want them deleted. If you tell us to delete your recordings, we will delete them. We will also tell your participant ID number to anybody else who has a copy of your recordings, and say that they need to delete your recordings. Nobody will ever use your recordings after that for any future research. If some computers have already learned from your recordings, it may not be possible to force the computer to unlearn what it has learned.

### **Will data collected from me be used for any other research?**

UIUC, Amazon, Apple, Google, Meta and Microsoft will use your speech recordings to make computers that better understand speech like yours.

In the future, we may let other people use your recordings, if they are willing to sign a contract. The contract will promise to use your speech only to help make better computers. The contract will also promise to protect your data. The contract will also promise that they will delete your recordings if we say that your recordings should be deleted. We will not ask you for your permission again. As long as new users of the data agree to the same terms, we will assume that you have already given permission.

### **Who should I contact if I have questions?**

Contact the UIUC research team leader, Professor Mark Hasegawa-Johnson ([jhasegaw@illinois.edu](mailto:jhasegaw@illinois.edu)), if you have any questions about this study or your part in it. You can also contact him if you have any worries or complaints about the research.

### **What are my rights as a research subject?**



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

If you have any questions about your rights as a research subject, including worries, complaints, or ideas, you may call the UIUC Office for the Protection of Research Subjects (OPRS) at 217-333-2670 or e-mail OPRS at [irb@illinois.edu](mailto:irb@illinois.edu).

If you would like to complete a brief survey to provide OPRS feedback about your experiences as a research participant, please follow the link [here](#) or through a link on the OPRS website: <https://oprs.research.illinois.edu/>. You will be able to provide feedback or concerns without giving your name, or you can give us your name and contact information.

**Please print this consent form if you would like a copy for your records.**

**Please read each item carefully and check each box to indicate your agreement.**

- I understand this consent form.
- I understand that my speech samples will be used to help computers better understand speech like mine. I understand that my speech samples are protected by a contract between UIUC and the current members of the Speech Accessibility Project. I understand that future users of the data will have to sign a similar contract before they can use the data.
- I am 18 years old or older.
- I have a valid email address that can use to send me my payments.
- I won't let other people record their speech in place of me.
- The only payment I will get is the payment for contributing speech samples, which will be \$60 for every batch of 200 samples I provide.
- My goal is to provide 600 samples total. My maximum payment will be \$180.
- If I need a caregiver to help me make recordings, I will tell UIUC, so that my caregiver can also get paid for participating.
- I can stop participating at any time. Any time I want to, I can tell UIUC to delete all of my recordings. If I tell UIUC to delete my recordings, they will delete them. They will also tell anybody else who has copies to delete them.

**Optional:** If you want to let us play your speech as part of a presentation or an online demo of the Speech Accessibility Project, you can check this box. This is not required.



University of Illinois Urbana-Champaign  
Online Consent Form



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

*I agree to let my speech be played during public presentations about the Speech Accessibility Project or as part of an online demo of the project.*

**Optional:** If you wish, we can give your name and your e-mail address or phone number to people who are planning other research studies, so you can be a participant in their studies too. This is not required.

*I agree to that my email address and phone number can be shared with other research teams, who may contact me to participate in their studies.*

By clicking the “I CONSENT” button below, I indicate my willingness to voluntarily take part in this study.

**I CONSENT**



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

### **Speech Accessibility Project: Individuals with disabilities helping researchers to improve technology**

#### **Consent Form for Recorded Caregivers**

You are being asked to participate in a voluntary research study that is called the “Speech Accessibility Project.” The purpose of this study is to help researchers at universities and companies to develop spoken-language user interfaces that work for people with atypical speech. Software programs that understand speech are developed using machine learning. Machine learning is a software development method that imitates the way humans learn, by testing the software repeatedly, and changing its behavior whenever it makes a mistake. In this way, the software learns to reduce the risk of similar mistakes in the future. By providing examples of the sounds of atypical speech, this study will make it possible for the software to learn how to make fewer mistakes when it encounters atypical speech in the future. Participating in this study is all online and will involve helping your friend or family member with a disability to record sentences and to upload the audio recordings to a secure encrypted storage system. We ask you to help your friend or family member with a disability to provide up to 600 speech samples or about 5 hours of your time which you can do any time online, at your own pace and convenience. Risks associated with this study are no greater than those in everyday life. There is no expected benefit of this research to you personally; the expected benefit to society is that companies and organizations using your speech recordings will be able to train better speech technology for people with speech and motor disabilities, which will make technology more accessible for everyone.

**Project Name:** The Speech Accessibility Project, which has the website:  
<https://speechaccessibilityproject.beckman.illinois.edu/>

**Principal Investigator Name and Title:** Professor Mark Hasegawa-Johnson, University of Illinois Urbana-Champaign (“UIUC”), Beckman Institute and Department of Electrical and Computer Engineering.

**Contact Information:** [speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu) is the team email address.

**Organizations that will perform this study:** University of Illinois Urbana-Champaign (UIUC) is the lead organization that will perform this study.

**Sponsor:** The Speech Accessibility Project is funded by a consortium that currently includes five companies: Amazon, Apple, Google, Microsoft, and Meta (parent company of Facebook). In the future, other companies, universities, and organizations may join this consortium as well.

#### **Why am I being asked?**

You are being asked to participate in this study because you are an adult who assists a person with a disability to interact with computers, and that person has already qualified to be in the study by successfully answering our pre-screening questions and providing a set of speech samples, and that person either is not able to read prompts from the screen, or does not wish to do so.

We plan to recruit more than 1,000 people to be in this study. Because the purpose of this study is to help organizations develop better speech interfaces for technologies, our team needs to gather many different types of speech samples from many types of voices.

Your participation in this research is voluntary. Your decision whether or not to participate will not affect your current or future dealings with the University of Illinois Urbana-Champaign, its partners, or its



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

sponsors of this project. If you decide to participate, you are free to withdraw at any time without affecting that relationship.

### What procedures are involved?

We will provide you with access to a website where you and the person that you are assisting can record up to 600 speech samples. Some of the prompts will instruct you to read a sentence out loud so that the person you're assisting can repeat it out loud. Some of the prompts will instruct you to ask the person you're assisting a question, so that they can speak their answer to the question. You and the person that you are assisting can do these recordings at any time, 24/7, at whatever pace you wish. We encourage you to record at different times of day so that we can learn how your speech might change throughout the day. And don't worry about giving us "bad" samples – all data are really helpful. The more diverse our data set is, the better!

You and the person that you are assisting have the option to work with UIUC to get help. UIUC can provide you with a mentor to help you do your recordings, give you advice about your microphone settings, and give you tips for when and where to do your recordings.

You may take up to three months to finish recording the 600 speech samples. If you will need more than three months, please contact the research team at [speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu).

We will ask the person you're assisting whether they consider U.S. English to be their first language. People who speak English as a second language are sometimes better recognized by devices that also know multiple languages. If you provide this information, it will be shared with researchers using the data.

We will also ask the person you're assisting to provide optional data about themselves, such as age, racial and ethnic identity, and questions about how their speech ability may have impacted their everyday life. They are not required to answer any of these questions. This information will only be used to confirm diverse representation in this data collection, and, if provided, will not be shared outside of the UIUC research team.

The UIUC research team will review all of your speech samples, and we will edit them to protect your privacy by removing any references to personally identifiable information about you or others that you might have said in your answers to our question prompts. We will also annotate your samples with "tags" that refer to technical aspects of the acoustic signals in your samples. We will not attach your name or email address to these samples; instead, we will assign a unique participant identification code to you and will annotate your samples with that code. Only the UIUC research team will know the "identity key" that relates your email address to your participant identification code.

The UIUC Speech Accessibility Project research team has established a secure, fully encrypted storage system at the UIUC Beckman Institute to store your samples. We will separately store your identifiable information in a separate secure database. We have also established a secure data sharing mechanism with the partners in our consortium. Right now, our partners are the five company sponsors of this project: Amazon, Apple, Google, Meta, and Microsoft. UIUC researchers and researchers at each of the partners will use the collected speech samples to train machine learning algorithms to perform better on speech recognition tasks. These partners will only have access to your speech samples and to no other data about you.



## University of Illinois Urbana-Champaign Online Consent Form



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

The five corporate sponsors of this research at UIUC have signed a legal agreement with the University that sets forth the permitted purposes of the speech data collected from this research project. These permitted purposes are only for research and technology development as stated here:

- Your samples will be kept forever, to assist in research and technology development in perpetuity.
- Your samples are not identified with your name or email address, but only with a unique participant code that is only known to the UIUC research team.
- Your samples will be analyzed by the UIUC research team to verify that the study is being done properly and to provide advice to other researchers about how to design or improve future studies.
- Your samples will be analyzed and combined with other people's samples to create aggregate results that will be published in the research literature.
- Your samples will be used to advance understanding of machine learning-related data variances associated with atypical speech variations, and eventually to develop products, technologies, and services by or on behalf of our sponsors to improve communications and increase access to technology, including the creation and training of machine learning solutions and artificial intelligence models for both research and commercial purposes.

In the future, any additional partner organizations in this consortium will sign data use agreements with the same terms and conditions. For these future uses of your data, we will not ask you to consent again. In other words, the future use of your de-identified data could happen without additional informed consent.

In summary, your speech samples will help researchers at UIUC and its partner organizations to develop better speech recognition technology algorithms and may eventually help commercial products perform more effectively for people with atypical speech.

### What are the potential risks and discomforts?

In general, we believe that the risks of your participation in this research study are no more than you would encounter in everyday life. However, we do recognize two issues that may be risks in this study, and we also have designed measures to minimize these risks.

First, you may get tired. To minimize your tiredness, we have arranged to let you record your speech samples over multiple sessions. You can log in to the system at a time and location convenient to you, and record at your own pace. You can also log out at any time, then log in again to continue recording as you wish. We hope that this flexibility and convenience will minimize any fatigue that you may experience.

Second, as is true for almost any online data collection research study, there is a risk that your speech samples could be released for purposes and to other individuals, companies, and organizations beyond the purposes to which you have given consent. Please know that the UIUC team has designed measures to minimize the possibility of unauthorized data release include the following:

- Participant code numbers will be used to annotate all speech recordings; your name and email address are not attached to any speech samples directly.
- All organizations using your speech recordings are required to sign a data use agreement, in which they agree that they will store the data securely, that they will prevent unauthorized access, and that they will not attempt to identify you or obtain any other identifying information about you.



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

- Researchers from outside UIUC will not know your name, e-mail address, state of residence, or any other identifying information about you. Identifying information will be stored separately from your speech recordings, in a database accessible only to UIUC research team members. In addition, as stated above, the UIUC team will edit and annotate your speech samples to protect your privacy.
- You are allowed to withdraw your voluntary consent at any time and request that all of your data be destroyed. When you notify the UIUC team that you wish to have all of your speech samples destroyed, UIUC will ensure that your speech recordings are removed from the UIUC secure system and will direct our partners that you wish your data to be removed from any copies they have as well.

### **Are there benefits to participating in the research?**

There is no immediate benefit of this research to you personally.

Information obtained from you in this research may help in the development of commercial products by UIUC, its research partners, or members of the consortium. There are no plans to provide financial compensation to you should this occur.

The expected benefit to society is that companies and organizations will be able to develop better spoken-language user interfaces so that people with atypical speech can be as easily understood as other people with typical speech.

### **What other options are there?**

You have the option to not participate in this study. This study is entirely voluntary.

### **Will my study-related information be kept confidential?**

The UIUC research team has made every effort to protect the audio records of your speech samples. The UIUC team will edit out identifying information from your speech samples (such as your name) and will only store your audio files with a participant code number. This participant code number is stored with your identifiable information on a secure server at the UIUC Beckman Institute, and this "identity key" information is only accessible by the UIUC research team. This information is necessary in order to pay you for your participation in this project.

As described previously, the UIUC team will also listen to all of your speech samples and edit out any personally identifiable information.

Faculty, staff, students, and others with permission or authority to see your study information (such as University auditors) will maintain its confidentiality to the extent permitted and required by laws and university policies.

Speech recordings will be distributed only to companies and organizations that have signed a data use agreement, agreeing to confidentiality terms specified above. No other personally identifiable information will be published or presented.

Because this is a research project, we do expect to publish research papers as a result of this study. These papers may be in the public domain. Research papers will never state your name or any identifying



## University of Illinois Urbana-Champaign Online Consent Form



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

information about you. Optionally, you may choose to allow the university to share anonymized samples of your voice as part of research publications or presentations, but this is not required.

### **Will I be reimbursed for any expenses or paid for my participation in this research?**

The person that you are assisting will be compensated up to \$180 as follows: \$60 after recording the 200<sup>th</sup> sample, an additional \$60 after recording the 400<sup>th</sup> sample, and an additional \$60 after recording the 600<sup>th</sup> sample.

In addition, since you are helping them to record their speech samples, you will also receive compensation for helping them. As a named caregiver you will be compensated with up to \$90 as follows: \$30 after they record the 200<sup>th</sup> sentence, an additional \$30 after they record the 400<sup>th</sup> sentence, and an additional \$30 after they record the 600<sup>th</sup> sentence.

Payments will be made through email using Amazon ecodes.

### **Can I withdraw or be removed from the study?**

Yes. You are free to withdraw your consent and discontinue participation at any time by notifying the UIUC research team by email or phone. The UIUC research team also have the right to stop your participation in this study without your consent if they believe it is in your best interests. If you withdraw consent, the UIUC research team will remove your speech recordings from the UIUC database and will notify the sponsors and partners that they should also remove your data from their own databases using only your anonymized participant code number. If you withdraw consent, your data will not be used for any future research. If the data you previously provided has been used to inform or create computer code or machine learning algorithms it may not be possible for your data to be removed from them.

### **Will data collected from me be used for any other research?**

As stated above, UIUC and its partners in the Speech Accessibility Project consortium will use your speech samples to advance understanding of machine learning-related data variances associated with atypical speech variations, and to develop products, technologies, and services to improve communications and increase access to technology, including computational use, and the creation and training of machine learning solutions and artificial intelligence models including for research and for commercial purposes.

It is possible that, in the future, other organizations will be granted access to the data, in addition to the founding consortium partners that are Amazon, Apple, Google, Meta, and Microsoft. These other organizations may be other universities or other companies. All organizations seeking access to the data will be required to sign data use agreements that specify the same terms as above: that your speech samples will only be used for research and technology development, that your data will be stored securely, and that you may withdraw consent at any time. You may not be asked to provide additional consent for these future uses of your data.

### **Who should I contact if I have questions?**

Contact the UIUC research team leader, Professor Mark Hasegawa-Johnson ([jhasegaw@illinois.edu](mailto:jhasegaw@illinois.edu)), if you have any questions about this study or your part in it, or if you have concerns or complaints about the research.

### **What are my rights as a research subject?**



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

## University of Illinois Urbana-Champaign Online Consent Form

If you have any questions about your rights as a research subject, including concerns, complaints, or to offer input, you may call the UIUC Office for the Protection of Research Subjects (OPRS) at 217-333-2670 or e-mail OPRS at [irb@illinois.edu](mailto:irb@illinois.edu). If you would like to complete a brief survey to provide OPRS feedback about your experiences as a research participant, please follow the link [here](#) or through a link on the OPRS website: <https://oprs.research.illinois.edu/>. You will have the option to provide feedback or concerns anonymously or you may provide your name and contact information for follow-up purposes.

**Please print this consent form if you would like to retain a copy for your records.**

**Please read each item carefully and check each box to indicate your agreement.**

- I have read and understand the above consent form.
- I understand that my speech samples are collected for the purpose of improving speech recognition technology and are protected by a legal agreement between UIUC and the current members of the Speech Accessibility Project coalition, and that future members of the data coalition and any future data licensees will also be bound by similar legal agreements.
- I certify that I am 18 years old or older.
- I understand that I must maintain a valid email address in order to participate in this study and to be paid for my participation.
- I certify that only I and the person that I am assisting will use our login to record speech samples. I will not allow other people to record speech samples on our behalf.
- I understand that the only financial compensation that I will receive for participating in this study is the payment for contributing speech samples, which will be \$30 for every batch of 200 samples that the person I am assisting provides.
- I understand that my goal is for the person I am assisting to provide 600 samples total, and that therefore my maximum payment will be \$90.
- I understand that at any time, I can end my participation and withdraw my consent. When I do so, the UIUC research team will remove all of my speech samples from the UIUC secure system. I also understand that it is UIUC's responsibility to ask all sponsors and partners to remove any copies of my speech samples from their own databases as well.

**Optional checkbox:** By checking the box below, you may also specifically grant your consent for the research team to play or provide audio files of your voice in a public research setting or publication.



IRB Number: 23183  
IRB Approval Date: 12/18/2023  
IRB Expiration Date: 03/02/2028

University of Illinois Urbana-Champaign  
Online Consent Form

*I consent to have my voice samples shared in public at research conferences and similar professional meetings, and to have digital audio files of my voice be shared as part of professional research publications.*

**Optional checkbox:** By checking the box below, you may also specifically grant your consent to share your contact information with other research teams, so that they may contact you to be a potential participant in their own research studies.

*I consent to have my email address and phone number be shared with other research teams, who may contact me to participate in their studies.*

By clicking the “I CONSENT” button below, I indicate my willingness to voluntarily take part in this study.

**I CONSENT**



Office for the Protection  
of Research Subjects

# Waiver of Documentation of Informed Consent

## For Requesting a Waiver of the Documentation of Informed Consent

All forms must be typewritten and submitted via email to [irb@illinois.edu](mailto:irb@illinois.edu).

### Section 1. PROTOCOL INFORMATION

1A. Primary Investigator:	Mark Hasegawa-Johnson
1B. Protocol Number:	23183
1C. Project Title:	Consortium NOVA HABILITAS (NOVA H)
1D. Is this research regulated by the US Food and Drug Administration?	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No

### Section 2. REQUEST FOR WAIVER OF DOCUMENTATION

A consent procedure which does not document obtained consent through a physical signature may be approved by the IRB under certain conditions. To request IRB approval of a consent procedure which does not document consent through a physical signature, provide a response to only one of the following. Note that the IRB may require the investigator to provide subjects with a written statement regarding the research, even though the documentation requirement may be waived.

**2A. The only record linking the subject and the research would be the consent document and the principal risk would be potential harm resulting from a breach of confidentiality. Each subject will be asked whether the subject wants documentation linking the subject with the research, and the subject's wishes will govern. (Note: A waiver of documentation of informed consent is not permissible under this category if the research is subject to FDA regulations.)**

**2B. The research presents no more than minimal risk of harm to subjects and involves no procedures for which written consent is normally required outside the consent.**

This research will be conducted entirely online with adult volunteer participants from all states in the USA except those that have laws restricting the use of "voice print" technologies (WA, IL, TX). Each participant will set up a login and password to the NOVA H Data Repository, will be assigned a unique user ID number, and will periodically upload speech samples via our website by speaking into their local computer's microphone and having the resulting audio file stored in our system. They will be compensated for their participation through the emailing of Amazon ecodes. After pre-screening, they will do the informed consent process online and check a series of checkboxes and a final button to provide consent to participate in the study.

**2C. The subjects or legally authorized representatives are members of a distinct cultural group or community in which signing forms is not the norm, the research presents no more than minimal risk of harm to subjects, and there is an appropriate alternative mechanism for documenting that informed consent was obtained.**



Office for the Protection  
of Research Subjects

# Certificate of Translation

## For Verifying the Translation of Research Documents

All forms must be typewritten, signed, and submitted via email to [irb@illinois.edu](mailto:irb@illinois.edu).

**When to use this form:** If research is conducted in a language other than English, submit this form with translated materials to indicate the credentials of the translator. The Certificate of Translation is required to verify that the translations are accurate. Those who translate the material are to provide a brief description of their qualifications, skills or experience for serving in this role and sign the certificate of translation form.

Please note the following:

- For research conducted in languages other than English, the University of Illinois IRB must have all versions of the research material (e.g. consents, recruitment, instruments, etc.) in both English and the language in which research is being conducted.
- It is acceptable for an investigator listed as research personnel to translate the research material if they are qualified.
- Researchers may wish to delay the initial translation until after the IRB has reviewed and approved the English versions. Doing so may help researchers avoid multiple translations.
- If the non-English documents are submitted to the IRB after initial approval, please submit an [Amendment Form](#) along with the translated material and a copy of the certificate of translation.

## Section 1. PROTOCOL INFORMATION

**1A. Principal Investigator:** Mark Hasegawa-Johnson

**1B. Protocol Number:** 23183

**1C. Project Title:** Nova Habilitas

## Section 2. TRANSLATOR

**2A. Translator's Name:** María Teresa Frontera

**2B. Translator Email Address:** maria.frontera.rodz@gmail.com

**2C. Translator's Qualifications:** Speech Language Pathologist

**2D. Language of translation:** Spanish to Spanish (minor edits)

**2E. List of document(s) translated:**

SpanishConsentForm\_MarkUp.pdf

**2F. Date(s) of translation(s):**

February 20, 2024

**2G. The translator declares that they are fluent in and understand the English language and the language of translation. The non-English documents for this study are a true and accurate translation of the English documents.** By signing below, I, the translator, agree with this statement.

Translator Signature

2/20/2024

Date



Office for the Protection  
of Research Subjects

# Certificate of Translation

## For Verifying the Translation of Research Documents

All forms must be typewritten, signed, and submitted via email to [irb@illinois.edu](mailto:irb@illinois.edu).

**When to use this form:** If research is conducted in a language other than English, submit this form with translated materials to indicate the credentials of the translator. The Certificate of Translation is required to verify that the translations are accurate. Those who translate the material are to provide a brief description of their qualifications, skills or experience for serving in this role and sign the certificate of translation form.

Please note the following:

- For research conducted in languages other than English, the University of Illinois IRB must have all versions of the research material (e.g. consents, recruitment, instruments, etc.) in both English and the language in which research is being conducted.
- It is acceptable for an investigator listed as research personnel to translate the research material if they are qualified.
- Researchers may wish to delay the initial translation until after the IRB has reviewed and approved the English versions. Doing so may help researchers avoid multiple translations.
- If the non-English documents are submitted to the IRB after initial approval, please submit an [Amendment Form](#) along with the translated material and a copy of the certificate of translation.

## Section 1. PROTOCOL INFORMATION

**1A. Principal Investigator:** Mark Hasegawa-Johnson

**1B. Protocol Number:** 23183

**1C. Project Title:** Nova Habilitas

## Section 2. TRANSLATOR

**2A. Translator's Name:** Carmen Ana Ramos-Pizarro, PhD, CCC-SLP

**2B. Translator Email Address:** Carmen.ramospizarro@udc.edu

**2C. Translator's Qualifications:** Advanced doctoral studies successfully completed in the USA. Lived in USA for the last 30 years.

**2D. Language of translation:** English to Spanish

**2E. List of document(s) translated:**

SAPRecruitment\_Parkinsons\_Spanish.pdf

**2F. Date(s) of translation(s):**

November 9, 2023

**2G. The translator declares that they are fluent in and understand the English language and the language of translation. The non-English documents for this study are a true and accurate translation of the English documents.** By signing below, I, the translator, agree with this statement.



Office for the Protection  
of Research Subjects

# Certificate of Translation

A handwritten signature in black ink, appearing to read "Cakamas".

Translator Signature

11/19/2023

Date



## Certificate of Accuracy

### Translation of **UIUC Parkinsons ConsentForm.pdf** from English to Spanish (Latin America)

As an authorized representative of GoTranscript LTD, a professional translation services agency, I, Karolis Toleikis, hereby certify that the above-mentioned document has been translated by an experienced, qualified and competent professional translator, fluent in the above-mentioned language pair and that, in my best judgement, the translated text truly reflects the content, meaning, and style of the original text and constitutes in every respect a complete and accurate translation of the original document. This document has not been translated for a family member, friend, or business associate.

This is to certify the correctness of the translation only. I do not make any claims or guarantees about the authenticity of the content of the original document. Further, GoTranscript LTD assumes no liability for the way in which the translation is used by the customer or any third party, including end-users of the translation.

A copy of this translation is attached to this certification.

Signed on **29 November 2023**

A handwritten signature in blue ink, appearing to read 'Karolis Toleikis'.

Karolis Toleikis, CEO of GoTranscript LTD

**Address:** 166, College Road, Harrow, Middlesex, HA11BH, United Kingdom

**Phone number:** +1 (831) 222-8398 **Email:** support@gotranscript.com **Website:** [gotranscript.com](http://gotranscript.com)



Office for the Protection  
of Research Subjects

# Certificate of Translation

## For Verifying the Translation of Research Documents

All forms must be typewritten, signed, and submitted via email to [irb@illinois.edu](mailto:irb@illinois.edu).

**When to use this form:** If research is conducted in a language other than English, submit this form with translated materials to indicate the credentials of the translator. The Certificate of Translation is required to verify that the translations are accurate. Those who translate the material are to provide a brief description of their qualifications, skills or experience for serving in this role and sign the certificate of translation form.

Please note the following:

- For research conducted in languages other than English, the University of Illinois IRB must have all versions of the research material (e.g. consents, recruitment, instruments, etc.) in both English and the language in which research is being conducted.
- It is acceptable for an investigator listed as research personnel to translate the research material if they are qualified.
- Researchers may wish to delay the initial translation until after the IRB has reviewed and approved the English versions. Doing so may help researchers avoid multiple translations.
- If the non-English documents are submitted to the IRB after initial approval, please submit an [Amendment Form](#) along with the translated material and a copy of the certificate of translation.

## Section 1. PROTOCOL INFORMATION

**1A. Principal Investigator:** Mark Hasegawa-Johnson

**1B. Protocol Number:** 23183

**1C. Project Title:** Nova Habilitas

## Section 2. TRANSLATOR

**2A. Translator's Name:** Carmen Ana Ramos-Pizarro, PhD, CCC-SLP

**2B. Translator Email Address:** Carmen.ramospizarro@udc.edu

**2C. Translator's Qualifications:** Advanced doctoral studies successfully completed in the USA. Lived in USA for the last 30 years.

**2D. Language of translation:** English to Spanish

**2E. List of document(s) translated:**

UIUC Parkinsons Consent Form (Spanish-latin).docx

**2F. Date(s) of translation(s):**

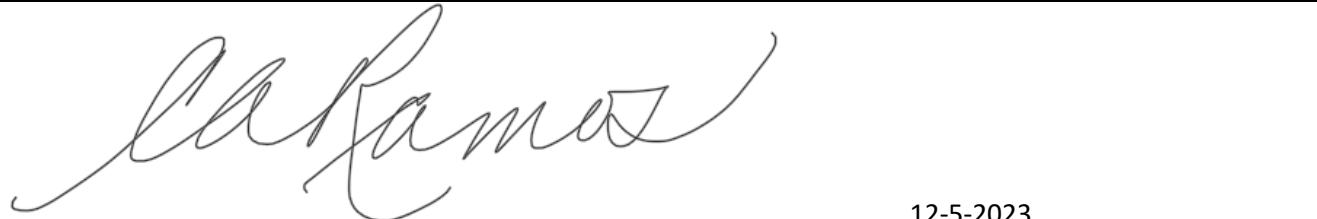
12-05-2023

**2G. The translator declares that they are fluent in and understand the English language and the language of translation. The non-English documents for this study are a true and accurate translation of the English documents.** By signing below, I, the translator, agree with this statement.



Office for the Protection  
of Research Subjects

# Certificate of Translation

 Cakamas	
Translator Signature	12-5-2023 Date

# **LSVT<sup>®</sup> GLOBAL**

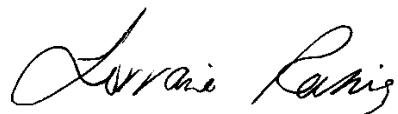
---

Professor Mark Hasegawa-Johnson  
Beckman Institute for Advanced Science and Technology  
University of Illinois Urbana-Champaign  
405 N. Mathews Ave, Urbana IL 61801

November 30, 2022

Dear Professor Hasegawa-Johnson,  
LSVT Global, Inc. is delighted to be a partner with UIUC in the Speech Accessibility Project. We look forward to collaborating on research procedures.

Sincerely,



Lorraine A. Ramig, PhD, CCC-SLP  
Chief Scientific Officer, LSVT Global, Inc.



*every victory counts®*

December 8, 2022

Mark Hasegawa-Johnson, PhD  
University of Illinois Urbana Champaign (UIUC)  
Address

**Re: Support for *Speech Accessibility Project***

Dear Professor Hasegawa-Johnson,

The Davis Phinney Foundation is delighted to partner with UIUC on the *Speech Accessibility Project*. We look forward to working with you to recruit people with Parkinson's to provide speech samples for the University's data repository and ultimately to make voice recognition technology more accessible for people living with Parkinson's. This will be a giant step forward in improving the quality of life for our Parkinson's community.

After your team at UIUC receives IRB approval and the software is functional and available, the Davis Phinney Foundation will assist in recruiting 200-300 adult volunteers living with Parkinson's to participate by providing speech samples.

We are looking forward to the opportunity to collaborate with you and your team.

Sincerely,

A handwritten signature in black ink, appearing to read 'Polly' followed by a period and a line.

Polly Dawkins

Executive Director  
720-259-0906  
[pdawkins@dpf.org](mailto:pdawkins@dpf.org)



TEAM GLEASON FOUNDATION  
Post Office Box 24493  
New Orleans, Louisiana 70184  
[www.TeamGleason.org](http://www.TeamGleason.org)  
EIN# 45-3689316

October 10, 2023

Mark Hasegawa-Johnson, PhD  
University of Illinois Urbana Champaign (UIUC)  
2011 Beckman Institute  
405 North Mathews Avenue  
Urbana, IL 61801

Re: Support for Speech Accessibility Project

Dear Professor Hasegawa-Johnson,

Team Gleason Foundation is delighted to partner with UIUC on the Speech Accessibility Project. We look forward to working with you to recruit people living with Amyotrophic Lateral Sclerosis (ALS) to provide speech samples for the University's data repository and ultimately to make voice recognition technology more accessible for people living with ALS. This will be a giant step forward in improving the quality of life for our ALS community. After your team at UIUC receives IRB approval and the software is functional and available, Team Gleason will assist in recruiting adult volunteers living with ALS to participate by providing speech samples. We are looking forward to the opportunity to collaborate with you and your team.

No White Flags,

Blair Casey  
Executive Director  
Team Gleason Foundation  
504-688-4730  
[blair@teamgleason.org](mailto:blair@teamgleason.org)



To Whom It May Concern,

I am writing this letter in support of the IRB application submitted by Professor Hasegawa-Johnson. I understand that this project will involve Lingraphica recruiting and setting up participants with aphasia to provide speech samples. Lingraphica fully supports the Speech Accessibility Project research study occurring virtually. We are happy to partner with UIUC and we're glad to be included in this project.

A handwritten signature in black ink that reads "Shawnise Carter".

---

Dr. Shawnise Carter  
Research Manager  
Speech Language Pathologist, CCC-SLP.D  
700 Alexander Park Drive, Suite 101  
Princeton, NJ 08540  
609.613.1879 – Voice/Text/Fax  
888.274.2742 X276 -Toll Free



80 Maiden Lane | 8th Floor  
New York | NY 10038  
212.683.6700  
[adaptcommunitynetwork.org](http://adaptcommunitynetwork.org)

---

March 11, 2024

Mark Hasegawa-Johnson, PhD  
University of Illinois Urbana (UIUC)  
2011 Beckman Institute  
405 North Mathews Avenue  
Urbana, IL 61801

Re: Support for Speech Accessibility Project

Dear Professor Hasegawa-Johnson:

ADAPT Community Network is delighted to partner with UIUC on the Speech Accessibility Project. We look forward to working with you to recruit people living with cerebral palsy to provide speech samples for the University's data repository. This will not only make voice recognition technology accessible for those living with cerebral palsy but will also improve their quality of life.

After your team at UIUC receives IRB approval and the software is functional and available, ADAPT will assist in recruiting adult volunteers with cerebral palsy to participate by providing speech samples.

We are looking forward to the opportunity to collaborate with you and your team.

Best,

*Roberta Koenigsberg*

Roberta Koenigsberg  
Chief Administrative Officer  
ADAPT Community Network  
80 Maiden Lane, 8<sup>th</sup> Floor  
New York, NY 10038  
[rkoenigsberg@adaptcommunitynetwork.org](mailto:rkoenigsberg@adaptcommunitynetwork.org)

# SPEECH ACCESSIBILITY PROJECT



## Share your voice with the Speech Accessibility Project!

Now recruiting individuals with Parkinson's to help speech recognition tools understand diverse voices, including those affected by Parkinson's and related neurological diagnoses, for example, MSA, PSP, CBD and post-DBS.

All personal information safeguarded at the University of Illinois Urbana-Champaign. Project funded by Amazon, Apple, Google, Meta, and Microsoft.

Record your speech from the comfort of your home. Mentors from LSVT Global are ready to help you through the participation process.

Get started today! Participants can earn up to \$180. Caregivers can earn up to \$90, as well.

**[go.illinois.edu/SAPregistration](http://go.illinois.edu/SAPregistration)**



Questions? [speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu) 1-888-309-6499



**Beckman Institute**  
UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN



# PROYECTO DE ACCESIBILIDAD DEL HABLA

## ¡Comparte tu voz con el Proyecto de Accesibilidad del Habla!

Estamos reclutando a personas con Parkinson que ayuden a que las herramientas de reconocimiento del habla puedan comprender diferentes voces, entre ellas las de personas con Parkinson y otros diagnósticos neurológicos relacionados, como MSA, LBD, PSP, CBD, y post-cirugía DBS.

Toda información personal será almacenada y protegida en la Universidad de Illinois Urbana-Champaign. Este proyecto cuenta con el respaldo financiero de las compañías Amazon, Apple, Google, Meta y Microsoft. Grabe su habla desde la comodidad de su hogar. Los mentores de LSVT Global están disponibles para ayudarle a través del proceso de participación.

¡Comience hoy mismo! Los participantes seleccionados pueden ganar hasta \$180 y sus cuidadores hasta \$90 también.

**[go.illinois.edu/SAPregistration](http://go.illinois.edu/SAPregistration)**



¿Preguntas? [speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu) 1-888-309-6499



**Beckman Institute**  
UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN



# SPEECH ACCESSIBILITY PROJECT

## Share your voice with the Speech Accessibility Project!

Now recruiting U.S. adults with Down syndrome (18+ years) to help commercial speech recognition tools understand different types of voices, including those of people with Down syndrome.

All personal information safeguarded at the University of Illinois Urbana-Champaign. Project funded by Amazon, Apple, Google, Meta, and Microsoft.

Record your speech from the comfort of your home, with caregiver support as needed.

Mentors from UIUC are ready to walk you through the participation process.

Participants can earn up to \$180. Caregivers, if needed, can earn up to \$90, as well.

[go.illinois.edu/SAPRegistration](http://go.illinois.edu/SAPRegistration)



Questions? [speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu)

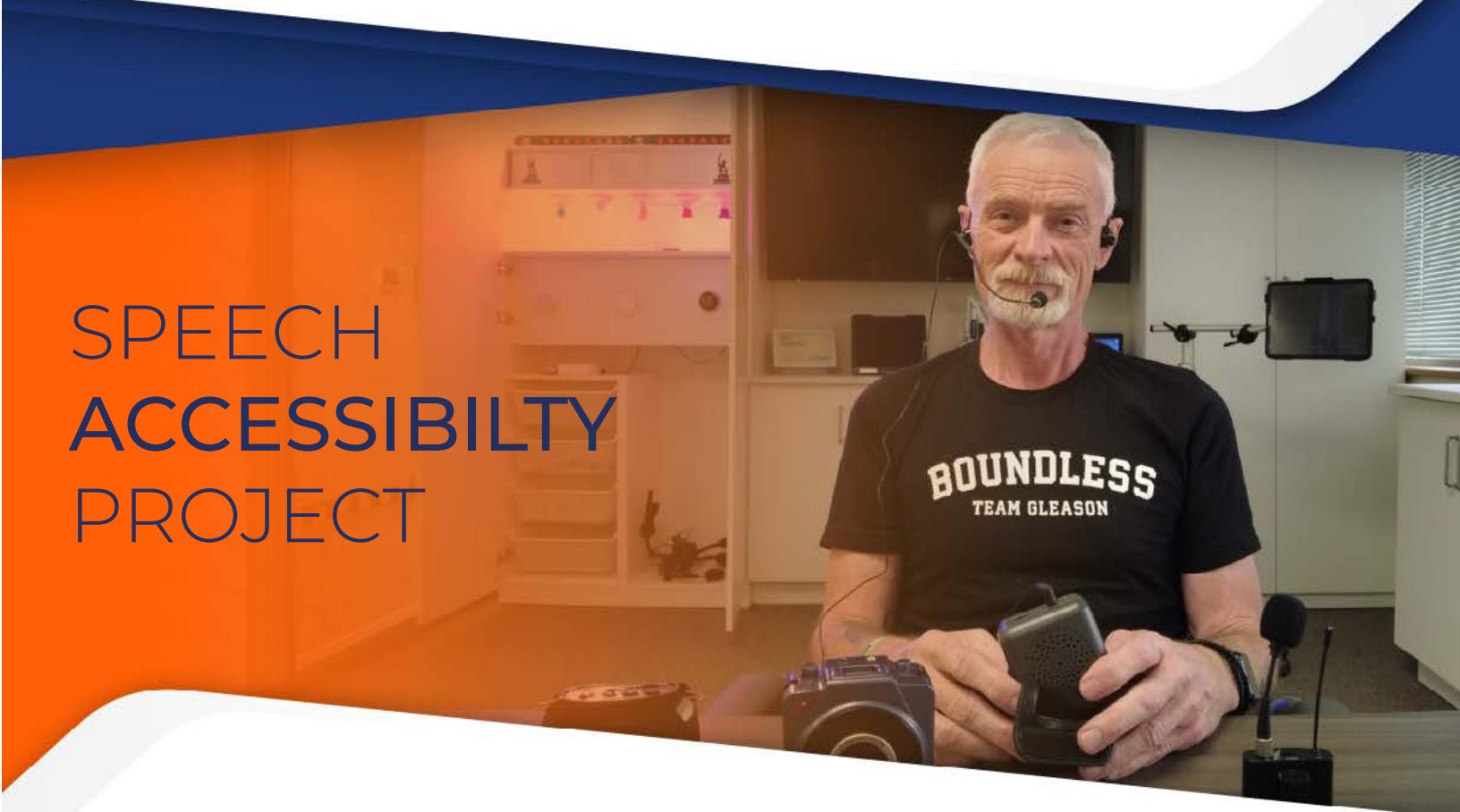
Phone #: 217-244-7700

Email: [DSSpeechAccess@illinois.edu](mailto:DSSpeechAccess@illinois.edu)



**Beckman Institute**  
UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN

# SPEECH ACCESSIBILITY PROJECT



## Share your voice with the Speech Accessibility Project!

Now recruiting people living with Amyotrophic Lateral Sclerosis (ALS) to help speech recognition tools understand diverse voices.

The University of Illinois Urbana-Champaign will safeguard your personal information as a part of this project funded by Amazon, Apple, Google, Meta, and Microsoft.

Record your speech from the comfort of your home. Mentors from Team Gleason are ready to help you through the participation process.

Participants can earn up to \$180.

Caregivers can earn up to \$90, as well.

[go.illinois.edu/SAPRegistration](http://go.illinois.edu/SAPRegistration)



Questions?

[speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu)

### PARTICIPATION CRITERIA

- Adults age 18 years or older
  - Diagnosis of ALS
  - Reads or speaks English
  - Has a valid email address
  - Excludes residents of Texas, Washington, or Illinois
- These states have privacy laws that don't allow us to collect voice prints*



**Beckman Institute**  
UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN

# SPEECH ACCESSIBILITY PROJECT



## Come be a part of this study!

Share your voice!

Help science and technology work better for people with speech challenges.

**[go.illinois.edu/SAPregistration](http://go.illinois.edu/SAPregistration)**

**WHO:** Adults of all ages, HAVE a speech and/or language impairment from stroke, CAN answer yes/no questions, CAN follow basic instructions, CAN read or repeat simple sentences, HAVE Internet access with computer or tablet, HAVE someone who can help you log on

**WHAT:** Record sentences and answer questions on our website. You will record 450 speech samples. You have up to three months to complete the recordings. It will take you about 7 hours. You can record a lot or a little at a time. You will be paid up to \$180. Your caregiver can earn up to \$90.

**WHERE:** In your own home on your computer or tablet.

**BENEFIT/ RISK:** There are no major risks or benefits for being in this study. You may get tired when you record.

Dr. Shawnise Carter, CCC-SLP, SLP.D | [scarter@lingraphica.com](mailto:scarter@lingraphica.com)



**Beckman Institute**

UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN



## Share your voice with the Speech Accessibility Project!

We're making speech recognition tools more accessible, and we need your help!

We're recording individuals with diverse voices, and we're recruiting U.S. adults who have:

- Parkinson's disease
- Down syndrome
- Amyotrophic lateral sclerosis
- Cerebral palsy
- Speaking difficulties after having a stroke

The University of Illinois Urbana-Champaign, a historic leader in accessibility, will safeguard all personal information. Amazon, Apple, Google, Meta, and Microsoft are funding this project.

Record your speech from the comfort of your home. Our mentors are ready to help you through the participation process.

Get started today! Participants can earn up to \$180. Caregivers can earn up to \$90, as well.

[go.illinois.edu/SAPregistration](http://go.illinois.edu/SAPregistration) | Questions? [speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu)  
1-888-309-6499



Beckman Institute

UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN

# SPEECH ACCESSIBILITY PROJECT



Share your voice  
with the Speech  
Accessibility Project!



Now recruiting individuals with Parkinson's to help speech recognition tools understand diverse voices, including those affected by Parkinson's and related neurological diagnoses, for example, MSA, PSP, CBD, and post-DBS.

All personal information is safeguarded at the University of Illinois Urbana-Champaign. Project funded by Amazon, Apple, Google, Meta, and Microsoft.

Record your speech from the comfort of your home. Mentors from LSVT Global are ready to help you through the participation process.

Participants can earn up to \$180. Caregivers can earn up to \$90, as well.

**SIGN UP FOR A CALL  
TO LEARN MORE!**



**GET STARTED  
TODAY!**

[go.illinois.edu/SAPregistration](http://go.illinois.edu/SAPregistration)

1-888-309-6499

**QUESTIONS?**

[speechaccessibility@beckman.illinois.edu](mailto:speechaccessibility@beckman.illinois.edu)

## **NOVA H Attachment 4: Participant Pre-Screening for Speech Samples**

*Initial pre-screening will be conducted with a combination of online and live engagement with potential participants. There are two parts to the pre-screening: Step 1 is a short questionnaire and Step 2 is providing a set of quality control speech samples. The questionnaire may be answered online by the potential participants directly (via a URL that is shared with them via email and is not public) or may be conducted as a telephone or Zoom interview with a member of the research team. Therefore, the Step 1 Pre-Screening Questionnaire shown below will be implemented both as an interactive web page and as an interview script.*

### **Step 1: Pre-Screening Questionnaire**

Thank you for your interest in the Speech Accessibility Project! To help us determine whether we can record your voice, please answer a few questions about yourself.

The information provided in these pre-screening questions will only be used to confirm your eligibility in the study and will remain confidential. This information will not be shared outside of the UIUC research team.

- 1. Do people or devices (e.g., Alexa, Siri) have a hard time understanding your speech?**  
<Answer field: Menu options Yes or No>
- 2. Have you been diagnosed with Parkinson's Disease?**  
<Answer field: Menu options Yes or No>
- 3. Are you 18 years of age or older?**  
<Answer field: Menu options Yes or No>
- 4. What state do you live in?**  
<Answer field: Drop-down menu listing all U.S. states, and the option "None of the above">
- 5. Do you have a valid email address?**  
<Answer field: Menu options Yes or No>
- 6. Do you have regular access to a web browser?**  
<Answer field: Menu options Yes or No>
- 7. Do you have the ability to read and speak English?**  
<Answer field: Menu options Yes or No>

*Note: ANSWER 1 and ANSWER 2 can be any answer; this is useful information but not screening criteria.*

IF ALL OF THE FOLLOWING CONDITIONS ARE TRUE:

ANSWER 3 = Yes

ANSWER 4 = any state except IL, TX, WA, "none of the above"

ANSWER 5 = Yes

ANSWER 6 = Yes

ANSWER 7 = Yes

THEN CONTINUING: "Thank you! So far it looks like you are eligible to be in the study."

ELSE FAILED PRE-SCREENING: "We're sorry. It looks like you are not eligible to be in our study. Thank you for your interest. Please stay in touch with us at <https://speechaccessibilityproject.beckman.illinois.edu>"

*IF CONTINUING, go to STEP 2 below. The text below is both on a web page and used as telephone/Zoom script by a member of the research team.*

### **Step 2: Quality Control Speech Samples**

Thank you! So far, it looks like you are eligible to be in our study. To confirm your final eligibility, the next step is for you to provide us some examples of your speech. We will give you access to a website where you will be prompted to record a small number of sentences (about 20 sentences). After you establish a login and password, the website will present you with a text prompt and the capability to record and save your speech samples. Your task is to make a set of voice recordings. The system will prompt you to read some text, and you will record yourself reading that text aloud. You can quit at any time and come back to this task later.

This task is performed entirely online at your own convenience. You may stop this task and restart it at any time at your own convenience, 24/7. You do not need to make all the recordings in one session; you can feel free to perform this task for just a minute or two each day spread over one week. We estimate that this task will take about 10 minutes of your time in total.

After you have submitted the full set of speech samples, the UIUC research team will evaluate the quality of your recordings. Within one week, the UIUC research team will contact you to either confirm that you may continue to participate in the study, to work with you to improve your recording environment, or to let you know that you will not be asked to participate any further in this study.

If you **are not** eligible to be in the study based on your speech samples, we will delete all of your data including your email address.

If you **are** eligible to participate in the study based on your speech samples, the speech samples provided for pre-screening will only be used to confirm your eligibility in the study, will remain confidential, and will not be shared outside of the UIUC research team.

Go to **this website (<insert URL here>)**. If you need help to provide these samples or have any questions, please contact (<insert name, email, telephone number of research team member>).

#### ***"This website" will include the following text:***

Welcome to the Speech Accessibility Project pre-screening website! At this site, you will be prompted to record a small number of sentences (about 20 sentences) that our research team will analyze in order to make sure that you are eligible to participate in our study.

First, please establish a login and password.

*<Login and password process>*

Thank you! Now, you can start recording your samples at any time. You may do these recordings at any time at your own convenience. We estimate that providing the full set of samples will take about 10 minutes total of your time. After you have submitted the full set of speech samples, the UIUC research team will evaluate the quality of your recordings. Within one week, the UIUC research team will contact you to either confirm that you may continue to participate in the study, to work with you to improve your recording environment, or to let you know that you will not be asked to participate any further in

this study. If you are not eligible to be in the study based on your speech samples, we will delete all of your data including your email address. If you have questions, please contact *<Insert name, telephone number, email address of research team member>*.

# NOVA H Attachment 7

Examples of research materials for speech sample data collection

- You will be speaking many different sentences and phrases.
- Our goal is to get many different types of speech from many different people.
- Sometimes, you will be reading aloud digital assistant commands. These are common phrases that people may ask their phone or smart speaker to help them with.
- Example: What is the weather today?

- Sometimes, you will read aloud sentences that have been randomly selected from novels.
- Example: You would never have believed that one small elbow could make such a big hole
- Some of the sentences may seem silly, or not make much sense. Please just read it as automatically as you can.

- Lastly, sometimes you will speak a little longer by answering some conversational questions.
- Example: What is a favorite tv show of yours and why?

❑ In summary, there are 3 sections.

1. 30 digital assistant commands
2. 10 sentences from novels
3. 5 open questions

# Block One

*Section 1:*

*Digital Assistant Commands*

# What song is this?

Pause the music.

# Resume.

Play music.

Stop.

Play some Beatles  
music.

Play my dance  
playlist.

Set the volume to 5.

Turn up the sound.

Skip backward  
2 minutes.

Play hip hop music.

Skip forward 30  
seconds.

# Decrease volume.

Skip this track.

Brighten the office lamp.

Turn on all switches.

Dim the living room  
lights.

Raise the  
temperature.

**Set the heat to 68.**

Turn off cooling.

Answer the door.

**Lock the front door.**

Broadcast  
‘I’m home.’

What's the  
weather like  
today?

Do I need an  
umbrella today?

# What's today's news?

Who won the NFL  
game yesterday?

What movies came  
out last week?

# Answer the call.

Hang up.

# Block One

*End of Section 1.*

*Thank You!*

# Block One

*Section 2:  
Sentences Selected from  
Novels*

# Block One

*Reminder:* Some of these sentences might not make sense. Read them as naturally as you can! It's OK to make mistakes!

I'm sure I don't know  
how it happened.

You would never have  
believed that one  
small elbow could  
make such a big hole.

But it is too late.

It's not more than half  
an hour since Uncle  
George and Aunt Bella  
went.

I'll have you ready in a  
twinkling.

But the fire and the  
children!

Father says that this  
was very wrong of  
him, and I suppose it  
was, since he says so.

But I don't see how  
Father could do  
anything wrong.

Anyway, he had a  
sister Esther whom he  
loved very much.

We had to hurry to fix  
the kite if we wanted  
to send it up before  
the wind fell.

# Block One

*End of Section 2.*

*Thank You!*

# Block One

## *Section 3: Open Questions*

# Block One

*Answer each question using  
one or more sentences.*

*Don't read the question out  
loud – just speak your  
answer!*

Tell us about one  
of your favorite  
bands or singers.

Tell us about one  
of your favorite TV  
shows.

What kind of music  
do you dislike, and  
why?

Tell us how to  
make your favorite  
sandwich.

Tell me all the steps involved in withdrawing money from a bank account, as if I had never done it before.

# Block One

*End of Section 3, and*

*End of Block One!*

*Thank You!*

## **NOVA H Attachment 9 optional metadata**

You have the option of providing extra information to help researchers interpret your speech recordings. Providing answers to such questions is not a requirement for participation in this study; if you prefer not to answer, you will be able to choose the option “Prefer not to answer.”

This information will only be used to confirm diverse representation in this data collection, and, if provided, will not be shared outside of the UIUC research team.

**1. Do you consider yourself to be a member of one of the following racial groups?**

- a. Answers: choose from
  - i. American Indian or Alaska Native,
  - ii. Asian,
  - iii. Black or African American,
  - iv. White,
  - v. Some other race (text box to allow description),
  - vi. Prefer not to answer

**2. Do you consider yourself to be Hispanic or Latino/a?**

- a. Answers: choose from Yes, No, Prefer not to answer

**3. What is your gender?**

- a. Answers: Male, Female, Non-binary, Prefer not to answer

**4. How old are you?**

- a. Answers: Age range in years (scroll menu), or Prefer not to answer

**5. What state do you live in?**

- a. Answers: Menu of US states (scroll menu), or Prefer not to answer

**6. Did you regularly converse in a language other than English before your 12<sup>th</sup> birthday?**

Answer: Yes, No, Prefer not to answer

If Yes: **what was/were the language(s)?**

Answer: Textbox

If Yes: **at what age did you start conversing in English?**

Answer: Age in years (scroll menu), or Prefer not to answer

**7. The next questions ask how easy or difficult it is for you to do certain things when you speak. Please answer as best you can. There are no wrong or right answers. It's just your opinion.**

**For each item, please rate how easy or difficult it is for you to do.**

	Very Easy	Somewhat Easy	Neither easy nor difficult	Somewhat Difficult	Very difficult	Prefer not to answer
--	--------------	------------------	----------------------------------	-----------------------	-------------------	----------------------------

Speak with a few familiar people at home (e.g. friends or family)?						
Speak with an unfamiliar person (e.g. waitstaff, store clerk)?						
Speak with a familiar person on the phone?						
Speak with a stranger over the phone?						
Speak in a noisy environment (e.g. a social gathering)?						
Speak while traveling in a car?						
Have a long conversation, an hour or longer?						
Speak when you're upset or angry?						

**8. How much do you agree or disagree with the statements below? Please answer as best you can. There are no wrong or right answers. It's just your opinion.**

	Strongly Agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree	Prefer not to answer
I am frustrated by my speech.						
I rely on others to help me communicate.						
I have to repeat myself often.						
Other people have difficulty hearing me.						
I avoid speaking when I am tired.						
My speech has impacted how often I participate in social activities.						

**9. Does your speech change throughout the day?**

- a. Answers: Yes, no unsure, prefer not to answer.
- b. If yes, please describe:
  - i. Answers: textbox