



# HHS Public Access

Author manuscript

*Behav Res Methods*. Author manuscript; available in PMC 2024 June 26.

Published in final edited form as:

*Behav Res Methods*. 2024 April ; 56(4): 2886–2900. doi:10.3758/s13428-023-02240-z.

## Toolkit to Examine Lifelike Language (TELL): An App to Capture Speech and Language Markers of Neurodegeneration

Adolfo M. García<sup>1,2,3,4</sup>, Fernando Johann<sup>2,4</sup>, Raúl Echevoyen<sup>2,4</sup>, Cecilia Calcaterra<sup>2,4</sup>, Pablo Riera<sup>5</sup>, Laouen Belloli<sup>5</sup>, Facundo Carrillo<sup>5</sup>

<sup>1</sup>Global Brain Health Institute, University of California, San Francisco, United States

<sup>2</sup>Cognitive Neuroscience Center, Universidad de San Andrés, Buenos Aires, Argentina

<sup>3</sup>Departamento de Lingüística y Literatura, Facultad de Humanidades, Universidad de Santiago de Chile, Santiago, Chile

<sup>4</sup>TELL Toolkit SA

<sup>5</sup>Instituto de Investigación en Ciencias de la Computación (ICC), CONICET-Universidad de Buenos Aires, Buenos Aires, Argentina

### Abstract

Automated speech and language analysis (ASLA) is a promising approach to capture early markers of neurodegenerative diseases. Yet, its potential remains underexploited in research and translational settings, partly due to the lack of a unified tool for data collection, encryption, processing, download, and visualization. Here we introduce the Toolkit to Examine Lifelike Language (TELL) v.1.0.0, a web-based app designed to bridge such a gap. First, we overview general aspects of its development, including safeguards for dealing with patient health information. Second, we list the steps to access and use the app. Third, we specify its data collection protocol, including a linguistic profile survey and 11 audio recording tasks. Fourth, we describe the outputs the app generates for researchers (downloadable files) and for clinicians (real-time metrics). Fifth, we survey published findings obtained through its tasks and metrics. Sixth, we refer to TELL's current limitations and prospects for expansion. Overall, with its current and planned features, TELL aims to facilitate ASLA for research and clinical aims in the neurodegeneration arena. Reviewers can gain access to TELL by visiting its website (<https://>

---

Correspondence concerning this article should be addressed to Adolfo M. García, Ph.D., Global Brain Health Institute, University of California, San Francisco; 505 Parnassus Ave, San Francisco, CA 94143, United States; [adolfo.garcia@gbhi.org](mailto:adolfo.garcia@gbhi.org).

#### Authors' Contribution

Adolfo M. García: conception, organization, figure design, writing of the first draft; Fernando Johann: review and critique; Raúl Echevoyen: review and critique; Cecilia Calcaterra: review and critique; Pablo Riera: writing of the first draft, review and critique; Laouen Belloli: writing of the first draft, review and critique; Facundo Carrillo: writing of the first draft, review and critique.

#### Competing Interests

Adolfo M. García, Fernando Johann, and Cecilia Calcaterra have received financial support from TELL SA. Raúl Echevoyen is consultant to TELL SA. Laouen Belloli, Pablo Riera, and Facundo Carrillo declare they have no financial interest.

#### Ethics Statement

No approval of research ethics committees was required to accomplish the goals of this study as it only describes software development and refers to previous literature.

#### Open Practices Statement

The materials described in the manuscript can be accessed with the credentials listed at the end of the abstract. This manuscript reports no original experiments.

tellapp.org/), clicking on the ‘Access TELL’ button on the homepage, and entering the following username (tell.examiner+BRM@gmail.com) and password (BRMreviewer1!). A demo version can be accessed here: <https://demo.sci.tellapp.org/>.

## Keywords

neurodegenerative diseases; speech and language; automated tools; multicentric research; translational science

---

## Introduction

The growth of neurodegenerative disorders is a prime public health concern. Currently incurable, Alzheimer’s disease, Parkinson’s disease, and frontotemporal dementia, among other conditions, are major causes of disability and death in high-income (Alzheimer’s Association, 2021; GBD 2017 US Neurological Disorders Collaborators, 2021) and low-income (Parra et al., 2018) countries. Their prevalence will likely double or triple by 2050 (Dorsey et al., 2018; Nichols et al., 2022), increasing caregiver burden (Cheng, 2017) and generating global costs of roughly USD 17 trillion (Nandi et al., 2022). The scenario is even direr in underserved world regions, given the dearth of equipment and staff needed for gold-standard diagnostic practices (Chávez-Fumagalli et al., 2021; Parra et al., 2018, 2023). Accordingly, a call has been raised to complement mainstream approaches with objective, affordable, and scalable digital innovations (Laske et al., 2015).

Automated speech and language analysis (ASLA) tools meet these imperatives. Participants are simply required to speak, producing acoustic (e.g., prosodic) and linguistic (e.g., semantic) features that can be automatically extracted from audio recordings and their transcriptions, respectively. The ensuing multidimensional datasets can be analyzed via inferential statistics or machine learning methods to fulfil central clinical aims (de la Fuente Garcia et al., 2020; Fraser et al., 2014, 2016; Luz et al., 2021; Nevler et al., 2019; Orimaye et al., 2017, 2018), such as differentiating among disorders (Cho et al., 2021; Sanz et al., 2022) and disease phenotypes (García et al., 2021, 2022a), predicting symptom severity (Al-Hameed et al., 2019; Eyigoz et al., 2020a; García et al., 2016) and brain atrophy patterns (Ash et al., 2013, 2019; Nevler et al., 2019), and identifying autopsy-confirmed pathology years before death (García et al., 2022b). Accordingly, ASLA has been noted as a promising resource to favor robust, equitable assessments of neurodegenerative diseases (de la Fuente Garcia et al., 2020; García et al., 2023; Laske et al., 2015).

However, the potential of ASLA remains underexploited in research and translational settings. First, investigation protocols are highly heterogeneous in terms of tasks, stimuli, procedures, and accompanying sociodemographic and linguistic profile data (Boschi et al., 2017; de la Fuente Garcia et al., 2020). Moreover, centralized systems are missing for harmonized data encryption, labeling, and storage. These issues undermine comparability and integration of findings across sites, precluding generalizability tests. Second, underlying technologies are unavailable for actual clinical use. Free resources abound for data preprocessing, analysis, and visualization, such as openSMILE (a toolkit capturing acoustic features like cepstral coefficients and formant frequencies) (Eyben et al., 2010) or

FreeLing (a language analysis tool suite providing morphological analysis and parsing functions, among others) (Padró & Stanilovsky, 2012). Nevertheless, their operation requires conditions that most clinicians lack, including software-specific knowledge, programming abilities, or substantial time for study –e.g., openSMILE and FreeLing presuppose Python and C++ skills, respectively. Finally, both research and translational efforts need to meet patient health information (PHI) standards. Briefly, optimal leveraging of ASLA for discovering and implementing novel markers requires the development of unified, user-friendly, regulation-compliant tools.

The Toolkit to Examine Lifelike Language (TELL) was developed to face these challenges. This web-based app offers an intuitive platform to collect, encrypt, analyze, download, and visualize speech and language data for research and clinical aims (Figure 1). It encompasses a linguistic profile survey and several tasks with varying motoric and cognitive demands, all proven to elicit discriminate verbal behavior in Alzheimer’s and Parkinson’s disease, among other conditions (García et al., 2021; 2022a; Sanz et al., 2022). Audio data and automated transcriptions are (i) made available for download and offline analysis, and (ii) automatically analyzed and plotted for clinical use –with diverse metrics sensitive to different brain disorders, disease presentations, and levels of severity (Al-Hameed et al., 2019; Cho et al., 2021; Eyigoz et al., 2020a; García et al., 2016, 2021, 2022b; Sanz et al., 2022). Also, TELL follows core PHI regulations, guaranteeing adequate treatment of protected health data.

In this paper, we introduce TELL v.1.0.0 and its key features, including PHI provisions. First, we outline basic aspects of the app’s development and steps for accessing it. Next, we describe the data collection protocol, which includes a linguistic profile survey and 11 audio recording tasks. We then detail TELL’s outputs, namely: downloadable files for researchers and real-time metrics for clinicians. Also, we review published research findings based on the app’s tasks and metrics. Moreover, we acknowledge its current limitations and highlight its prospects for future development. To conclude, we summarize TELL’s contributions to the global quest for early markers of neurodegenerative diseases. Reviewers can gain access to TELL by visiting its website (<https://tellapp.org/>), clicking on the ‘Access TELL’ button on the homepage, and entering the following username (tell.examiner+BRM@gmail.com) and password (BRMreviewer1!). A demo version can be accessed here: <https://demo.sci.tellapp.org/>.

## Development and General Features

TELL was developed by experts in neurolinguistics, brain health, and data science. It was built by leveraging open-access tools (e.g., Django and React) and custom code (e.g., the acoustic metrics detailed in Table 2) provided by Sigmind, a digital health service provider specialized in solutions for neuropsychiatric disorders, including computerized clinical histories and speech analytics. The app is meant to favor culturally diverse, well-powered ASLA research across several neurodegenerative disorders. Its current version (v.1.0.0) was designed to be operated by examiners (researchers or clinicians) rather than by examinees (patients, healthy controls). It is entirely web-based, thus requiring no downloads or installations. It runs on multiple devices and is optimized for Google Chrome.

The app is hosted on a cloud computing platform and is built on a REST architecture, which offers flexible resources for building distributed systems (Webber et al., 2010). Its backend consists in a Django project that uses several libraries –including Django Rest Framework, a powerful and flexible toolkit for building application programming interfaces (APIs). The backend is deployed on Amazon Web Services (AWS) ElasticBeanstalk, whose orchestration simplifies the deployment process and enables swift scalability. TELL’s intuitive, user-friendly front-end is implemented on AWS Amplify, a development platform for web applications, mobile apps, and backends.

Data storage is based on an Amazon RDS PostgreSQL database, a reliable solution for handling PHI. Audio files are stored in AWS S3, a highly secure and scalable object storage service that enables TELL to store and retrieve any amount of data from anywhere on the web. PHI is secured via multiple layers of encryption at different levels using AWS Key Management Service. This ensures that PHI is kept confidential and secure, both in storage and during transmission over networks. In this sense, TELL was built following the Health Insurance Portability and Accountability Act (HIPAA), ensuring the confidentiality, integrity, and availability of PHI through various administrative, physical, and technical safeguards. These include regular risk assessments (conducted to identify and address potential security issues) as well as data backup and disaster recovery plans (implemented to guarantee the non-vulnerability of PHI). Importantly, access to PHI is limited to authorized personnel, who must complete HIPAA compliance training and follow strict controls and authorizations in their use of the app.

TELL’s underlying modular structure allows for continual updates of its algorithms, tasks, and metrics. Front-end functions are intuitive and require minimal user intervention. These are currently available in English, Spanish, Portuguese, and French, with new languages being added based on project-specific requirements. Its tasks and metrics have been validated by our team and other laboratories, in populations from both high-income countries (e.g., United States, Germany, Czech Republic) and low/middle-income countries (e.g., Argentina, Chile, Colombia) (Eyigoz et al., 2020a; García et al., 2016, 2021, 2022a, 2022b; Sanz et al., 2022). Access, details, news, and publications are provided on the app’s website (<https://tellapp.org/>). Key functions of its current version are detailed below.

## Accessing the App

Access to TELL is requested through the website’s contact form (<https://tellapp.org/contact/>). This guarantees the efficiency of all functions, which run on paid online services with a usage cap. Users simply need to contact the TELL team, describe the project in which they will use the app, specify how many sessions will be required, and commit to open science practices for collaborative research. They will then receive an e-mail with a username and link to activate their account and create a personalized password. The app can then be accessed by clicking on the “Access TELL” button on the website’s homepage (<https://tellapp.org/>). See Figure 2.

We note that TELL is provided by a commercial company (TELL Toolkit SA), as required to manage the funds secured for its initial development. Yet, ours is a basic and translational

research initiative, offered at no cost for scientific investigation (free research licenses are issued with no enrolment or subscription fees). Although commercial exploits are beyond TELL's central scientific motivation, funds will be pursued to cover the costs of its underlying technologies (Amazon Web Services, Google resources, and an Open AI API with a usage cap) and to support its further development. Indeed, it is because of such costs that the tool requires setting up accounts (otherwise, we would face the risk of financial unviability due to more requests than we can pay for). This simple process allows us to ensure that TELL users will have optimal access to all functions for the entirety of their projects.

## Data Collection

### Creating Participants

Upon accessing TELL, users can create new participants by entering their names or identification codes (Figure 3). Participant information can be entered in specific fields, and relevant files (e.g., neuropsychological tests) can be uploaded as attachments for future reference. Participants are listed alphabetically and they can be browsed in a search field. Each participant name is accompanied by two icons, granting access to a linguistic profile survey and the speech recording protocol.

### Linguistic Profile Survey

TELL's linguistic profile survey (Figure 4) is an adaptation and extension of the Bilingual Language Profile (Faroqi-Shah et al., 2018; Gertken et al., 2014). The survey comprises 18 items tapping on four dimensions, namely: *language history* (e.g., languages used, schooling in each), *use* (e.g., time using each language), *proficiency* (in speaking, understanding, writing, and reading), and *perceived difficulties* (e.g., age of speech difficulty onset). Depending on the item, responses are entered in open-ended fields, through multiple-option menus or via Likert scales. Once the survey is complete, a .csv file is automatically produced with all responses presented in successive columns. The survey is currently available in English, Spanish, Portuguese, and French. Administration time ranges from 5 to 10 minutes, depending on the participant's condition and disease severity.

### Speech Recording Protocol

Users can access the speech recording protocol in each available language. Since acquisition conditions may vary widely across settings, the protocol begins with recommendations that follow and extend recent guidelines (Busquet et al., 2023) to optimize data quality and comparability across locations:

1. Sit in a quiet room with no distractions, facing the microphone head on.
2. Mute all phones and remove all noise sources from the room.
3. If available, use a condenser unidirectional microphone with pop filters.
4. Speak close to the microphone, keeping the same distance across tasks.
5. If using a headset, make sure the microphone does not touch your face or facial hair.

6. Remove any jewelry or accessories that might make noise as you speak.
7. Speak with your normal pace and volume.

Upon reading these instructions, users must grant permission for the app to access the device's microphone. The protocol comprises successive speech tasks. They all follow the same structure, with a title and instructions on top, a stimulus in the middle (when required), and a 'record' button at the bottom. Examiners must read the instructions to the participants, ensure that they have understood their indications, and then press 'record'. For most tasks, a minimum of 30 seconds is required before the recording can be interrupted, with no upper bound. Once the participant finishes the task, the examiner can save the recording by pressing the 'record' button again, which also triggers the following task. Tasks can be skipped at the examiner's discretion.

TELL comprises 11 tasks, which can be used or skipped at will (Table 1). Two *spontaneous speech tasks* require enrollees to describe (a) a typical day and (b) a pleasant memory, yielding varied acoustic and linguistic patterns based on the participant's individual experience (Boschi et al., 2017). *Semi-spontaneous speech tasks* involve describing (c) pictures of daily scenes and narrating (d) an audio story as well as (e) a silent animated film –eliciting partly predictable patterns (Boschi et al., 2017), with higher executive/memory demands for the narration tasks. *Non-spontaneous speech tasks* require (f) reading a paragraph (which restricts acoustic and linguistic variation and lowers executive demands) (Rusz et al., 2013), (g) repeating the syllables /pataka/ (to assess diadochokinetic skills), and (h) uttering the vowel /a/ until breath runs out (to test for vocal stability). Two *complementary tasks* have participants (i) speak about the uses of a wheel-less car (to assess creative thinking) and (j) describe their experience of the disease (to identify condition-specific patterns in their account of daily functionality). Finally, participants are asked to (k) record 15 seconds of silence, to capture background noise for removal from other tasks. All these tasks have been reported in studies on diseases as diverse as Alzheimer's disease and Parkinson's disease (García et al., 2021, 2022b; Sanz et al., 2022). Yet, users can choose to administer all of them or any specific subset depending on their goals and hypotheses.

Except for reading, which uses specific texts for each language, every task employs the same cross-culturally adequate stimuli for all currently available languages (English, Spanish, Portuguese, and French). Depending on the participant's condition and disease severity, the full protocol takes between 15 and 25 minutes to complete. Figure 5 illustrates the speech recording interface based on the video retelling task.

Audio files are recorded as Wave (.wav) files, a single-channel format with an ideal sampling rate of 44.1 Hz and an ideal bit-depth of 24 bits –two conditions that are supported by most modern devices. While recordings' characteristics can vary depending on factors like microphone quality and environmental conditions, .wav files are uncompressed, meaning that no information is lost during the coding process. Such files are temporarily stored in virtual memory before being sent to the TELL backend. Once encrypted and transmitted, audios (and, with them, all unencrypted PHI) are automatically deleted from the device and stored as a highly durable object in an encrypted S3 bucket in the backend (encryption ensures that audio files remain confidential and protected from unauthorized

access). Additionally, TELL regularly performs backups of its S3 buckets to ensure that all data is recoverable in the event of malfunction or data loss. Audio-based transcriptions are made through Whisper (v. 9.67.1), Open AI's cutting-edge speech recognition service (Radford et al., 2023). Whisper employs a transformer neural network model to convert audio signals into written words, yielding high transcription accuracy in varying recording conditions. The resulting texts can be subject to multiple preprocessing and processing steps, depending on research needs. Whisper code is hosted on our own servers.

## Audio preprocessing

The audio preprocessing pipeline consists of four steps: (1) common channel normalization, (2) loudness normalization, (3) denoising, and (4) voice activity detection (VAD).

Common channel normalization is performed with the ExSound eXchange library (<https://sox.sourceforge.net/>) to standardize different audio acquisition conditions. The resulting output serves as a lower bound of acceptable audio quality. Normalization consists of the following steps: conversion to .wav format, conversion to mono (single channel), encoding in GSM FULL RATE with a bit rate of 13 kbps, resampling to 16 kHz (Favaro et al., 2023; Haulcy & Glass, 2020), and compression with a factor of 8 and bandpass filtering from 200 Hz to 3.4 kHz (Arias-Vergara et al., 2019). The bandpass filter emulates the characteristics of narrowband transmissions, which is sufficient to represent the main harmonics of human speech (Cox et al., 2009). Since Python cannot handle 13-bit audio, files are decoded back to 16 bits using ExSound eXchange.

Loudness normalization is performed to ensure that all audio files have equivalent sound intensities, compensating for variations caused during signal acquisition, such as the position of the recording device relative to the interviewee (Haulcy & Glass, 2020; Luz et al., 2020). A normalization filter from the FFmpeg multimedia framework (<https://ffmpeg.org/ffmpeg.html>) is used, following the EBU R128 standard (Favaro et al., 2023).

Denoising (background noise reduction) is applied to significantly decrease the level of any external noise present in the recording (Luz et al., 2020). This is achieved through a speech enhancement signal optimization model that combines two sequentially processed models: one that operates on full-band spectra and another that processes the audio in sub-bands. This approach leverages and integrates the advantages of both methods (Hao, 2021). The complete model, known as FullSubNet (<https://github.com/Audio-WestlakeU/FullSubNet>), combines the full-band and sub-band models and has been pretrained using neural networks with datasets containing "clean" speech (free from background noise) and datasets of dynamically mixed noise signals (Li et al., 2023).

The VAD calculates the probability that a given audio segment corresponds to human speech. This is achieved by separating audio into segments that contain speech and segments that contain silence (or non-speech) (Haulcy & Glass, 2020; Hecker et al., 2022). Additionally, it enables the analysis of different metrics derived from these two types of segments, such as the average duration of voiced segments and the number of pauses, among others. The Silero VAD model (<https://github.com/snakers4/silero-vad>) is used for this segmentation process.

## Output for research and clinical purposes

Each task includes calculations of audio-derived features (including timing and pitch-related features) and transcript-derived features (predominant words, affective valence, emotional magnitude). These features are listed in Table 2. Together with the survey responses, these make up the full set of data offered by the app. The source code for all features is available at [https://github.com/telltoolkit/TELL\\_source\\_code](https://github.com/telltoolkit/TELL_source_code)

TELL v.1.0.0 includes three timing features: speech rate (words per minute), silence duration (median of the duration of all silence segments), and silent pauses per word. Speech and silence segments are established through a voice activity detector running on pyannote 1.1.0 (Bredin et al., 2020). This algorithm is based on a Sinc + LSTM neural network whose output indicates the probability that an audio segment includes a speech signal. The output is binarized to determine which portions of the recording are phonated or silent utilizing onset and offset probability thresholds and values of segments' minimum durations. Speech rate is computed as the number of words in the transcription divided by the total speech time in minutes (upon removal of silent segments). Silence duration is the mean of the silence segments. The calculation of silences per word is computed as the number of silent segments divided by the number of words.

The app also calculates recordings' main pitch and pitch variations. Pitch is computed with openSMILE using eGeMAPSv02 recipe via sub-harmonic-sampling and Viterbi smoothing (Eyben et al., 2010). The signal has a sampling rate of 100 Hz and takes minimum and maximum values of 30 and 500 Hz, respectively. Main pitch and pitch variations consist in the mean and standard deviation (in Hz) of all voiced segments, respectively. The pitch time series is conserved in the data pipeline if needed.

On-the-fly computations are also performed on the audio transcripts. Predominant words in a transcription are displayed through a word cloud, in which each item's size is proportional to its frequency. Our pipeline captures all words, irrespective of their category, and retains their specific form without lemmatization. Affective valence and emotional magnitude are measured via analyzeSentiment (<https://cloud.google.com/natural-language/docs/analyzing-sentiment?hl=en>), a sentiment analysis algorithm within Google's Cloud Natural Language API. This multilingual tool was trained on a large collection of texts from the General Inquirer corpus as well as corpora with manually labeled sentiment scores, such the SentiWordNet lexicon, the Amazon product reviews corpus, and the Stanford Sentiment Treebank. Affective valence and emotional magnitude are derived for any input text through term frequency-inverse document frequency (a statistical measure that quantifies the importance of a word in a document) combined with part-of-speech tagging and a module to recognize words with strong sentiment. Based on this information, the affective valence metric quantifies each text's emotional content on a continuous scale from -1 to 1, texts being mostly negative if close to -1, neutral if close to 0, and mostly positive if close to 1. For its part, the emotional magnitude metric provides the absolute value of the amount of emotion expressed in the text, regardless of its polarity. Texts with a value close to 0 express no emotional content, with increasingly higher values indicating greater emotionality.



TELL allows users to download data for offline research purposes (Figure 6). All responses to the linguistic profile survey are made available in a .csv file. Then, data from the speech protocol can be downloaded for each task separately or as a bulk for the whole protocol. This includes (i) raw audio recordings (.wav files, as detailed above), automated transcriptions (via Open AI's Whisper), and a .csv file containing basic speech and language metrics (described in the following section). Files are anonymized and labeled with a standardized name for easy access (e.g., participant\_code\_story\_retelling.wav). Note that .wav files are lossless (surpassing the quality of compressed formats, such as .mp3) and supported by a vast array of hardware and software options.

Moreover, metrics are graphically depicted for quick, intuitive revision by examiners (Figure 7), and they are accompanied by brief and extended explanations, suggested papers, and interpretive notes based on descriptions of benchmark control data. Acoustic measures further allow tracking longitudinal changes through charts that plot variations in time. Note that many of these metrics are not available for the syllable repetition and the sustained vowel tasks, as they do not involve actual linguistic production. Also, no metric is available for the silence recoding task, intended solely for offline detection of acoustic conditions in the recording environment.

## Applications and Validations

The metrics included in TELL have been validated through multiple studies from our team and others. For example, speech timing features, including speech rate, discriminate persons with nonfluent and logopenic variant primary progressive aphasia from healthy persons (Ballard et al., 2014; Cordella et al., 2019; Nevler et al., 2019). Moreover, this feature distinguishes nonfluent from semantic variant primary progressive aphasia patients, and even predicts the former's underlying pathology several years before death (García et al., 2022b). Also, speech rate correlates with the volume of the hippocampus and other core atrophy regions in Alzheimer's disease (Jonell et al., 2021; Riley et al., 2005), and with superior/inferior frontal atrophy in frontotemporal dementia (Ballard et al., 2014; Cordella et al., 2019; Nevler et al., 2019). Also, abnormalities of pitch and pause are central for identifying persons with Parkinson's disease and discriminating between those with and without mild cognitive impairment (García et al., 2021). Pitch-related alterations, such as narrowed range of fundamental frequency, differentiate nonfluent variant primary progressive aphasia patients from both healthy controls and semantic variant primary progressive aphasia patients, while also predicting left inferior frontal atrophy (the neuroanatomical epicenter of the syndrome) (Nevler et al., 2019). Reduced pitch is also sensitive to behavioral variant frontotemporal dementia in early disease stages (Nevler et al., 2017). In addition, reduced word counts are key contributors to machine learning models for predicting Alzheimer's disease onset (Eyigoz et al., 2020b), identifying persons at risk for this disorder (Fraser et al., 2019), and differentiating between Alzheimer's disease patients with and without vascular pathology (Rentoumi et al., 2014).

Of note, identification of neurodegenerative diseases becomes even higher (often surpassing 90% accuracy) when multiple acoustic and linguistic features are combined rather than analyzed in isolation (Fraser et al., 2014; Nevler et al., 2019; Zimmerer et al., 2020). Beyond

the metrics currently embedded in TELL, additional developments in the field of ASLA have proven useful for characterizing, differentiating, phenotyping, and monitoring diverse neurodegenerative disorders, paving the way for powerful updates (Al-Hameed et al., 2019; Ash et al., 2013, 2019; Ballard et al., 2014; Cho et al., 2021; Cordella et al., 2019; Eyigoz et al., 2020a; Ferrante et al., accepted; García et al., 2016, 2021, 2022a; Faroqi-Shah et al., 2020; Fraser et al., 2014; Nevler et al., 2017, 2019; Norel et al., 2020; Ruzs & Tykalová, 2021; Themistocleous et al., 2021).

Note that TELL, like ASLA tools in general, does not and cannot replace mainstream diagnostic or prognostic procedures. The contribution of speech and language metrics lies in their capacity to reveal markers that can be combined with others to favor early detection of disease-related changes. This is a key objective in the global fight against neurodegeneration, given that timely detection can reduce affective burden, increase planning time for neuroprotective changes (Dubois et al., 2016; Isaacson et al., 2018, 2019), optimize pathology-targeted therapies (several compounds are in phase 2/3 trials) (Cummings et al., 2020), and generate financial savings (fostering routine over emergency care) (Alzheimer's Association, 2021; GBD 2017 US Neurological Disorders Collaborators, 2021). In addition to its research applications, then, TELL can support relevant actions in the clinical arena.

## Current Limitations and Future Developments

TELL cannot be reliably used with browsers other than Google Chrome. Moreover, some of its metrics require multiple recordings to generate visual results. These issues will be addressed in future versions. Additional developments include the implementation of novel fine-grained metrics validated in recent research and superior speech-to-text algorithms (indeed, Whisper is not trained on patient populations and transcription errors are common for persons with motor speech difficulties, non-Anglophone individuals, and speakers of particular dialects). We are also planning to streamline our front end and to include pre-trained models for classifying participants into healthy and pathological neurocognitive profiles. Furthermore, new technologies will be explored to replace TELL's current paid underlying technologies with open-source tools. This will allow for more direct access to the tool while enabling academic users to work directly with experimental code within the platform (e.g., via a plugin architecture or other alternatives).

Further challenges concern the potential biases introduced by variable recording conditions (ranging from varying devices and noise sources to inconsistent mouth-to-microphone proximity) (Busquet et al., 2023). These can be minimized through machine learning models for bias correction (Busquet et al., 2023) or with state-of-the-art preprocessing pipelines, including common resampling, amplitude normalization, denoising, and channel standardization, as reported in recent works (García et al., 2021; Pérez-Toro et al., 2022). Future versions of TELL will incorporate these steps into its built-in audio processing algorithms to enhance comparability despite inconsistent acquisition settings. These developments, in addition to the recommendations in the "Speech Recording Protocol" section and strategic items included in the app's survey (questions on device type, use of

face mask, smoking habits, and throat surgery), should enable strict controls of variance sources for multicentric studies.

Another limitation relates to longitudinal tracking. Some of the app's tasks may be efficiently reused for specific measures across time points (e.g., speech rate in the paragraph reading task). Yet, data from other combinations of tasks and metrics may be affected by learning or familiarization effects (e.g., counts of predominant words in picture description tasks). This could be mitigated by including several randomizable materials (e.g., pictures) with consistent core properties (e.g., emotionality), as proposed by the Ghent Semi-spontaneous Speech Paradigm (Van Der Donckt et al., 2023). We will initiate stimulus validation efforts for future versions of TELL to include multiple, well controlled, randomizable stimuli for at least some of its tasks (e.g., picture description, video narration). This will increase data reliability for longitudinal research and patient monitoring.

Finally, while TELL is initially focused on neurodegenerative disorders, we are aware that ASLA can capture powerful markers in other populations. For example, speech and language markers allow classifying persons with different psychiatric conditions (Agurto et al., 2020; Bedi et al., 2014; Mota et al., 2012, 2014) as well as tracking the severity (Mota et al., 2017) and predicting the onset (Bedi et al., 2015) of psychotic symptoms. Also, specific features distinguish between children with and without neurodevelopmental conditions, such autism spectrum disorder (MacFarlane et al., 2022). Future versions of TELL will incorporate tasks and metrics tailored to these groups, expanding its reach for both academic and clinical purposes.

## Concluding Remarks

The plea for scalable and equitable cognitive markers is central in the agenda of neurodegeneration research. Digital tools, thus, become critical for the generation and application of relevant discoveries. TELL aims to support these missions by facilitating ASLA to scientists and clinicians. Its current version, as described here, is our first (and certainly not our last) step in this direction.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Funding

Adolfo García is an Atlantic Fellow at the Global Brain Health Institute (GBHI) and is partially supported by the National Institute On Aging of the National Institutes of Health (R01AG075775); ANID (FONDECYT Regular 1210176, 1210195); GBHI, Alzheimer's Association, and Alzheimer's Society (Alzheimer's Association GBHI ALZ UK-22-865742); Universidad de Santiago de Chile (DICYT 032351GA\_DAS); and the Multi-partner Consortium to Expand Dementia Research in Latin America (ReDLat), which is supported by the Fogarty International Center and the National Institutes of Health, the National Institutes of Aging (R01AG057234, R01AG075775, R01AG21051, and CARDS-NIH), Alzheimer's Association (SG-20-725707), Rainwater Charitable Foundation's Tau Consortium, the Bluefield Project to Cure Frontotemporal Dementia and the Global Brain Health Institute. The contents of this publication are solely the responsibility of the authors and do not represent the official views of these institutions.

## References

- Agurto C, Pietrowicz M, Norel R, Eyigoz EK, Stanislawski E, Cecchi G, & Corcoran C (2020). Analyzing acoustic and prosodic fluctuations in free speech to predict psychosis onset in high-risk youths. *Annu Int Conf IEEE Eng Med Biol Soc*, 2020, 5575–5579. 10.1109/embc44109.2020.9176841
- Ahmed S, Haigh AM, de Jager CA, & Garrard P (2013). Connected speech as a marker of disease progression in autopsy-proven Alzheimer's disease. *Brain*, 136(Pt 12), 3727–3737. 10.1093/brain/awt269 [PubMed: 24142144]
- Al-Hameed S, Benaissa M, Christensen H, Mirheidari B, Blackburn D, & Reuber M (2019). A new diagnostic approach for the identification of patients with neurodegenerative cognitive complaints. *PLOS ONE*, 14(5), e0217388. 10.1371/journal.pone.0217388 [PubMed: 31125389]
- Alzheimer's Association (2021). 2021 Alzheimer's disease facts and figures. *Alzheimers Dement*, 17(3), 327–406. 10.1002/alz.12328 [PubMed: 33756057]
- Arias-Vergara T, Vasquez-Correa JC, Gollwitzer S, Orozco-Arroyave JR, Schuster M, Nöth E (2019). Multi-channel convolutional neural networks for automatic detection of speech deficits in cochlear implant users. In Nyström I, Hernández Heredia Y, Milián Núñez V (eds.), *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications. CIARP 2019. Lecture Notes in Computer Science*, 11896. Springer, Cham. 10.1007/978-3-030-33904-3\_64
- Ash S, Evans E, O'Shea J, Powers J, Boller A, Weinberg D, ... Grossman M. (2013). Differentiating primary progressive aphasia in a brief sample of connected speech. *Neurology*, 81(4), 329–336. 10.1212/WNL.0b013e31829c5d0e [PubMed: 23794681]
- Ash S, Nevler N, Phillips J, Irwin DJ, McMillan CT, Rascovsky K, & Grossman M (2019). A longitudinal study of speech production in primary progressive aphasia and behavioral variant frontotemporal dementia. *Brain Lang*, 194, 46–57. 10.1016/j.bandl.2019.04.006 [PubMed: 31075725]
- Ballard KJ, Savage S, Leyton CE, Vogel AP, Hornberger M, & Hodges JR (2014). Logopenic and Nonfluent Variants of Primary Progressive Aphasia Are Differentiated by Acoustic Measures of Speech Production. *PLOS ONE*, 9(2), e89864. 10.1371/journal.pone.0089864 [PubMed: 24587083]
- Bedi G, Carrillo F, Cecchi GA, Slezak DF, Sigman M, Mota NB, ... Corcoran CM (2015). Automated analysis of free speech predicts psychosis onset in high-risk youths. *npj Schizophrenia*, 1(1), 15030. 10.1038/npjshz.2015.30 [PubMed: 27336038]
- Bedi G, Cecchi GA, Slezak DF, Carrillo F, Sigman M, & de Wit H (2014). A window into the intoxicated mind? Speech as an index of psychoactive drug effects. *Neuropsychopharmacology*, 39(10), 2340–2348. 10.1038/npp.2014.80 [PubMed: 24694926]
- Boschi V, Catricalà E, Consonni M, Chesi C, Moro A, & Cappa SF (2017). Connected speech in neurodegenerative language disorders: A Review. *Frontiers in Psychology*, 8(269). 10.3389/fpsyg.2017.00269
- Bowen LK, Hands GL, Pradhan S, & Stepp CE (2013). Effects of Parkinson's disease on fundamental frequency variability in running speech. *J Med Speech Lang Pathol*, 21(3), 235–244. [PubMed: 25838754]
- Bredin H, Yin R, Coria JM, Gelly G, Korshunov P, Lavechin M, ... Gill MP (2020). Pyannote.Audio: Neural Building Blocks for Speaker Diarization. 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain.
- Busquet F, Efthymiou F, & Hildebrand C (2023). Voice analytics in the wild: Validity and predictive accuracy of common audio-recording devices. *Behavior Research Methods*. 10.3758/s13428-023-02139-9
- Carrillo F, Sigman M, Fernández Slezak D, Ashton P, Fitzgerald L, Stroud J, ... Carhart-Harris RL (2018). Natural speech algorithm applied to baseline interview data can predict which patients will respond to psilocybin for treatment-resistant depression. *J Affect Disord*, 230, 84–86. 10.1016/j.jad.2018.01.006 [PubMed: 29407543]
- Chávez-Fumagalli MA, Shrivastava P, Aguilar-Pineda JA, Nieto-Montesinos R, Del-Carpio GD, Peralta-Mestas A, ... Lino Cardenas CL (2021). Diagnosis of Alzheimer's disease in developed

- and developing countries: Systematic review and meta-analysis of diagnostic test accuracy. *J Alzheimers Dis Rep*, 5(1), 15–30. 10.3233/adr-200263 [PubMed: 33681713]
- Cheng ST (2017). Dementia Caregiver Burden: a Research Update and Critical Analysis. *Curr Psychiatry Rep*, 19(9), 64. 10.1007/s11920-017-0818-2 [PubMed: 28795386]
- Cho S, Nevler N, Ash S, Shellikeri S, Irwin DJ, Massimo L, ... Liberman M (2021). Automated analysis of lexical features in frontotemporal degeneration. *Cortex*, 137, 215–231. 10.1016/j.cortex.2021.01.012 [PubMed: 33640853]
- Cho S, Shellikeri S, Ash S, Liberman MY, Grossman M, Nevler N, & Nevler N (2021). Automatic classification of AD versus FTLN pathology using speech analysis in a biologically confirmed cohort. *Alzheimer's & Dementia*, 17(S5), e052270. <https://doi.org/10.1002/alz.052270>
- Cordella C, Quimby M, Touroutoglou A, Brickhouse M, Dickerson BC, & Green JR (2019). Quantification of motor speech impairment and its anatomic basis in primary progressive aphasia. *Neurology*, 92(17), e1992–e2004. 10.1212/wnl.00000000000007367 [PubMed: 30944238]
- Cox RV, Neto SFDC, Lamblin C, & Sherif MH (2009). ITU-T coders for wideband, superwideband, and fullband speech communication. *IEEE Communications Magazine*, 47(10), 106–109. 10.1109/MCOM.2009.5273816
- Cummings J, Lee G, Ritter A, Sabbagh M, & Zhong K (2020). Alzheimer's disease drug development pipeline: 2020. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 6(1), e12050. <https://doi.org/10.1002/trc2.12050> [PubMed: 32695874]
- de la Fuente García S, Ritchie CW, & Luz S (2020). Artificial intelligence, speech, and language processing approaches to monitoring Alzheimer's disease: A systematic review. *Journal of Alzheimer's Disease*, 78, 1547–1574. 10.3233/JAD-200888
- Dodge HH, Mattek N, Gregor M, Bowman M, Seelye A, Ybarra O, ... Kaye JA (2015). Social markers of mild cognitive impairment: Proportion of word counts in free conversational speech. *Curr Alzheimer Res*, 12(6), 513–519. 10.2174/1567205012666150530201917 [PubMed: 26027814]
- Dorsey ER, Elbaz A, Nichols E, Abd-Allah F, Abdelalim A, Adusar JC, ... Murray CJL (2018). Global, regional, and national burden of Parkinson's disease, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Neurology*, 17(11), 939–953. 10.1016/S1474-4422(18)30295-3 [PubMed: 30287051]
- Dubois B, Padovani A, Scheltens P, Rossi A, & Dell'Agnello G (2016). Timely diagnosis for Alzheimer's disease: A literature review on benefits and challenges. *Journal of Alzheimer's Disease*, 49, 617–631. 10.3233/JAD-150692
- Ellis C, Holt YF, & West T (2015). Lexical diversity in Parkinson's disease. *Journal of Clinical Movement Disorders*, 2(1), 5. 10.1186/s40734-015-0017-4 [PubMed: 26788341]
- Eyben F, Wöllmer M, & Schuller B (2010). Opensmile: the munich versatile and fast open-source audio feature extractor. *Proceedings of the 18th ACM international conference on Multimedia, Firenze, Italy*. 10.1145/1873951.1874246
- Eyigoz E, Courson M, Sedeño L, Rogg K, Orozco-Arroyave JR, Nöth E, ... García AM (2020a). From discourse to pathology: Automatic identification of Parkinson's disease patients via morphological measures across three languages. *Cortex*, 132, 191–205. <https://doi.org/10.1016/j.cortex.2020.08.020> [PubMed: 32992069]
- Eyigoz E, Mathur S, Santamaria M, Cecchi G, & Naylor M (2020b). Linguistic markers predict onset of Alzheimer's disease. *eClinicalMedicine*. 10.1016/j.eclinm.2020.100583
- Faroqi-Shah Y, Sampson M, Pranger M, & Baughman S (2018). Cognitive control, word retrieval and bilingual aphasia: Is there a relationship? *Journal of Neurolinguistics*, 45, 95–109. <https://doi.org/10.1016/j.jneuroling.2016.07.001>
- Faroqi-Shah Y, Treanor A, Ratner NB, Ficek B, Webster K, & Tsapkini K (2020). Using narratives in differential diagnosis of neurodegenerative syndromes. *Journal of Communication Disorders*, 85, 105994. 10.1016/j.jcomdis.2020.105994 [PubMed: 32388191]
- Favaro A, Moro-Velázquez L, Butala A, Motley C, Cao T, Stevens RD, ... Dehak N. (2023). Multilingual evaluation of interpretable biomarkers to represent language and speech patterns in Parkinson's disease. *Frontiers in Neurology*, 14, 1142642. 10.3389/fneur.2023.1142642
- Ferrante FJ, Migeot JA, Birba A, Amoroso L, Pérez G, Hesse E, Tagliazucchi E, Estienne C, Serrano C, Slachevsky A, Matallana D, Reyes P, Ibáñez A, Fittipaldi S, González Campo C & García AM

(accepted). Multivariate word properties in fluency tasks reveal markers of Alzheimer's dementia. *Alzheimer's & Dementia*.

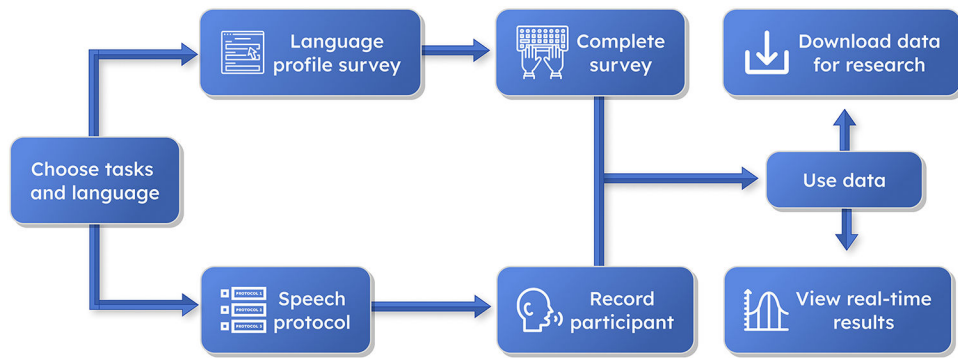
- Fraser KC, Lundholm Fors K, Eckerström M, Öhman F, & Kokkinakis D (2019). Predicting MCI status from multimodal language data using cascaded classifiers. *Frontiers in Aging Neuroscience*, 11, 205–205. 10.3389/fnagi.2019.00205 [PubMed: 31427959]
- Fraser KC, Meltzer JA, & Rudzicz F (2016). Linguistic features identify Alzheimer's disease in narrative speech. *Journal of Alzheimer's Disease*, 49(2), 407–422. 10.3233/jad-150520
- Fraser KC, Meltzer JA, Graham NL, Leonard C, Hirst G, Black SE, & Rochon E (2014). Automated classification of primary progressive aphasia subtypes from narrative speech transcripts. *Cortex*, 55, 43–60. 10.1016/j.cortex.2012.12.006 [PubMed: 23332818]
- García AM, Arias-Vergara T, J CV-C, Nöth E, Schuster M, Welch AE, ... Orozco-Arroyave JR (2021). Cognitive determinants of dysarthria in Parkinson's disease: An automated machine learning approach. *Movement Disorders* 36, 2862–2873. 10.1002/mds.28751 [PubMed: 34390508]
- García AM, Carrillo F, Orozco-Arroyave JR, Trujillo N, Vargas Bonilla JF, Fittipaldi S, ... Cecchi GA (2016). How language flows when movements don't: An automated analysis of spontaneous discourse in Parkinson's disease. *Brain Lang*, 162, 19–28. <https://doi.org/10.1016/j.bandl.2016.07.008> [PubMed: 27501386]
- García AM, Escobar-Grisales D, Vásquez Correa JC, Bocanegra Y, Moreno L, Carmona J, & Orozco-Arroyave JR (2022a). Detecting Parkinson's disease and its cognitive phenotypes via automated semantic analyses of action stories. *npj Parkinson's Disease*, 8(1), 163. 10.1038/s41531-022-00422-8
- García AM, Welch AE, Mandelli ML, Henry ML, Lukic S, & Torres Prioris MJ (2022b). Automated detection of speech timing alterations in autopsy-confirmed nonfluent/agrammatic variant primary progressive aphasia. 99(5), e500–e511. 10.1212/wnl.0000000000200750
- García AM, de Leon J, Tee BL, Blasi DE & Gorno-Tempini ML (2023). Speech and language markers of neurodegeneration: A call for global equity. *Brain*. doi: 10.1093/brain/awad253
- GBD 2017 US Neurological Disorders Collaborators (2021). Burden of neurological disorders across the US from 1990-2017: A global burden of disease study. *JAMA Neurology*, 78(2), 165–176. 10.1001/jamaneurol.2020.4152 [PubMed: 33136137]
- Gertken LM, Amengual M, & Birdong D (2014). Assessing language dominance with the Bilingual Language Profile. In Leclercq P, Edmonds A, & Hilton H (Eds.), *Measuring L2 Proficiency: Perspectives from SLA*. Multilingual Matters.
- Hao X, Su X, Horaud R and Li X (201). Fullsubnet: A full-band and sub-band fusion model for real-time single-channel speech enhancement. ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, ON, Canada, 2021, pp. 6633–6637, doi: 10.1109/ICASSP39728.2021.9414177
- Haulcy R, & Glass J (2020). Classifying Alzheimer's disease using audio and text-based representations of speech. *Frontiers in Psychology*, 11, 624137. 10.3389/fpsyg.2020.624137 [PubMed: 33519651]
- Hecker P, Steckhan N, Eyben F, Schuller BW, & Arnrich B (2022). Voice analysis for neurological disorder recognition: A systematic review and perspective on emerging trends. *Frontiers in Digital Health*, 4, 842301. 10.3389/fdgth.2022.842301 [PubMed: 35899034]
- Isaacson RS, Ganzer CA, Hristov H, Hackett K, Caesar E, Cohen R, ... Krikorian R (2018). The clinical practice of risk reduction for Alzheimer's disease: A precision medicine approach. *Alzheimer's & Dementia*, 14(12), 1663–1673. 10.1016/j.jalz.2018.08.004
- Isaacson RS, Hristov H, Saif N, Hackett K, Hendrix S, Melendez J, ... Krikorian R (2019). Individualized clinical management of patients at risk for Alzheimer's dementia. *Alzheimer's & Dementia*, 15(12), 1588–1602. 10.1016/j.jalz.2019.08.198
- Jonell P, Moëll B, Håkansson K, Henter GE, Kucherenko T, Mikheeva O, ... Beskow J (2021). Multimodal capture of patient behaviour for improved detection of early dementia: Clinical feasibility and preliminary results. *Frontiers in Computer Science*, 3. 10.3389/fcomp.2021.642633
- Laske C, Sohrabi HR, Frost SM, López-de-Ipiña K, Garrard P, Buscema M, ... O'Bryant SE (2015). Innovative diagnostic tools for early detection of Alzheimer's disease. *Alzheimer's & Dementia*, 11(5), 561–578. 10.1016/j.jalz.2014.06.004

- Li J, Song K, Li J, Zheng B, Li DS, Wu X, ... Meng HM (2023). Leveraging pretrained representations with task-related keywords for Alzheimer's disease detection. *ArXiv*, abs/2303.08019.
- Luz S, Haider F, de la Fuente Garcia S, Fromm D, & MacWhinney B (2021). Editorial: Alzheimer's dementia recognition through spontaneous speech. *Frontiers in Computer Science*, 3. 10.3389/fcomp.2021.780169
- Luz S, Haider F, de la Fuente S, Fromm D, & MacWhinney B (2020). Alzheimer's dementia recognition through spontaneous speech: The ADReSS challenge. *Proceedings of Interspeech 2020*, 2172–2176. doi:10.21437/Interspeech.2020-2571
- MacFarlane H, Salem AC, Chen L, Asgari M, & Fombonne E (2022). Combining voice and language features improves automated autism detection. *Autism Research*, 15(7), 1288–1300. 10.1002/aur.2733 [PubMed: 35460329]
- Macoir J, Hudon C, Tremblay MP, Laforce RJ, & Wilson MA (2019). The contribution of semantic memory to the recognition of basic emotions and emotional valence: Evidence from the semantic variant of primary progressive aphasia. *Social Neuroscience*, 14(6), 705–716. 10.1080/17470919.2019.1577295 [PubMed: 30714843]
- Meilán JJ, Martínez-Sánchez F, Carro J, López DE, Millian-Morell L, & Arana JM (2014). Speech in Alzheimer's disease: can temporal and acoustic parameters discriminate dementia? *Dementia and Geriatric Cognitive Disorders*, 37(5-6), 327–334. 10.1159/000356726 [PubMed: 24481220]
- Mendez MF, Carr AR, & Paholpak P (2017). Psychotic-like speech in frontotemporal dementia. *Journal of Neuropsychiatry and Clinical Neuroscience*, 29(2), 183–185. 10.1176/appi.neuropsych.16030058
- Mota NB, Copelli M, & Ribeiro S (2017). Thought disorder measured as random speech structure classifies negative symptoms and schizophrenia diagnosis 6 months in advance. *npj Schizophrenia*, 3(1), 18. 10.1038/s41537-017-0019-3 [PubMed: 28560264]
- Mota NB, Furtado R, Maia PPC, Copelli M, & Ribeiro S (2014). Graph analysis of dream reports is especially informative about psychosis [Article]. *Scientific Reports*, 4, 3691. 10.1038/srep03691 [PubMed: 24424108]
- Mota NB, Vasconcelos NA, Lemos N, Pieretti AC, Kinouchi O, Cecchi GA, ... Ribeiro S (2012). Speech graphs provide a quantitative measure of thought disorder in psychosis. *PLoS One*, 7(4), e34928. 10.1371/journal.pone.0034928 [PubMed: 22506057]
- Nandi A, Counts N, Chen S, Seligman B, Tortorice D, Vigo D, & Bloom DE (2022). Global and regional projections of the economic burden of Alzheimer's disease and related dementias from 2019 to 2050: A value of statistical life approach. *eClinicalMedicine*, 51. 10.1016/j.eclinm.2022.101580
- Nevler N, Ash S, Irwin DJ, Liberman M, & Grossman M (2019). Validated automatic speech biomarkers in primary progressive aphasia. *Annals of Clinical and Translational Neurology*, 6(1), 4–14. 10.1002/acn3.653 [PubMed: 30656179]
- Nevler N, Ash S, Jester C, Irwin DJ, Liberman M, & Grossman M (2017). Automatic measurement of prosody in behavioral variant FTD. *Neurology*, 89(7), 650–656. 10.1212/wnl.0000000000004236 [PubMed: 28724588]
- Nichols E, Steinmetz JD, Vollset SE, Fukutaki K, Chalek J, Abd-Allah F, ... Vos T (2022). Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. *The Lancet Public Health*, 7(2), e105–e125. 10.1016/S2468-2667(21)00249-8 [PubMed: 34998485]
- Norel R, Agurto C, Heisig S, Rice JJ, Zhang H, Ostrand R, ... Cecchi GA (2020). Speech-based characterization of dopamine replacement therapy in people with Parkinson's disease. *npj Parkinson's Disease*, 6(12). <https://doi.org/10.1038/s41531-020-0113-5>
- Orimaye SO, Wong JSM, Golden KJ, Wong CP, & Soyiri IN (2017). Predicting probable Alzheimer's disease using linguistic deficits and biomarkers. *BMC Bioinformatics*, 18(1), 34–34. 10.1186/s12859-016-1456-0 [PubMed: 28088191]
- Orimaye SO, Wong JS-M, & Wong CP (2018). Deep language space neural network for classifying mild cognitive impairment and Alzheimer-type dementia. *PLOS ONE*, 13(11), e0205636. 10.1371/journal.pone.0205636 [PubMed: 30403676]

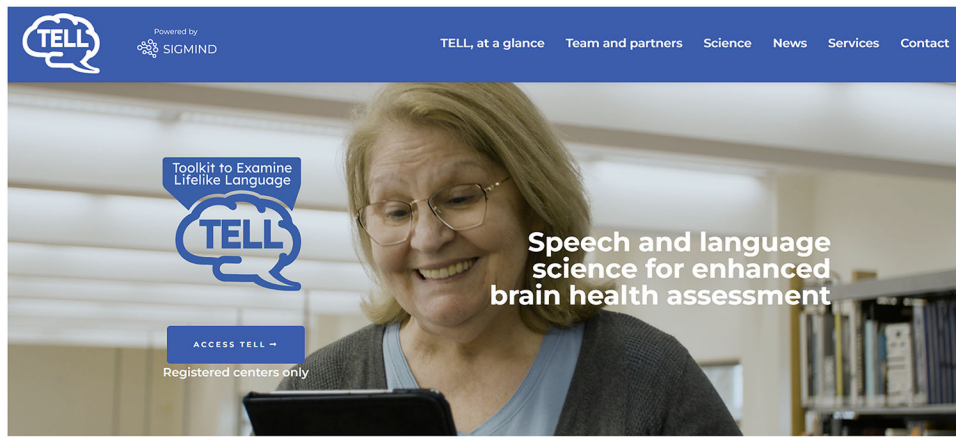
- Padró L & Stanilovsky E (2012). FreeLing 3.0: Towards wider multilinguality. In Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12), 2473–2479, Istanbul, Turkey. European Language Resources Association (ELRA).
- Paek EJ (2021). Emotional valence affects word retrieval during verb fluency tasks in Alzheimer's dementia. *Frontiers in Psychology*, 12, 777116. 10.3389/fpsyg.2021.777116
- Parra MA, Baez S, Allegrí R, Nitrini R, Lopera F, Slachevsky A, ... Ibáñez A (2018). Dementia in Latin America: Assessing the present and envisioning the future. *Neurology*, 90(5), 222–231. 10.1212/wnl.0000000000004897 [PubMed: 29305437]
- Parra M, Orellana P, León T, Cabello V, Henriquez F, Gomez R, ..., Durán-Aniotz C (2023). Biomarkers for dementia in Latin American countries: Gaps and opportunities. *Alzheimer's & Dementia*, 19(2), 721–735. <https://doi.org/10.1002/alz.12757>
- Pell MD, Cheang HS, & Leonard CL (2006). The impact of Parkinson's disease on vocal-prosodic communication from the perspective of listeners. *Brain & Language*, 97(2), 123–134. 10.1016/j.bandl.2005.08.010 [PubMed: 16226803]
- Pérez-Toro PA, Klumpp P, Hernández A, Arias-Vergara T, Lillo P, Slachevsky A, ... Orozco-Arroyave JR (2022). Alzheimer's detection from English to Spanish using acoustic and linguistic embeddings. 23rd Interspeech Conference, Incheon, Korea, 2483–2487.
- Radford A, Kim JW, Xu T, Brockman G, McLeavey C, & Sutskever I (2023). Robust speech recognition via large-scale weak supervision. Proceedings of the 40th International Conference on Machine Learning, 202, 28492–28518. Accessed on June 6, 2023.
- Rentoumi V, Raoufian L, Ahmed S, de Jager CA, & Garrard P (2014). Features and machine learning classification of connected speech samples from patients with autopsy proven Alzheimer's disease with and without additional vascular pathology. *Journal of Alzheimer's Disease*, 42 Suppl 3, S3–17. 10.3233/jad-140555
- Riley KP, Snowdon DA, Desrosiers MF, & Markesbery WR (2005). Early life linguistic ability, late life cognitive function, and neuropathology: findings from the Nun Study. *Neurobiology of Aging*, 26(3), 341–347. <https://doi.org/10.1016/j.neurobiolaging.2004.06.019> [PubMed: 15639312]
- Rusz J, & Tykalová T (2021). Does cognitive impairment influence motor speech performance in de novo Parkinson's disease? *Movement Disorders*, 36(12), 2980–2982. <https://doi.org/10.1002/mds.28836> [PubMed: 34921455]
- Rusz J, Cmejla R, Tykalova T, Ruzickova H, Klempir J, Majerova V, ... Ruzicka E (2013). Imprecise vowel articulation as a potential early marker of Parkinson's disease: effect of speaking task. *J Acoust Soc Am*, 134(3), 2171–2181. 10.1121/1.4816541 [PubMed: 23967947]
- Sanz C, Carrillo F, Slachevsky A, Forno G, Gorno Tempini ML, Villagra R, ... García AM (2022). Automated text-level semantic markers of Alzheimer's disease. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 14(1), e12276. <https://doi.org/10.1002/dad2.12276>
- Seixas Lima B, Levine B, Graham NL, Leonard C, Tang-Wai D, Black S, & Rochon E (2020). Impaired coherence for semantic but not episodic autobiographical memory in semantic variant primary progressive aphasia. *Cortex*, 123, 72–85. 10.1016/j.cortex.2019.10.008 [PubMed: 31760339]
- Singh S, Bucks RS, & Cuerden JM (2001). Evaluation of an objective technique for analysing temporal variables in DAT spontaneous speech. *Aphasiology*, 15(6), 571–583. 10.1080/02687040143000041
- Themistocleous C, Webster K, Afthinos A, & Tsapkini K (2021). Part of speech production in patients with primary progressive aphasia: An analysis based on natural language processing. *Am J Speech Lang Pathol*, 30(1s), 466–480. 10.1044/2020\_ajslp-19-00114 [PubMed: 32697669]
- Tosto G, Gasparini M, Lenzi GL, & Bruno G (2011). Prosodic impairment in Alzheimer's disease: assessment and clinical relevance. *J Neuropsychiatry Clin Neurosci*, 23(2), E21–23. 10.1176/jnp.23.2.jnp21
- Van Der Donckt J, Kappen M, Degraeve V, Demuyneck K, Vanderhasselt M, & Van Hoecke S (2023). Ecologically valid speech collection in behavioral research: The Ghent Semi-spontaneous Speech Paradigm (GSSP). *Behavior Research Methods*, forthcoming. <https://doi.org/10.31234/osf.io/e2qxw>



- Wang J, Zhang L, Liu T, Pan W, Hu B, & Zhu T (2019). Acoustic differences between healthy and depressed people: a cross-situation study. *BMC Psychiatry*, 19(1), 300. 10.1186/s12888-019-2300-7 [PubMed: 31615470]
- Webber J, Parastatidis S, & Robinson I (2010). *REST in Practice: Hypermedia and Systems Architecture*. O'Reilly Media, Incorporated. <https://books.google.cl/books?id=5CjJci4UfMC>
- Wilson SM, Henry ML, Besbris M, Ogar JM, Dronkers NF, Jarrold W, ... Gorno-Tempini ML (2010). Connected speech production in three variants of primary progressive aphasia. *Brain*, 133(Pt 7), 2069–2088. 10.1093/brain/awq129 [PubMed: 20542982]
- Zimmerer VC, Hardy CJD, Eastman J, Dutta S, Varnet L, Bond RL, ... Varley RA (2020). Automated profiling of spontaneous speech in primary progressive aphasia and behavioral-variant frontotemporal dementia: An approach based on usage-frequency. *Cortex*. <https://doi.org/10.1016/j.cortex.2020.08.027>



**Figure 1.** TELL's Basic Workflow. The app offers a linguistic profile survey and a speech protocol, whose outputs can be downloaded for research purposes and visualized in real time for clinical uses.



**Figure 2.** Accessing TELL. Users can access the app through its website’s homepage upon registration.





Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Participants

Last and first names	Record	Questionnaire
Cristi Daniel		
González Catalina		

**Figure 3.**

Participant List. Participants are saved and listed prior to data collection. They can be searched by name or other metadata and, once located, immediate access can be gained to their survey and recording data.

LINGUISTIC PROFILE QUESTIONNAIRE - TELL

We would like to ask you some questions concerning your language skills. There are no right or wrong answers. Please answer each question honestly. Thank you very much!

---

**GENERAL INFORMATION**  
GENERAL INFORMATION

Last name and first name (or participant code)

Gender

Age

Years of education (counting from elementary school onwards)

**LINGUISTIC BACKGROUND**  
LINGUISTIC BACKGROUND

At what age did you start learning your first language

At what age did you start learning your second language

**USE OF LANGUAGES**  
THE TOTAL USE OF ALL LANGUAGES IN EACH QUESTION MUST REACH 100%.

In an average week, what percentage of the time do you use your first language with your family and friends?

In an average week, what percentage of the time do you use your second language with your family and friends?

In an average week, what percentage of the time do you use your other languages with your family and friends?

**PROFICIENCY**  
FOR THIS SECTION, WE WOULD LIKE YOU TO CONSIDER YOUR LANGUAGE PROFICIENCY BY CHECKING A BOX FROM 0 TO 6.

How well do you SPEAK your first language?

How well do you SPEAK your second language?

How well do you UNDERSTAND your first language?

How well do you UNDERSTAND your second language?

How well do you READ in your first language?

How well do you READ in your second language?

How well do you WRITE in your first language?

How well do you WRITE in your second language?

**DIFFICULTIES**

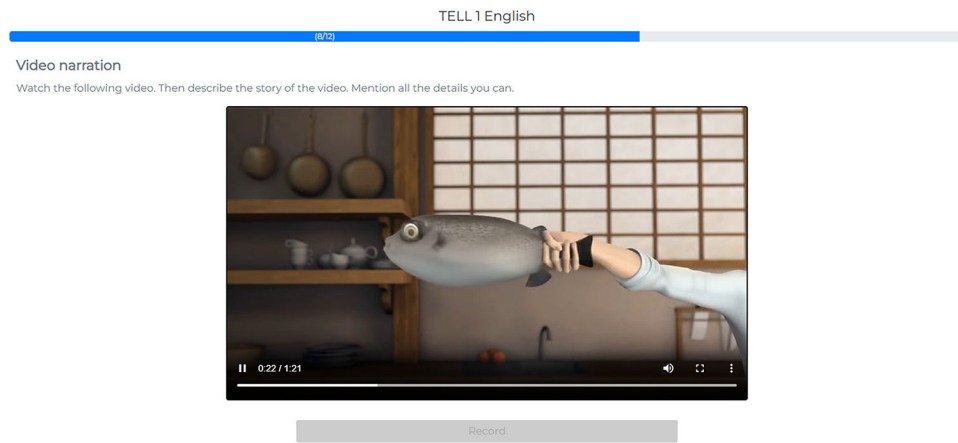
Have you ever had problems producing or understanding in any of your languages? Please answer YES or NO.

If YES, what kind of difficulties did you notice?

At what age did you start noticing these difficulties?

How much do these difficulties affect your daily activities (domestic, professional, recreational)? Choose a number from below:

**Figure 4.** TELL’s Language Profile Survey. A simple instrument provides self-report data on language history, use, proficiency, and perceived difficulties.



**Figure 5.** TELL’s Speech Recording Interface. Each task includes a title, an instruction, and a ‘record’ button. These elements are accompanied by specific stimuli in certain tasks (as in ‘video narration’, shown here).

TELL 1 English

Typical day

Pleasant memory

Picture description 1

Picture description 2

Story retelling

Paragraph reading

Sustained vowel

Syllables

Video narration

Car without wheels

Testimony


### PATIENT NARRATION

» Recording date

April 12, 2023 11:37 PM

» Task

Watch the following video. Then describe the story of the video. Mention all the details you can. When you are done, click the 'Record' button again to stop the recording.



» Audio

▶ 0:00 ◀

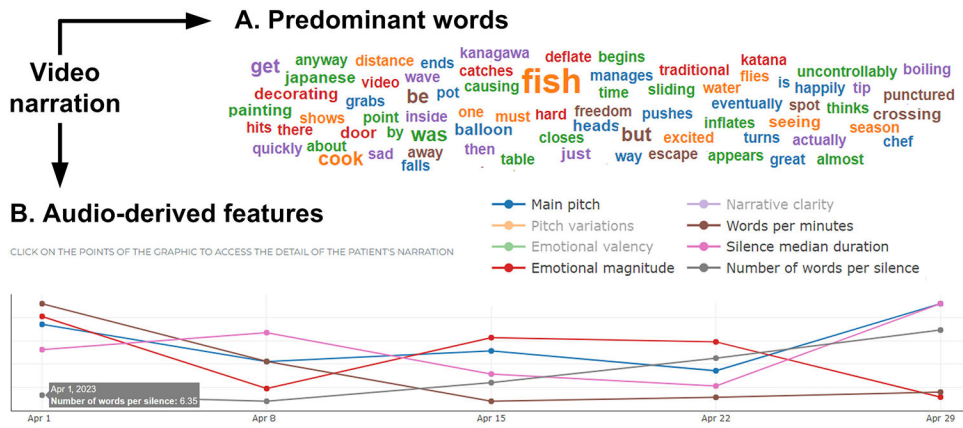
» Transcript

The video shows a fish about to be cooked in boiling water, the fish almost falls into the pot, but the cook, who seems to be Japanese catches a just-in-time. That's when the fish inflates like a balloon and flies away trying to escape at one point. The fish season the distance, what he thinks is the sea. So he gets excited and quickly, he gets that way. To get there, it must cross through a traditional Japanese sliding door but the chef closes the door just as the fishes Crossing and grabs it. The fish manages to get across and heads happily toward its freedom but it turns out that the see the fish was seeing was actually a decorative painting of The Great Wave off. Kanagawa the fish hits, the painting hard and it pushes it backwards, causing it to be punctured by the tip of a katana. That was decorating a table. The fish then begins to deflate uncontrollably like a balloon and eventually ends up inside the cook spot. Anyway, its kind of sad.

### RECENT ACTIVITY

Monday, February 20, 2023 7:37 PM	TELL 1 English	<a href="#">See metrics obtained</a>	<a href="#">↓</a>
Friday, February 10, 2023 4:54 PM	TELL 1 English	<a href="#">See metrics obtained</a>	<a href="#">↓</a>

**Figure 6.** TELL's Task Summary and Data Download Menus. The prompts and outputs of each task are jointly presented in a menu for quick overview and download.



**Figure 7.** Examples of Data Visualization. The figure illustrates the graphs generated for the (A) ‘predominant words’ metric and (B) several audio-derived metrics. The latter are identified by different colors on a single plot. The specific value of each point in each variable is revealed by placing the mouse (or a finger, in the case of touch-screens) on the data point. Note that, while the figure illustrates successive assessments through weekly measurements, longitudinal differences in neurodegenerative diseases may likely require several months to become evident.



**Table 1**

Tasks Included in TELL v.1.0.0.

Task	Instructions	Stimulus	Minimum time
Typical day	Describe what your daily routine is like. Tell me everything you do in a typical day, from when you wake up to when you go to sleep. Provide as much detail as you can.	None	30 sec
Pleasant memory	Think of a pleasant memory in your life. It can be a trip, a special moment, a celebration, or anything that you remember with joy. Describe it in as much detail as you can.	None	30 sec
Picture description	Look at this picture carefully and, as you look at it, describe the situation in as much detail as you can. You can start wherever you want.	Pictures of daily scenes	30 sec
Story retelling	You will now hear a short story. Then I will ask you to narrate it in your own words. Provide as much detail as possible.	Recorded story	30 sec
Paragraph reading	Start recording and read the following paragraph. Read it normally, from beginning to end. When you are done, click the 'Record' button again to stop the recording.	Written paragraph	30 sec
Sustained vowel	Say the vowel 'A' continuously, for as long as you can, until your breath runs out. When you are done, click again on the 'Record' button to stop the recording.	None	None
Syllables	Repeat the syllables 'PATAKA' for as long as you can, until you run out of breath. When you are done, click the 'Record' button again to stop the recording.	None	None
Video narration	Watch the following video. Then describe the story of the video. Mention all the details you can.	Animated silent film	30 sec
Car without wheels	Answer the following question: what could a car without wheels be used for? List as many uses as you can. Let your imagination run wild and describe all the ideas that come to mind.	Picture of a wheel-less car	30 sec
Testimony	What is it like living with your disease? What changes have you had in your life because of it?	None	30 sec
Silence recording	Lastly, let's do a silent recording. Just press 'Record' and say nothing. Stop the recording after 15 seconds.	None	15 sec

**Table 2**

Metrics Included in TELL v.1.0.0. The app includes eight metrics validated in abundant works.

Metric	Source	Description	References
Speech rate	Audio and transcript	This metric captures speech production timing. High values may be linked to disinhibition, fluent aphasias or manic episodes. Low values could indicate motor speech disorders.	García et al. (2021) García et al. (2022b) Meilán et al. (2014)
Silence duration	Audio	This metric reflects the length of silences between sentences, words, or word segments. Silences with abnormally long duration are observed in neurodegenerative and neuropsychiatric conditions. Sudden shorter silences may be associated with hypomanic episodes.	García et al. (2021) García et al. (2022b) Singh et al. (2001)
Silences per word	Audio and transcript	This metric represents the proportion of non-speech segments to lexical items. It captures how often people pause during speech, providing insights on fluency.	García et al. (2021) García et al. (2022b) Meilán et al. (2014)
Main pitch	Audio	This metric quantifies the fundamental pitch of vocal fold vibration, with lower values associated to motor speech or affective anomalies.	Agurto et al. (2020) Nevler et al. (2017) Nevler et al. (2019)
Pitch variation	Audio	Pitch varies considerably in normal speech. Decrease in this variation might reflect motor, cognitive, or affective deficits.	Bowen et al. (2013) García et al. (2021) Wang et al. (2019)
Predominant words	Transcript	The size of each word represents how many times it was said in the recording. The larger the word, the more frequent it was. Hence, its associated concept is predominant in speech.	Dodge et al. (2015)
Affective valence	Transcript	Values express negativity if close to -1, neutrality if close to 0, and positivity if close to 1.	Ellis et al. (2015) Macoir et al. (2019) Paek (2021)
Emotional magnitude	Transcript	Values close to 0 express null emotionality. The higher the value, the greater the emotionality. Persons with Alzheimer's and Parkinson's disease tend to have more neutral values.	Carrillo et al. (2018) Paek (2021) Pell et al. (2006) Tosto et al. (2011)