Schizophrenia

ARTICLE OPEN

Check for updates

# Collecting language, speech acoustics, and facial expression to predict psychosis and other clinical outcomes: strategies from the AMP® SCZ initiative

Zarina R. Bilgrami[1], Eduardo Castro[2], Carla Agurto[2], Einat Liebenthal[3,4], Michaela Ennis[4], Justin T. Baker[3,4], Isabelle Scott[5,6], Beau-Luke Colton[5,6], Kang Ik K. Cho[7], Linying Li[1], Zailyn Tamayo[8], Mara Henecks[8], Habiballah Rahimi Eichi[3,4], Tae'lar Henry[1], Jean Addington[9], Luis K. Alameda[10,11], Celso Arango[12], Nicholas J. K. Breitborde[13,14], Matthew R. Broome[15,16], Kristin S. Cadenhead[17], Monica E. Calkins[18], Eric Yu Hai Chen[19], Jimmy Choi[20], Philippe Conus[10,21], Barbara A. Cornblatt[22,23], Lauren M. Ellman[24], Paolo Fusar-Poli[11,25], Pablo A. Gaspar[26], Carla Gerber[27], Louise Birkedal Glenthøj[28], Leslie E. Horton[29], Christy Hui[19], Joseph Kambeitz[30], Lana Kambeitz-Ilankovic[30], Matcheri S. Keshavan[3], Sung-Wan Kim[31,32], Nikolaos Koutsouleris[11,33], Jun Soo Kwon[34,35], Kerstin Langbein[36], Daniel Mamah[37], Covadonga M. Diaz-Caneja[12], Daniel H. Mathalon[38,39], Vijay A. Mittal[40], Merete Nordentoft[41,42], Godfrey D. Pearlson[1,8], Jesus Perez[43,44], Diana O. Perkins[45], Albert R. Powers III[8,46], Jack Rogers[15], Fred W. Sabb[47], Jason Schiffman[48], Jai L. Shah[49,50], Steven M. Silverstein[51], Stefan Smesny[36], William S. Stone[4,52], Walid Yassin[4,52], Gregory P. Strauss[53], Judy L. Thompson[54,55], Rachel Upthegrove[15], Swapna Verma[56], Jijun Wang[57], Daniel H. Wolf[18], Patrick D. McGorry[5,6], Rene S. Kahn[58], John M. Kane[23,59], Alan Anticevic[8,46], Carrie E. Bearden[60,61], Dominic Dwyer[5,6], Tashrif Billah[7], Sylvain Bouix[7,62], Ofer Pasternak[7], Martha E. Shenton[4,7], Scott W. Woods[8], Barnaby Nelson[5,6], Accelerating Medicines Partnership® Schizophrenia (AMP® SCZ)*, Guillermo A. Cecchi[2], Cheryl M. Corcoran[58,63] and Phillip M. Wolff[1,63]✉

Speech-based detection of early psychosis is progressing at a rapid pace. Within this evolving field, the Accelerating Medicines Partnership® in Schizophrenia (AMP® SCZ) is uniquely positioned to deepen our understanding of how language and related behaviors reflect early psychosis. We begin with detailed standard operating procedures (SOPs) that govern every stage of collection. These SOPs specify how to elicit speech, capture facial expressions, and record acoustics in synchronized audio–video files—both on-site and through remote platforms. We then explain how we chose our sampling tasks, hardware, and software, and how we built streamlined pipelines for data acquisition, aggregation, and processing. Robust quality-assurance and quality-control (QA/QC) routines, along with standardized interviewer training and certification, ensure data integrity across sites. Using natural language processing parsers, large language models, and machine-learning classifiers, we analyzed Data Release 3.0 to uncover systematic grammatical markers of psychosis risk. Speakers at clinical high risk (CHR) produced more referential language but fewer adjectives, adverbs, and nouns than community controls (CC), a pattern that replicated across sampling tasks. Some effects were task-specific: CHR participants showed elevated use of complex syntactic embeddings in two elicitation conditions but not the third, underscoring the importance of the language sampling task. Together, these results demonstrate how computational linguistics can turn everyday speech into a scalable, objective biomarker, paving the way for earlier and more precise detection of psychosis.

Video Link: https://vimeo.com/1112291965?fl=pl&fe=sh

*Schizophrenia* (2025)11:125 ; https://doi.org/10.1038/s41537-025-00669-z

## INTRODUCTION: THE POTENTIAL OF LANGUAGE AS A BIOMARKER

Psychosis is often indicated by distinct changes in speech, characterized by patterns of language that suggest illogical thinking, limited content, and loose associations (Andreasen[1–4]. The early appearance of these language patterns during the prodromal phase indicates that they may potentially serve as early indicators of psychosis[5,6]. In recognition of its importance, assessment of spoken language has been incorporated into standardized diagnostic tools for psychosis, including the Positive and Negative Symptoms Scale (PANSS)[7]; the Structured Interview for Psychosis-Risk Syndromes (SIPS)[8] and the Comprehensive Assessment of At-Risk Mental States (CAARMS)[9]. Additionally, tools have been developed for the direct assessment of atypical language production, such as the Scale for the Assessment of Thought Disorder, Language and Communication (TLC)[1] the Thought Disorder Index (TDI)[10], and the thought and language disorder (TALD) scale[11].

The value of language biomarkers has recently been enhanced with the advent of natural language processing (NLP) and artificial intelligence (AI) techniques. Early demonstrations of these methods have shown how they can facilitate the identification of early psychosis through the analysis of discourse coherence and syntactic complexity[12,13], semantic density and content[14], and

A full list of author affiliations appears at the end of the paper.

Published in partnership with the Schizophrenia International Research Society

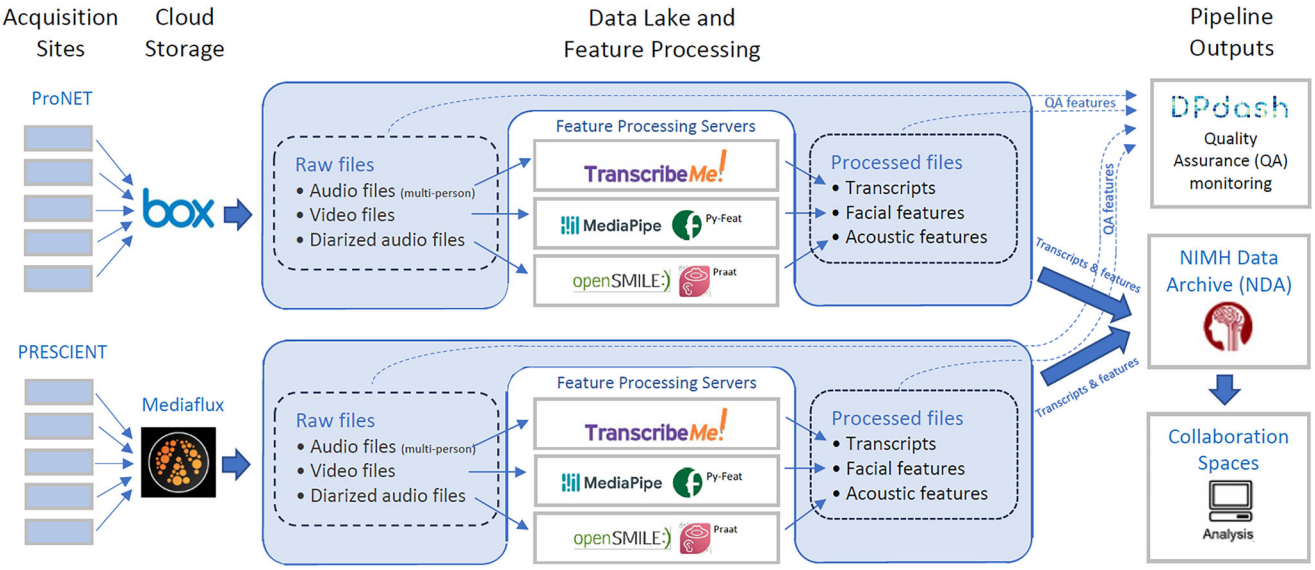


**Fig. 1 AVL processing pipeline.** Files collected by the sites are uploaded to a secure cloud storage system. Raw files are separated into combined audio, diarized audio, and video files, which are sent to feature processing services or servers to produce transcripts, acoustic analyses, and facial analyses. The results of quality checks, conducted at the initial submission of the files to the data aggregate server and later, after the files are processed for features, are sent to the data visualization platform DPdash for QA/QC monitoring. Finalized files are sent to the NIMH data archive (NDA), which conducts the final curating of the data prior to releasing it to the collaboration server for further analysis, and to the general research community.

speech connectedness[15]. Furthermore, computational methods have successfully extracted psychosis indicators based on prosodic features, speech pauses[16,17], and even facial expressions and movements[18,19]. Notably, NLP and machine learning (ML) have played a crucial role in identifying linguistic indicators of psychosis across phylogenetically distinct languages, like English and Mandarin Chinese, suggesting the possibility of deep-seated markers of the condition across diverse linguistic systems[20].

While promising, the demonstrated benefits of automated approaches in the extraction of spoken language and facial expression biomarkers have been based on small-scale studies, which are vulnerable to the risk of statistical overfitting. There is thus a pressing need to collect language samples on a larger scale. The pursuit of such comprehensive sampling is crucial for the identification of robust and repeatable language biomarkers. This endeavor is supported by the Accelerating Medicines Partnership® in Schizophrenia (AMP® SCZ) initiative. This project is a collaborative effort involving two specialized research networks focused on collecting data from individuals at clinical high risk (CHR) for psychosis—the Psychosis Risk Outcomes Network (ProNET), which operates across 28 sites, and the Prediction Scientific Global Consortium (PRESCIENT), active across 15 sites. The third project, the Psychosis Risk Evaluation, Data Integration, and Computational Technologies: Data Processing, Analysis, and Coordination Center (PREDICT-DPACC), is dedicated to the aggregation, processing, and analysis of the data. This includes the construction and maintenance of servers and software platforms, along with rigorous quality assessment and control (QA/QC) monitoring.

In this paper, we describe the methodologies and automated processing systems developed for extracting language samples from a large-scale international cohort, encompassing a diverse range of languages. We focus on the standard operating procedures (SOPs) developed for eliciting spoken language, as well as for capturing facial expressions and speech acoustics through both in-person and virtual audio-video recordings. These procedures include an in-depth description of interview strategies, recording equipment, and digital platforms, a discussion of software packages for the initial analyses of these novel biomarkers, a review of the QA/QC measures implemented to ensure data quality, and a summary of the training programs and certification processes used to prepare study staff. Lastly, preliminary findings are reported to show how different interview styles may yield different kinds of information. The latter suggests that certain interview methods might be especially effective for extracting distinct information critical to the early detection of psychosis.

## AUDIO/VIDEO/LANGUAGE (AVL) PROCESSING PIPELINE

The collection of language samples, facial expressions, and speech acoustics is made possible by means of an AVL processing pipeline. The main parts of this pipeline are depicted in Fig. 1.

The pipeline starts by gathering audio and video samples from the various data acquisition sites. The collected data are then uploaded to cloud storage. Once there, the data are transferred to aggregation servers managed by the ProNET and PRESCIENT research networks, which are then harmonized with each other by the DPACC. At this stage, the data undergo quality control checks, which are subsequently forwarded to the data visualization platform, DPdash. Following quality control, features are extracted. Multi-speaker audio files are transcribed by a transcription service using human transcribers, video files are processed for feature extraction by an audio/video (A/V) server, and diarized, single-speaker audio files (see below) are analyzed for acoustic features. The outcomes of these feature extraction processes are consolidated on the aggregation servers and subsequently submitted to the National Institute of Mental Health (NIMH) Data Archive (NDA) for analysis on a collaboration server by researchers in the project and the broader research community. The original raw A/V files are stored on the aggregation servers and are not distributed to the broader research community without special case agreements with the NDA to preserve anonymity.

### Types of language sampling

At the data acquisition sites, three distinct types of language samples are systematically collected, each fulfilling a specific role within the study and adhering to detailed, standardized protocols.

The first type is the PSYCHS, a semi-structured clinical interview that captures responses directly related to questions assessing psychosis risk. The second type is an open-ended qualitative interview designed to gather spontaneous and narrative language samples, offering deeper insights into participants' expressive capabilities. Lastly, participants record daily diary entries independently, providing regular snapshots of their everyday language use and behaviors. Comprehensive descriptions of each language sample type are provided below.

*PSYCHS clinical interview*. The Positive SYmptoms and Diagnostic Criteria for the CAARMS Harmonized with the SIPS (PSYCHS) clinical interview is a semi-structured interview and constitutes the primary assessment tool for the AMP SCZ project. It is used for case identification, positive symptom ratings, and the determination of the primary outcome of psychosis transition. The PSYCHS harmonizes two widely used measures in clinical high-risk research: the Structured Interview for Psychosis-Risk Syndromes (SIPS)[8] and the Comprehensive Assessment of At-Risk Mental States (CAARMS)[21], which have been foundational tools in the field for over two decades. The PSYCHS assesses 15 distinct positive symptoms categorized into three primary groups: attenuated delusions, attenuated hallucinations, and attenuated thought disorder. The interviews are led by trained and certified research assistants and include a mix of verbatim inquiries and semi-structured follow-up queries (see also in this volume for more detail[22]. The duration of these interviews varies depending upon participants' symptoms and communication style, ranging from 30 min for those who do not have many symptoms (including community controls [CCs]) to over 2 hours for those who experience many psychotic-like symptoms. These interviews are conducted at multiple time points during the study: at screening/baseline and at 1, 2, 3, 6, 12, 18, and 24-month follow-ups, as well as upon confirmed transition to psychosis. The standardized structure across time points makes these language samples well-suited for longitudinal analysis. Although the full interview is recorded, only the first 30 min are transcribed manually. This duration was chosen as a balance—long enough to yield a meaningful sample of linguistic behavior but constrained to remain within budget. A potential limitation of the PSYCHS is its tendency, particularly in controls, to prompt denials of symptoms rather than elicit richer expressions of personal experiences or conversational speech.

*Open-ended qualitative interviews*. Open-ended interviews are conversational interactions in which participants have the freedom to select topics they find relevant to their experience. This approach aligns with the phenomenological interview techniques developed by Davidson and colleagues[23], in which the interviewer's role is to enable a natural, expansive conversation, remaining neutral to avoid influencing the participants' responses. The primary objective is to delve into the participant's personal narrative, drawing out anecdotes and encouraging spontaneous, unrestrained dialog. In this setting, the participant is regarded as an expert on their own life, while the interviewer acts as a facilitative witness, prompting the participant to provide in-depth insights into their lived experiences. Such interviews are distinct from therapeutic interactions, as they neither seek to achieve specific goals nor induce changes. These Open interviews are conducted both at baseline and during the 2-month follow-up. In Release 3, 915 participants completed baseline interviews and 461 completed the 2-month follow-up. The average interview length was 17.59 min (SD = 4.73) at baseline and 17.35 min (SD = 5.12) at follow-up. The inherent flexibility of such open-ended interviews allows for the capture of content that may be overlooked in more structured PSYCHS interviews, due to their prescriptive nature.

*Daily diaries*. Daily diaries represent a distinctive form of language sampling, consisting of audio-only recordings captured on smartphones. Participants are prompted to record a diary entry once per day at the conclusion of their ecological momentary assessment survey (see also, this volume[24]). Although participants may record more than once per day, only the first entry is transcribed. To keep file sizes manageable for mobile data plans, each entry was limited to 4 min. Across both ProNET and Prescient sites, 1281 participants generated 17,346 English-language diary entries, with an average length of 1.81 min (SD = 0.631). Because of their spontaneous nature, daily diaries offer the potential to capture richer and more varied insights into participants' everyday experiences, revealing nuances that scheduled interviews might otherwise miss.

## Methods for recording AVL samples

Interviews are recorded using two distinct methods, tailored to the interview's setting (onsite or remote) and type (open-ended or PSYCHS). This dual-method approach is summarized in Table 1. Given that the project started during the Covid-19 pandemic, as much flexibility for remote data collection is included as part of the protocol and left to the discretion of the sites. The choice of recording technology is influenced by the need to conduct acoustic analyses, which require audio files that feature only one voice. The process of isolating individual voices in an interview is known as diarization and is readily accomplished in online platforms like Zoom, where audio inputs from different speakers are naturally segregated by different input machines. Critically and uniquely, Zoom allows these audio streams to be saved to separate but synchronized audio files.

Open-ended interviews, which tend to elicit more natural conversations than the more structured PSYCHS interviews, were exclusively conducted using Zoom to benefit from its diarization capabilities. For onsite interviews, this meant having the interviewer and participant in separate rooms to maintain audio separation. Remotely conducted PSYCHS were always conducted via Zoom, but for onsite interviews, interviewers were given a choice between Zoom and a handheld digital recorder to enhance ease of recording. At the time of this publication, Zoom was used at the data acquisition sites three times more often than the handheld recorders, though some sites opted to use only the recorder.

While the digital recorder is less intrusive, it does not support recording facial data or diarization of speech streams. Nevertheless, as open-ended interviews already provide rich data for facial and acoustic analysis, capturing these data from the PSYCHS interviews was deemed optional. The 1 EVISTR digital recorder was recommended to sites as it can save recordings in WAV format at 1536 kbps and has a Micro USB port for efficient charging and downloading.

To optimize the sound quality in the recording of Open interviews, it was recommended that sites purchase two

**Table 1.** Recording method for onsite and remote open-ended and PSYCHS interviews and diaries.

| | Open-ended Interviews | PSYCHS interviews | Daily diaries |
|---|---|---|---|
| Onsite | Zoom (video and audio) | Handheld recorder or Zoom (audio only) | – |
| Remote | Zoom (video and audio) | Zoom (video and audio for ProNET network; audio only for PRESCIENT Network) | MindLAMP (audio only) |

WirelessFinest Monaural Headset Headphones with Microphones and a USB plug, one for the participant and the other for the interviewer, along with alcohol wipes for cleaning the microphone headsets after each use. Zoom Audio settings are changed to record voice with the highest fidelity possible and Recording settings are changed to enable the collection of diarized audio. Detailed instructions on these settings are available in the study's SOP available on the Accelerating Medicines Partnership Schizophrenia (AMP SCZ) website (www.ampscz.org).

For daily diaries, participants used the MindLAMP app on their smartphone. This is detailed in the companion paper on digital phenotyping in this special issue, led by Drs. John Torous and Justin Baker. This approach takes advantage of the widespread availability and user familiarity of smartphones, making the recording process both convenient and accessible. It also enables the collection of natural speech samples, capturing participants' spontaneous responses in real-world contexts outside of structured conversations. While these responses may not explicitly address classic symptoms of psychosis, they can still potentially offer valuable insights into participants' thought processes by revealing the topics they choose to discuss when speaking freely without prompts or an interlocuter.

### Processing of language, facial, and acoustic data types

Once interview recordings are completed, study staff at each site manually upload the files to designated cloud storage systems. ProNET utilizes Box (www.box.com) for this purpose, while PRESCIENT employs MediaFlux (https://www.arcitecta.com/mediaflux/). These cloud storage systems function as portals, enabling acquisition sites to securely upload data into the AVL pipeline.

Before processing the interview files, details of the interview session must be entered into databases managed by either REDCap or RPMS applications. REDCap run sheets are used by the ProNET research network, while RPMS run sheets are used by the PRESCIENT research network. These run sheets capture a variety of variables, including the recording environment (e.g., large/small room, outdoors, car, other), recording mode (remote or onsite; in the same or separate rooms), digital recording method (zoom, digital recorder, or other), type of device used by the participant (laptop/desktop, phone/tablet), any deviations from the established protocol, and the perceived quality of the recording.

With consent provided and documented and the run sheets filled out, the primary audio and video files undergo a preliminary quality control (QC) analysis as described below. This analysis, using specific programs detailed in the subsequent sections, focuses on key aspects such the length of the interview (0–80 min), the overall decibel level (40–90 dB), the number of faces detected in the video, and the percentage of frames containing two faces. The complete list of these features and their acceptable ranges can be found in Table S1. Any discrepancies or values falling outside the predetermined acceptable ranges are highlighted in the DPDash dashboard, triggering warning emails to study staff. The dashboard, managed by the DPACC, is closely monitored by study personnel, facilitating timely interventions at acquisition sites when necessary. (The dashboard is further described in detail in a companion paper in this special issue on the study-wide data flow, processing, and visualization, led by Drs. Sylvain Bouix and Justin Baker.) For the interview audio files to be transcribed, they must meet certain criteria, including a minimum duration of 10 min and a decibel level of at least 40 dB. Similarly, daily diary audio files must be at least 40 dB.

The process of collecting and processing interview files and daily diaries thus involves multiple parallel processing streams, each dedicated to analyzing different aspects of the data-language, speech acoustics, and facial expressions. Each of these streams employs a distinct set of programs and adheres to specific QC protocols, ensuring comprehensive and thorough analysis of the various data types.

### Language processing

The language processing stream of this project involves several critical steps, including language identification, transcription, redaction of sensitive information, and quality control. Transcription is handled by the HIPAA-compliant transcription service company TranscribeMe! using human transcribers.

*Language identification*. Initially, the language of each interview has to be identified before transcription. This is accomplished using a look-up table, which specifies the language spoken at each acquisition site. The identified language is added to the audio file's filename, ensuring it is directed to the appropriate team of transcribers at TranscribeMe!. The languages in this project include English, German, Danish, Korean, Cantonese, Mandarin, French, Spanish, and Italian. Of note, there are a few sites (e.g., Montreal), for which more than one language (e.g., English or French) may be used.

*Transcription process*. The transcription of audio files encompasses both Zoom recordings and digital recorder files from PSYCHS in-person sessions and is conducted in a "full" verbatim style. This approach captures speech in writing with the greatest possible accuracy, preserving utterances exactly as spoken. As such, the transcripts include filler words, grammatical errors, and nonlinguistic utterances. Each speaker is sequentially labeled (e.g., S1 and S2) according to the order in which they first started to speak in the interview, with S1 typically being a research team member. The transcripts are also detailed with timestamps at the second-level accuracy. The transcript is divided into entries, with each entry representing a change in the speaker. For illustrative purposes, a fictional example of an open-ended interview transcript, demonstrating this format, is provided in Fig. 2.

*Redaction of identifying information*. Human editors from TranscribeMe! carefully review the transcripts for protected health information (PHI) and personally identifiable information (PII). This includes names, geographic details smaller than a state, specific dates related to individuals, contact information, and

S1: 00:00:30.345 So how have things been going for you lately?
S2: 00:00:33.043 Good.
S1: 00:00:35.019 Good? What kinda things have you been up to?
S2: 00:00:38.524 Um, well, uh, I don't know. Not much.
S1: 00:01:02.706 Not much? How about, like-- what kinda stuff did you do this summer?
S2: 00:01:13.123 I went to {REDACTED} with my family.
S1: 00:01:15.864 Oh, yeah. What kinda things did you do there?
S2: 00:01:22.435 Well, we went to the lake, and we, like, we, we paddle boarded there, and we swam with my g-grandma--

**Fig. 2** A fictional example of an open-ended interview transcript complete with speaker labeling, timestamps, verbatim encodings, and redactions.

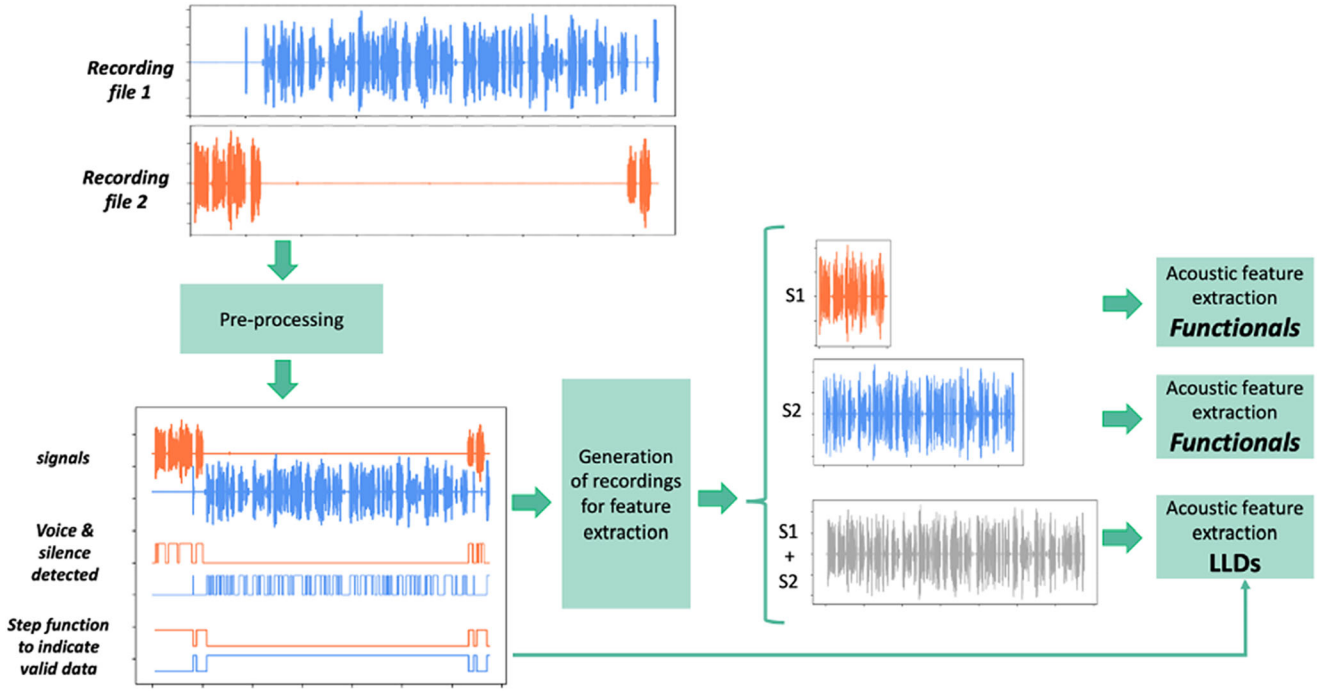


**Fig. 3 Audio files undergo a series of processes to identify acoustic features.** Zoom allows the audio from each speaker to be saved to separate files, here labeled Recording file 1 and Recording file 2. These files are then renamed S1 and S2, corresponding to the order in which the participants speak, with S1 designating the first speaker and S2 the second. During a pre-processing step, a step function is used to identify valid speech signals. The resulting recordings are used to extract two types of acoustic features: low-level descriptors (LLDs) and higher-level 'functional' features, the latter of which represents global properties of a participant's acoustic signal.

any unique identifying numbers (see Supplemental Materials for the PII/PHI redaction guidelines used in this project). Redacted information is initially marked with curly brackets, which are later replaced with the word "REDACTED." To avoid introducing inconsistencies or artifacts into the dataset, redacted words are not replaced with descriptive placeholders (e.g., replacing 'Atlanta' with {CITY} or a different city like {CHICAGO}), as such substitutions could affect measures of interest such as vagueness, concreteness, and conceptual coherence. Preliminary findings suggest that large language models (LLMs) interpret the symbol '{REDACTED}' as indicating omitted content, and therefore it is unlikely to significantly impact analyses involving LLMs. An analysis of the AMP SCZ Release 2 data revealed that only 0.17% ($SD = 0.0038$) of words were redacted, suggesting minimal potential impact on overall results. Interestingly, redactions were significantly more frequent in Open interviews ($M = 0.00278$, $SD = 0.000135$) than in PSYCHS interviews ($M = 0.000830$, $SD = 0.000099$), $t(19{,}997) = 11.853$, $p < 0.001$, indicating that participants disclosed more PII during Open interviews than during PSYCHS interviews.

*Quality control and transcript review process*. Once the transcripts are generated and stored on a TranscribeMe! server, they are downloaded back to the aggregate server data lake as part of the AVL pipeline. Here, further QC measures are applied, assessing aspects such as the percentage of redacted utterances, inaudible words (i.e., words that the transcribers could not identify), number of speakers, number of words per speaker, and the total number of conversational turns (see Table S1 for a full list of the variables and the range of acceptable values). Values outside the established acceptable ranges are highlighted in DPdash managed so that they can be easily detected by study staff.

As an additional layer of QC, the first ten transcripts from each site are sent back to the Box and Mediaflux cloud storage services for further review by staff at the data acquisition sites. These additional checks confirmed that the transcription service's PII/PHI redaction guidelines were applied reliably. If any instances of missed PHI/PII were identified during the review process, feedback was provided to TranscribeMe! to help prevent similar errors in the future.

## Speech acoustics processing

The processing of speech acoustics in this project is designed to provide a relatively full characterization of each speaker's acoustic, prosodic, and voice quality features. The extraction of these features is based on the diarized audio files (one for each speaker in the interviewer) generated from the open-ended and PSYCHS Zoom interviews.

*Acoustic preprocessing*. The audio files are first converted to WAV format at 44.1 KHz using the ffmpeg library. Praat (Version 6.3.17[25]; is then employed to detect silences and sounds, setting a minimum pitch threshold of 75 Hz and a silence threshold of −25 dB. Silences under 200 ms are not classified as pauses but are labeled as speech. Further, since the assignment of pauses by one speaker depends on the voice activity of the other speaker, no pause is assigned following an interruption when the other speaker is actively talking. While the interruption or speech sound is still considered a valid signal, it does not result in a subsequent pause.

The voice and pause signals from all participants are further aligned to ascertain which segments are relevant for feature extraction, utilizing binary step functions to avoid the misattribution of pauses. This process creates individual audio files for each speaker, as well as an audio file containing all participants present during the interview (see Fig. 3).

*Feature extraction/generation*. For the extraction of acoustic features, we utilize Praat v6.3.17[25] and openSMILE-python v2.4.2[26]. The feature set includes a combination of

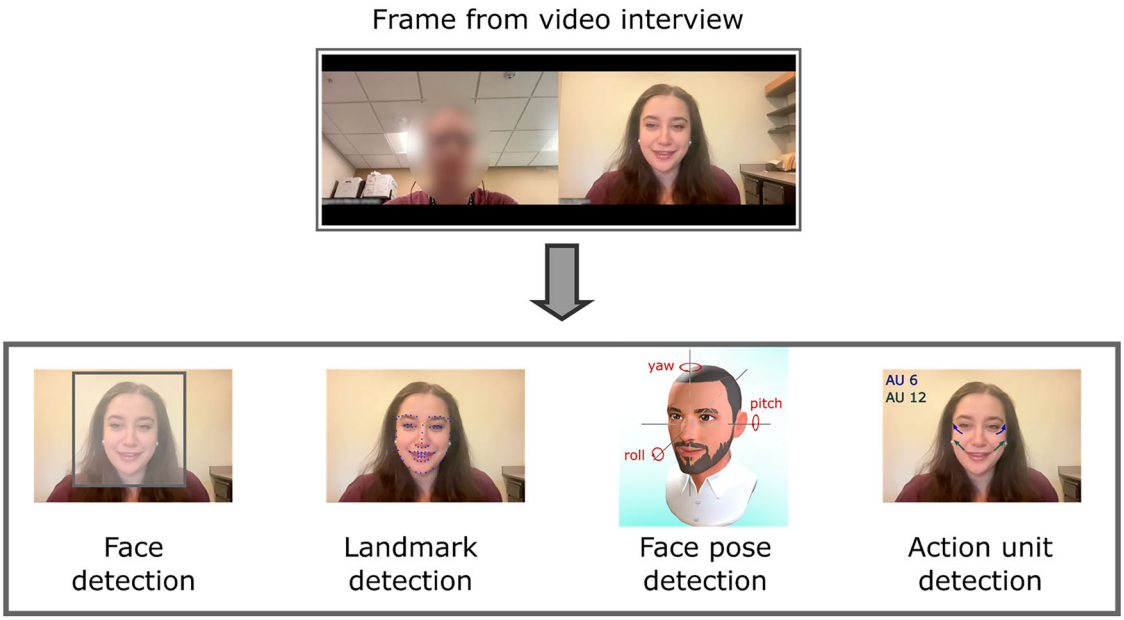


**Fig. 4** Face processing involves a sequence of four stages: face detection, landmark detection, face pose detection, and action unit detection.

ComParE_2016[27] and eGeMAPSv02[28], covering four domains: cepstral (e.g., mel frequency cepstral coefficients[MFCC]), spectral (e.g., harmonicity), prosodic (e.g., loudness variation), and voice quality (e.g., jitter). This composite feature set has proven effective in characterizing CHR individuals[17]. The extraction process is twofold. Initially, low-level descriptors (LLDs) are extracted every 10 ms in 20 ms windows, including labels for each participant based on their speaking time (see Fig. 3). This approach facilitates integration with other modalities, such as video features used to specify facial expressions. Secondly, global features, or 'functionals,' are extracted for each participant, representing summary statistics of the LLD features ($N = 6443$) over the entire duration of the interview.

Additional temporal features are computed, focusing on speech tempo given its relevance to emotion analysis and negative symptoms in CHR individuals[16], as well as the detection of syllables along with their duration[29]. These examinations involve analyzing the distribution of pause durations with respect to eight functional measures including median, interquartile range, 5th and 95th percentiles, skewness and kurtosis, and total number of samples[16]. Speech rate and articulation rate are also calculated, with and without pauses.

*Post-processing*. Upon completion of preprocessing and feature extraction, two sets of files are generated for each recording: one containing all LLD features ($N = 85$ features) for all participants in a speech task and a second CSV file for each participant with the functional features ($N = 6461$ features). This comprehensive analysis provides a detailed characterization of the acoustic features of participants' speech.

## Facial feature extraction

The extraction of facial features is designed to accurately capture and characterize the nuanced facial movements of participants during the interviews. This process, specifically applied to Zoom video files, is carried out through a four-step sequence, as illustrated in Fig. 4.

*Ethics declaration*. The individual depicted in the identifiable image included in this manuscript is an author of the paper and has provided informed consent for the publication of their image.

*Face detection*. The first step involves detecting faces in each video frame using MediaPipe's (version 0.9.2.1) Face Detection algorithm[30], an approach that supports multiple faces—a crucial feature for interviews with two or more participants. To ensure complete coverage of detected faces, we adjust the facial regions of interest (ROIs) by stretching their width and height through a scalar factor, thereby avoiding any improper cropping and ensuring full facial coverage.

*Landmark detection and face pose detection*. The extraction of facial landmarks and poses involves two processes. Firstly, a dense estimate of the 3D surface of the face is computed from 2D coordinates in the image plane. MediaPipe's Face Mesh algorithm[31] implementation of this process results in 468 landmarks. Second, a similarity transformation (translation, rotation, and isotropic scaling) is used to estimate the orientation of the face, and hence its pose.

*Action unit detection*. The last module in the pipeline estimates the intensity of facial muscle actions (action units), which are components of basic facial expressions and emotions[32]. This module is implemented using Py-Feat's (version 0.5.1) XGBoost model. Since Py-Feat modules work with a sparser landmark representation ($N = 68$), we provide a subset of the 468 detected landmarks from MediaPipe to Py-Feat.

*Algorithm choices*. The choice of face-processing software is guided by the need to balance efficiency and accuracy. While MediaPipe offers robust solutions for the initial stages of our pipeline, it does not offer a solution for the generation of facial action units. To fill this gap, we incorporated Py-Feat. Py-Feat could have been used for the initial stages of the pipeline, but its execution time is not optimal. We also evaluated OpenCV's solvePnP for face pose detection (https://opencv.org/). While solvePnP is fast, its performance is less reliable than other programs tested. Due to this and the fact that face pose is already embedded in MediaPipe's Face Mesh, we opted for the latter algorithm. Our pipeline thus integrates what we determined to be the best modules from different toolboxes.

*Integration of facial and speech features*. To effectively combine facial features with speech-based data, we use all of the frames in

the video to extract facial features in 40 ms windows, which is possible because the videos are recorded at a rate of 25 frames per second. The extracted data includes coordinates and dimensions of face ROIs, confidence levels for face detection, head orientation angles (pitch, roll, and yaw), and the intensity of action units, with the exception of action units 7 and 20 due to poor estimation. For each window, we also know the frame index, and the number of faces detected. The detailed analysis of features are aligned with features generated from language and speech acoustics to enable a cross-model characterization of the communicative context. It is worth mentioning that this cross-model integration is not available at our data submission to the NDA, but it could be straightforwardly achieved by a researcher with access to the data by using the provided timestamps.

## MITIGATING THREATS TO DATA QUALITY

The AVL processing pipeline is a complex system susceptible to various challenges. These challenges can range from recording issues like echoes and microphone malfunctions to procedural errors such as incorrect Zoom settings or file misplacements. To address and minimize these challenges, we implemented two primary strategies: comprehensive training and vigilant monitoring of the AVL pipeline.

### Training and certification

Training sessions for conducting Open and PSYCHS interviews are organized in a group format across all sites. These sessions cover not only the interview techniques but also standard data collection procedures as outlined in the SOP. Key aspects of this training include equipment setup, optimizing video and audio settings, and data uploading protocols.

Interviewers are instructed to use study identifiers instead of participant names during recordings to protect participant identities. They are also guided to pause rather than stop recordings during interviews to maintain continuity. For interviews conducted over multiple days, the pipeline considers only the first session. Even if the interview questions are continuous, sessions held on different days represent inherently distinct language samples, as symptoms and mood can fluctuate over time. Simply concatenating these sessions could negatively impact statistical analyses examining relationships between different parts of the interview.

Post-training, staff undergo individual certification processes, which assess their proficiency in interview techniques and adherence to SOP procedures. Knowledge of settings is verified through demonstrations, and open-ended interview style competency is evaluated using mock interviews scored against a specific rubric (see details in Supplemental Materials, Table S2). PSYCHS competency is assessed separately, as described in the companion paper in this special issue on clinical assessments[22].

### Manual and automated quality control

To safeguard data quality, we developed an automated organization, processing, and quality control of the data system. The code for this system flags major upload issues such as missing files from one of the modalities, incorrect file formats, or missing metadata, which are then corrected manually. If no fatal flaws are discovered, it maps the interview date to a study day corresponding to the number of days since the participant consented to study participation. This format prevents revealing potentially identifiable real-date information while maintaining an invertible mapping to the stored raw data. The QC variables calculated for language, speech acoustics, and facial processing are displayed on the DPdash system, using color coding for easy identification of potential issues. Regular summaries of these data are circulated among project staff for central monitoring. Additionally, weekly QC meetings involving the two research networks and the DPACC are held to review reports and troubleshoot and strategize solutions.

## ETHICAL CONSIDERATIONS

The collection and analysis of audiovisual data in the AMP SCZ study necessitates careful consideration of several ethical issues. These include privacy concerns, potential bias, partnerships with individuals and communities, data ownership, and maintaining equity and inclusion, especially in the face of possible biases in computational models.

### Privacy considerations

Given that the data will be stored at the National Institute of Mental Health Data Archive (NDA) and made available to the wider research community, there are considerable privacy concerns. To minimize the risk of identifying individuals, the raw audiovisual data are not stored. Instead, tools such as PRATT, OpenSMILE, MediaPipe, and Py-Feat are utilized to extract standard acoustic and facial expression variables. Although cepstral (vocal) features and facial landmarks could potentially identify individuals, reconstructing someone's identity from these variables is highly challenging. Furthermore, all transcripts processed by TranscribeMe! are thoroughly de-identified prior to archiving. The process is augmented by manual spot-checks at each site to ensure that any PHI or PII is effectively redacted.

As advancements in computational tools emerge, revisiting some of the original raw audio and video files for further analysis could be considered. Such re-analyses could be achieved through direct collaboration with the research networks. Alternatively, to allow for such future analyses, we aim to implement in coordination with the NDA a privacy-preserving Federated Learning approach[33]. Such an approach would allow for additional fine-tuning of the features extracted and could be used to improve the models by providing only abstract information as in the current approach, but in an iterative learning framework[34].

### Partnership and data ownership

The AMP SCZ consortium includes a diverse group of stakeholders, including individuals with lived experiences (see also in this volume[35]: and partnerships with organizations such as the National Alliance on Mental Illness (NAMI). These collaborations have been instrumental in shaping the design of our study, underscoring our dedication to inclusive research practices. To ensure data transparency and availability, we publish SOPs on the AMP SCZ website and facilitate access to the data via the NDA. Moreover, participants are provided access to their personal data, including smartphone data and daily diaries, further emphasizing our commitment to transparency and participant engagement.

## GRAMMATICAL PROFILING OF THREE LANGUAGE SAMPLE TYPES TO DISTINGUISH CHR FROM CC INDIVIDUALS

Research has consistently demonstrated that the language produced by individuals with schizophrenia differs systematically from that of healthy controls[2,36,37]. Two of the most widely replicated linguistic features are the overuse of pronouns—particularly first-person singular forms—and reduced syntactic complexity, especially a lower frequency of embedded clauses[38–41]. These patterns have been observed across a broad range of clinical populations. Pronoun overuse has been reported in individuals with schizophrenia[42–44], schizophrenia-spectrum disorders[45], individuals experiencing first-episode psychosis (FEP[46–48], and individuals with formal thought disorder[49]. This convergence suggests that elevated pronoun use—especially of the first-person singular—may represent a transdiagnostic marker

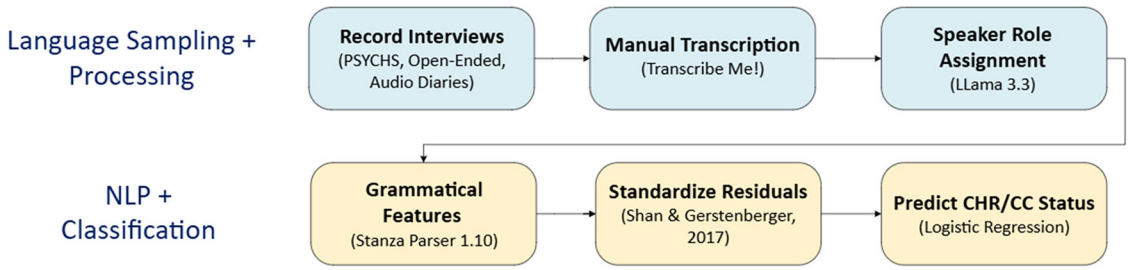


**Fig. 5** Pipeline for extracting grammatical features, syntactic dependencies, and parts of speech from language samples to assess their ability to distinguish CHR individuals from CCs.

across the psychosis continuum. Likewise, reductions in syntactic complexity have been documented in chronic schizophrenia[50–52], schizophrenia-spectrum disorders[38], first-episode psychosis[53], and in formal thought disorder[54,55]. These findings indicate that disruptions in grammatical organization may serve as an additional candidate biomarker of psychosis risk.

Despite the consistency of these findings, most studies have focused on individuals in later stages of illness. For linguistic features to function as early indicators of psychosis, it is essential to determine whether they are also present in CHR—those who exhibit early signs of vulnerability but have not yet transitioned to full psychosis.

A notable study by Corcoran et al[12]. found that pronoun usage predicted conversion among CHR individuals. Interestingly, rather than overusing pronouns, CHR converters showed a reduction in overall pronoun use. Similar reductions have been reported in other studies[44,45,56]. Furthermore, inconsistencies remain regarding whether elevated pronoun use is limited to first-person forms or extends to second- and third-person usages[39,40,57]. Inconsistencies have also been observed in measures of syntactic complexity. These divergences raise two key issues. First, while increased pronoun use and reduced syntactic complexity are frequently observed, they are not universal[39]. Second, such discrepancies may reflect differences in clinical subgroups—or variation in language elicitation methods.

NLP techniques, including automated grammatical parsing, offer scalable tools for examining referential and syntactic patterns across large language samples. In the present study, we apply these tools to investigate pronoun use and syntactic complexity in individuals at CHR for psychosis. Unlike prior studies, which primarily focus on individuals with established diagnoses, our emphasis is on those at elevated risk but who have not yet converted. Detecting these features in CHR individuals—even in attenuated form—would support their use as early linguistic markers of illness onset. Crucially, our study also allows us to examine the influence of task type on the measurement of these markers. Prior studies of pronoun use have employed a range of elicitation methods, including spontaneous monologues[42], emotionally evocative narratives[43], open-ended interviews[12], free speech[45], picture description tasks[46], autobiographical and dream narratives[44,58], and structured interviews[56]. Studies of syntactic complexity have used similarly diverse tasks, including structured interviews[41,50,59], narrative retellings[55], picture-sequence descriptions[38,51,53,54,60], and free-form conversations[52].

These tasks differ in cognitive demands, emotional salience, and discourse structure, all of which may influence language production and the reliability of derived linguistic markers. A strength of the present study is its use of multiple language-sampling contexts—including structured interviews, open-ended interviews, and audio diaries—enabling a more rigorous test of the robustness and task sensitivity of linguistic features associated with psychosis risk.

## Methods

*Participants*. Language samples were collected from participants enrolled in the Accelerating Medicines Partnership® Schizophrenia (AMP® SCZ) project. All participants completed PSYCHS assessments, open-ended interviews, and audio diaries. Transcripts were obtained from AMP SCZ Public Data Release 3.0 and are available through the NIMH Data Archive (NDA) via the AMP SCZ public website. The dataset includes 172 individuals identified as being CHR and 43 CC participants. All participants provided oral and written informed consent in accordance with institutional review board guidelines. The project was approved by the governing institutional review board at each site and is registered at clinicaltrials.gov (NCT05905003).

*Procedures*. The full pipeline for language sampling, processing, and subsequent NLP and ML classification is depicted in Fig. 5. The process begins with the collection of recorded language samples. Manual transcription is conducted on the first 30 min of the interview by trained human coders by the HIPAA-compliant transcription service company TranscribeMe!. As shown previously in Fig. 2, the transcripts are partitioned by speaker. The person who speaks first is labeled S1, and the person who speaks second is labeled S2, and so on. In further analyses, the labels S1 and S2 are identified with respect to their role in the conversation. Typically, S1 is the interviewer and S2 is the interviewee, but this is not always the case. The process of assigning conversational roles is conducted using the LLM, LLaMA 3, based on 70 billion parameters[61], run locally on an offline machine (i.e., no data were shared with Meta). LLaMA 3 is provided with sequences of conversation between speakers and asked to determine whether S1 is the interviewer and S2 is the interviewer or S1 is the interviewee and S2 is the interviewee. LLaMA 3's answers to these yes-no questions are consequently used to assign the conversational roles of interviewer and interviewee to the labels S1 and S2. When LLaMA 3's judgments were compared against those of a human, it performed with greater than 98% accuracy on a set of 200 transcripts.

After assigning speaker-role information, we proceeded with the extraction of linguistic features. Transcribed speech samples from both CHR and CC participants were processed using the Stanza NLP toolkit[62]. The feature set included all syntactic and lexical variables identified by the parser that appeared with non-zero counts in at least half of the participants, resulting in a total of 102 linguistic features. Of these, 77 features were derived from the Universal Dependencies (UD) framework, a cross-linguistic initiative that provides a consistent set of syntactic categories and dependency relations for grammatical annotation[63]. The remaining 24 features were based on the Penn Treebank (PTB) part of speech tagset, which was developed to annotate English corpora with fine-grained lexical and syntactic categories[64]. A complete list of the extracted features, along with their corresponding tags and illustrative examples, is provided in Appendix A.

To assess whether the frequency of linguistic features differs between CHR and CC participants, it is essential to account for

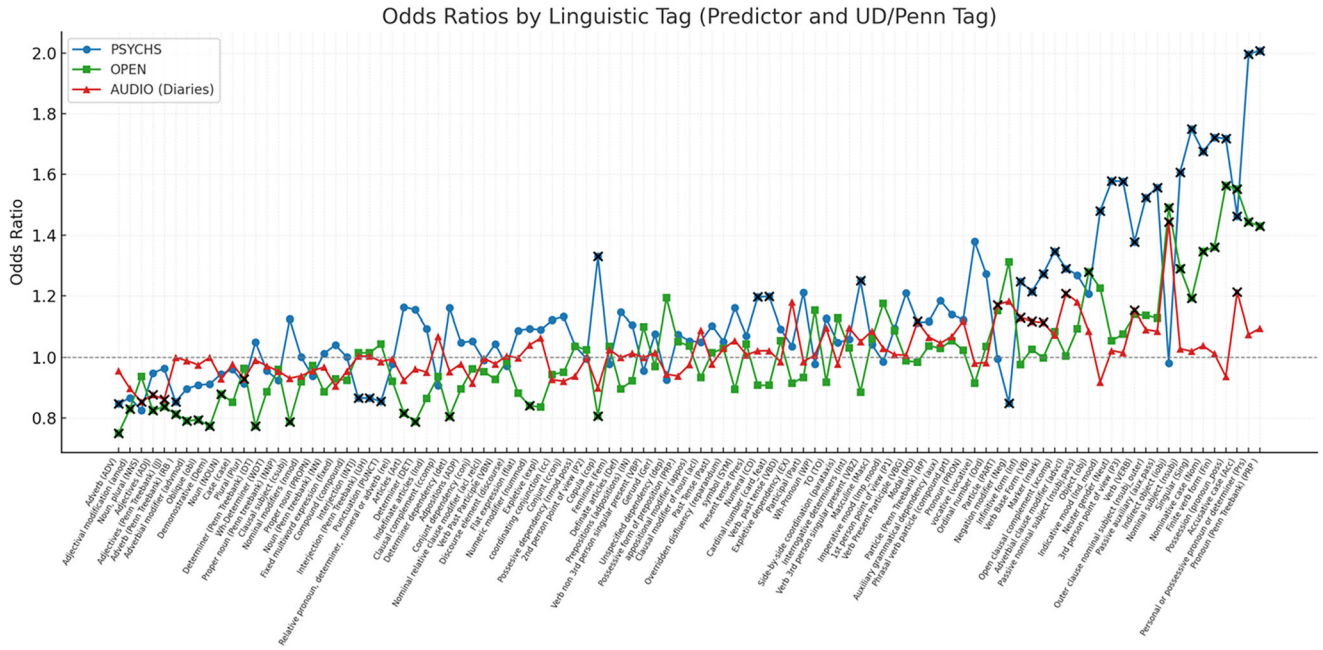


**Fig. 6 ORs for grammatical features, syntactic dependencies, and parts of speech across three language sample types—PSYCHS interviews (blue), open-ended interviews (green), and audio diaries (audio).** An OR of 1 indicates no association with CHR status. ORs greater than 1 suggest that higher feature values are associated with increased odds of being classified as CHR, while ORs less than 1 indicate a negative association. Features marked with an "x" were statistically significant based on the Wald *z*-test.

variation in interview length. This was accomplished by computing standardized residuals, a widely used approach for evaluating cell-level differences in contingency tables. In this context, the frequency matrix is structured with participants as rows and linguistic features as columns, where each cell represents the observed count of a specific feature for a given participant. To normalize these counts, we first compute the raw residuals by subtracting the expected frequency from the observed frequency in each cell. However, raw residuals are not directly comparable across cells, as their magnitude depends on the scale of the expected count—an absolute difference of 10 may be meaningful in one context but negligible in another. To address this, each raw residual is divided by the square root of its expected frequency, yielding a standardized residual. This transformation adjusts for cell-wise variability and allows residuals to be compared on a common scale. Under the null hypothesis of independence, these standardized residuals approximately follow a standard normal distribution[65], enabling the use of parametric statistical methods in subsequent analyses.

With standardized residuals computed, logistic regression was used to evaluate how well each linguistic variable, on its own, could predict whether an individual was classified as CHR or CC. For each feature, an odds ratio (OR) was calculated to indicate the change in the odds of being CHR associated with a one-unit increase in that feature. An OR greater than 1 suggests that higher values of the feature increase the likelihood of CHR classification, while an OR less than 1 indicates a decreased likelihood of being CHR (and thus a higher likelihood of being CC). The statistical significance of each feature was assessed using the Wald *z*-test applied to the regression coefficient. A significant result provides evidence that the feature is reliably associated with CHR status.

## Results and discussion

Figure 6 illustrates how specific linguistic features relate to the likelihood of an individual being classified as CHR for psychosis. The figure shows ORs for grammatical features, syntactic dependencies, and parts of speech across three different types of language samples: PSYCHS interviews (blue), Open interviews (green), and audio diaries (orange). An OR of 1 suggests no association with CHR status and values greater than 1 indicate that higher values of a given feature increase the odds of being CHR, while values less than 1 imply that higher values of the feature are associated with lower odds of CHR status and increased odds of being CC status.

As predicted, CHR individuals produced more pronouns than CC. This increase was not limited to a single type of pronoun but extended across a broad range of grammatical features, suggesting a widespread alteration in how CHR individuals refer to themselves and others in speech. Across at least two elicitation methods—PSYCHS interviews, Open interviews, and audio diaries—pronoun-related features yielded ORs greater than 1. In other words, higher pronoun usage was more typical of CHR individuals and reduced the likelihood of classification as CC. Several specific linguistic features supported this pattern. The personal pronoun (PRP) category showed strong effects in both the PSYCHS and Open interviews ($p < 0.00001$ and $p = 0.000315$, respectively), with a marginal trend in the Diaries ($p = 0.09923$). Similarly, the personal/possessive determiner (Prs) category was significantly elevated in CHR speech in the PSYCHS ($p < 0.00001$) and Open interviews ($p = 0.00029$), though not in the Diaries ($p = 0.19902$). Features marking grammatical case reinforced the trend. Use of nominative case (Nom) pronouns—typically indicating subject position—was significantly higher in CHR speech in both PSYCHS ($p < 0.00001$) and Open interviews ($p = 0.00271$), but not in Diaries ($p = 0.52264$). Accusative case (Acc) pronouns—marking object position—were significantly elevated in all three contexts: $p = 0.00082$ (PSYCHS), 00059 (Open), and 0.00267 (Diary). Other features added support, especially in the PSYCHS samples. Neuter gender (Neut) and third-person perspective (P3) both reached strong significance in PSYCHS ($p < 0.00001$ for each), though effects were weaker or nonsignificant in the Open and Diary samples. Not all pronoun-related features followed this pattern. First-person (P1) and second-person (P2) pronouns did not differ reliably between CHR and CC groups. For P1, *p*-values were 0.18292 (PSYCHS), 0.09683 (Open), and 0.83267 (Diary); for P2,

**Table 2.** Mean number of words, sentences, and words per sentence for each interview with associated standard deviations.

| | Group | *N* | Words (M ± SE) | Sentences (M ± SE) | Words/sent (M ± SE) |
|---|---|---|---|---|---|
| PSYCHS | CHR | 172 | 2192.81 ± 103.12 | 256.06 ± 8.76 | 8.51 ± 0.391 |
| | CC | 43 | 1209.26 ± 160.04 | 185.05 ± 12.45 | 5.55 ± 0.464 |
| Open-ended | CHR | 172 | 2249.22 ± 118.09 | 186.26 ± 6.23 | 12.28 ± 0.372 |
| | CC | 43 | 2556.12 ± 137.69 | 202.05 ± 10.72 | 12.98 ± 0.427 |
| Audio diary | CHR | 172 | 6373.29 ± 1014.64 | 507.14 ± 79.36 | 12.85 ± 0.273 |
| | CC | 43 | 4140.14 ± 1265.84 | 347.81 ± 111.41 | 11.70 ± 0.658 |

none of the effects reached significance. This finding diverges from meta-analyses by Elleuch et al. (2025), which reported increased first-person pronoun use among individuals with schizophrenia[39]. Additionally, the UD PRON tag showed only marginal significance in the Diary samples ($p = 0.05234$) and was nonsignificant in PSYCHS ($p = 0.13901$) and Open ($p = 0.75338$) interviews. Despite some inconsistencies across individual features and contexts, the overall direction of effects across multiple grammatical categories—PRP, Prs, Nom, Acc, Neut, and P3—was largely consistent: greater pronoun use was associated with increased psychosis risk. This convergence suggests that alterations in pronoun usage may reflect deeper disruptions in perspective-taking and discourse structure, potentially serving as a meaningful linguistic marker of emerging psychopathology.

Contrary to expectations, we found evidence for higher—not lower—syntactic complexity among CHR individuals. However, as we discuss below, this effect likely reflects differences in the amount of language elicited from CHR and CC participants across the different elicitation methods. The strongest indicators of increased complexity were grammatical features associated with syntactic embedding. Four features stood out: markers (mark), adverbial clausal modifiers (advcl), infinitive forms (Inf), and open clausal complements (xcomp). In addition, increases in the use of verbs (VERB & VB) offered indirect support for heightened syntactic complexity, as embedded structures require additional verb phrases. In all cases, ORs exceeded 1.0, indicating that more frequent use of these features was associated with higher odds of CHR classification.

For example, marker dependencies (mark)—significant in both the PSYCHS ($p = 0.0107$) and audio diary conditions ($p = 0.02082$)—typically involve subordinating conjunctions such as "that," as in "He says that the party will be canceled." Adverbial clausal modifiers (advcl), which were also significant in PSYCHS ($p = 0.02015$) and diaries ($p = 0.00967$), appear in constructions like "She left because she was tired," where the embedded clause modifies the main clause. Infinitive constructions (Inf) were elevated in CHR speech in both PSYCHS ($p = 0.04103$) and audio diaries ($p = 0.00769$), as in "She hopes to win the competition." That same sentence also illustrates an open clausal complement (xcomp)—which reached significance in PSYCHS ($p = 0.00606$)—where the verb "win" is syntactically dependent on the matrix verb "hopes." Finally, CHR individuals used more verbs than CC participants, as indicated by increased frequencies of the VERB part of speech tag in PSYCHS ($p = 0.0088$) and audio diaries ($p = 0.04265$), and of the base verb form (VB) tag in PSYCHS ($p = 0.04552$) and diaries ($p = 0.01256$).

In contrast, several other embedding-related features, where differences might have been expected,showed no evidence of differential usage between CHR and CC individuals. These included clausal subjects (csubj), nominal relative clauses (acl:relcl), relative pronouns (rel), and clausal complements (ccomp). The one exception was csubj, which reached significance in the open interviews ($p = 0.03548$) but not in the PSYCHS ($p = 0.43062$) or audio diary conditions ($p = 0.41783$). Notably, all of these non-significant features were associated with ORs below 1.0. Had they been significant, they would have suggested reduced syntactic complexity among CHR speakers. One possibility is that these features are beginning to decline in CHR speech—consistent with prior reports of syntactic simplification—but have not yet diminished enough to yield statistically robust effects.

Where significant differences did emerge, the grammatical profile pointed to increased syntactic complexity in CHR speech. However, this interpretation is complicated by differences in speech quantity. During the structured PSYCHS interviews, CHR participants were considerably more talkative (see Table 2), producing an average of 2193 words and 256 sentences, compared to 1209 words and 185 sentences for controls. This greater output naturally inflated their average sentence length—with 8.51 words per sentence for CHR vs 5.55 words per sentence for controls—with all group differences highly significant ($p < 0.0001$). In contrast, CHR and CC participants produced similar amounts of language during the open-ended interviews: approximately 2249 words and 186 sentences for CHR, and 2556 words and 202 sentences for CC. With output essentially matched, no group differences in sentence length or syntactic complexity were observed. The audio diaries fell in between. Although word and sentence counts were comparable, CHR participants produced slightly longer sentences (12.9 vs 11.7 words), an effect that was only marginally significant ($p = 0.071$).

Greater syntactic complexity in CHR speech seems to arise chiefly when CHR and control participants differ in how much they say. During PSYCHS interviews, CCs often answered probes like "Have you ever felt suspicious of other people?" with a terse "No." The parser tagged these single-word replies as interjections (INTJ, e.g., uh), accounting for their higher frequency in control speech in the PSYCHS ($p < 0.0001$). CHR participants, by contrast, typically offered longer, more nuanced answers; that extra verbal output appeared to introduce more embedded clauses and other signs of syntactic elaboration.

Not every grammatical feature was amplified in CHR speech; several, in fact, declined. Adjectives appeared less often: in the Universal tagset (ADJ) this drop reached significance in the Open interviews ($p = 0.021$) and the audio diaries ($p = 0.015$), and the same pattern held for the PTB tag JJ ($p = 0.034$ and 0.008, respectively). The adjectival-modification dependency amod echoed the decline in the Open interviews ($p = 0.013$). Adverbs showed a similar contraction. Both the Universal ADV tag and the PTB RB tag fell significantly in PSYCHS ($p = 0.032$ for both) and even more sharply in the open interviews ($p < 0.001$), while the advmod dependency confirmed the reduction in the open interviews ($p = 0.006$). Nouns, too, thinned out: overall noun frequency (NOUN) decreased in the open interviews ($p = 0.042$) and trended downward in the diaries, and plural nouns (NNS) were notably lower in PSYCHS ($p = 0.050$) and in the diaries ($p = 0.017$). Together, these converging reductions—in nouns, their adjectival modifiers, and adverbs—point to a broader attenuation of descriptive detail in CHR speech.

Overall, there was a moderate degree of consistency in the prevalence of linguistic features across the different language elicitation methods. As shown in Fig. 6, ORs from the PSYCHS and Open-ended interviews were strongly correlated, $r(99) = 0.577$, $p < 0.0001$, indicating substantial agreement. A similarly strong correlation was observed between Open-ended interviews and Audio diaries, $r(99) = 0.468$, $p < 0.0001$. In contrast, the correlation between PSYCHS and Audio diaries was weaker, though still statistically significant, $r(100) = 0.217$, $p = 0.029$. These findings suggest that while there is meaningful overlap in the linguistic patterns captured across tasks, the method of language elicitation can substantially influence which features are detected.

## CONCLUSION

The language sampling methodologies implemented in the AMP SCZ demonstrate how large-scale language collection efforts can be successfully carried out. The initial analysis of data generated through this initiative illustrates how grammatical profiling may be used to identify individuals at risk for psychosis.

A particularly robust and consistent finding was the increased use of pronouns among CHR participants. This elevation remained statistically significant even after adjusting for overall speech volume, suggesting that it was not merely a by-product of verbosity. Notably, the increase extended beyond first-person pronouns. CHR individuals showed elevated usage across a broad set of grammatical categories—including personal, possessive, nominative, accusative, neuter, and third-person forms—indicating a generalized disruption in referential processes and perspective-taking rather than a narrowly focused amplification of self-reference.

In contrast, the evidence for increased syntactic complexity among CHR individuals was more context dependent. While initial analyses pointed to greater grammatical elaboration, particularly through embedded clauses and subordinating constructions, this effect was largely confined to settings where CHR participants also spoke more. When verbal output was more closely matched, as in the open-ended interviews, no meaningful differences in complexity were observed. These findings suggest that syntactic complexity may be driven more by task structure and speech volume than by inherent grammatical differences.

Together, these results support two central conclusions. First, pronoun overuse emerges as a stable and context-independent linguistic marker of psychosis risk. Its presence across elicitation contexts and independence from speech quantity suggest it may reflect a fundamental shift in language use during early illness stages. Second, syntactic complexity should be interpreted with caution. Although it may appear elevated in certain contexts, this effect is likely a by-product of greater verbal engagement rather than an intrinsic linguistic feature of CHR individuals. As psychosis progresses and speech becomes more limited, syntactic complexity may diminish, consistent with previous findings in chronic psychotic disorders. Further, participants included in this study contributed all three language samples, including daily diaries. However, those who declined to complete diaries may represent a subgroup characterized by reduced verbal output, and, in turn, reduced complexity, which was not captured in this analysis.

Beyond the rise in pronouns and certain embedded-clause constructions, CHR speech also shows marked decreases in content words—most notably adjectives and adverbs, and to a lesser extent nouns. These changes are likely interdependent: fewer nouns reduce opportunities for adjectives to serve as modifiers and may encourage substitution with neuter pronouns such as it in place of missing referents. The combined decline in adjectives, adverbs, and nouns—paired with a surge in pronouns—suggests a shift away from words whose meaning is inherent (e.g., nouns, adjectives, and adverbs) toward words whose meaning depends heavily on context (e.g., pronouns). Even the observed increase in syntactic complexity may be partly compensatory, as speakers lengthen sentences and embed more clauses to express ideas that would otherwise rely on a richer stock of precise content words.

Ultimately, the findings underscore the potential of everyday language as a non-invasive and scalable marker of emerging psychopathology. Advances in NLP and AI now enable the identification of such subtle features. By combining robust collection protocols with computational analysis, this study demonstrates the feasibility of using speech as a clinically informative biomarker—one that may help identify individuals at risk for psychosis well before overt symptoms emerge.

## DATA AVAILABILITY

Data used in this study are accessible through scheduled releases at the NDA AMP SCZ Data Repository (https://nda.nih.gov/ampscz).

Received: 28 August 2024; Accepted: 21 August 2025;
Published online: 15 October 2025

## REFERENCES

1. Andreasen N. C. Scale for the assessment of thought, language, and communication (TLC). *Schizophr. Bull.* https://doi.org/10.1093/schbul/12.3.473 (1986).
2. Andreasen, N. C. Thought, language, and communication disorders: II. Diagnostic significance. *Arch. Gen. Psychiatry* **36**, 1325–1330 (1979).
3. Andreasen, N. C. Thought, language, and communication disorders. I. Clinical assessment, definition of terms, and evaluation of their reliability. *Arch. Gen. Psychiatry*. **36**, 1315–1321 (1979).
4. Andreasen, N. C. Scale for the assessment of negative symptoms (SANS). *Br J Psychiatry*. **155**, 53–58 (1989).
5. Bearden, C. E., Wu, K. N., Caplan, R. & Cannon, T. D. Thought disorder and communication deviance as predictors of outcome in youth at clinical high risk for psychosis. *J. Am. Acad. Child Adolesc. Psychiatry* **50**, 669–680 (2011).
6. Gooding, D., Ott, S., Roberts, S. & Erlenmeyer-Kimling, L. Thought disorder in mid-childhood as a predictor of adulthood diagnostic outcome: findings from the New York High-Risk Project. *Psychol. Med.* **43**, 1003–1012 (2013).
7. Kay, S. R., Fiszbein, A. & Opler, L. A. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr. Bull.* **13**, 261–276 (1987).
8. Woods S. W., Walsh B. C., Powers A. R. & McGlashan T. H. Reliability, validity, epidemiology, and cultural variation of the structured interview for psychosis-risk syndromes (SIPS) and the scale of psychosis-risk symptoms (SOPS). In: *Handbook of Attenuated Psychosis Syndrome Across Cultures* 85–113 (Springer International Publishing, 2019).
9. Yung, A. R. et al. Prediction of psychosis. A step towards indicated prevention of schizophrenia. *Br. J. Psychiatry Suppl.* **172**, 14–20 (1998).
10. Johnston, M. H. et al. Scoring manual for the thought disorder index. *Schizophr. Bull.* **12**, 483 (1986).
11. Kircher, T. et al. A rating scale for the assessment of objective and subjective formal thought and language disorder (TALD). *Schizophr. Res.* **160**, 216–221 (2014).
12. Corcoran, C. M. et al. Prediction of psychosis across protocols and risk cohorts using automated language analysis. *World Psychiatry*. https://doi.org/10.1002/wps.20491 (2018).
13. Bedi, G. et al. Automated analysis of free speech predicts psychosis onset in high-risk youths. *npj Schizophrenia*. https://doi.org/10.1038/npjschz.2015.30 (2015).
14. Rezaii, N., Walker, E. & Wolff, P. A machine learning approach to predicting psychosis using semantic density and latent content analysis. *npj Schizophrenia***5**, 1–12 (2019).
15. Spencer T., et al. Lower speech connectedness linked to incidence of psychosis in people at clinical high risk. *Schizophr. Res.* https://doi.org/10.1016/j.schres.2020.09.002 (2020).
16. Stanislawski, E. R. et al. Negative symptoms and speech pauses in youths at clinical high risk for psychosis. *npj Schizophrenia* **7**, 1–3 (2021).
17. Agurto, C. et al. Analyzing acoustic and prosodic fluctuations in free speech to predict psychosis onset in high-risk youths. In: *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* 5575–5579 (IEEE, 2020).
18. Gupta, T., Haase, C. M., Strauss, G. P., Cohen, A. S. & Mittal, V. A. Alterations in facial expressivity in youth at clinical high-risk for psychosis. *J. Abnorm. Psychol.* **128**, 341–351 (2019).
19. Loch, A. A. et al. Detecting at-risk mental states for psychosis (ARMS) using machine learning ensembles and facial features. *Schizophr. Res.* **258**, 45–52 (2023).

20. Agurto, C. et al. Are language features associated with psychosis risk universal? A study in Mandarin-speaking youths at clinical high risk for psychosis. *World Psychiatry.* **22**, 157–158 (2023).
21. Yung, A. R. et al. Mapping the onset of psychosis: the comprehensive assessment of at-risk mental states. *Aust. N Z J Psychiatry* **39**, 964–971 (2005).
22. Addington, J. et al. Sample ascertainment and clinical outcome measures in the accelerating medicines partnership® schizophrenia program. *Schizophrenia***11**, 54 (2025).
23. Davidson, L. Phenomenological research in schizophrenia: From philosophical anthropology to empirical science. *J. Phenomenol. Psychol.* **25**, 104–130 (1994).
24. Wigman, J. T. W. et al. Digital health technologies in the accelerating medicines partnership® schizophrenia program. *Schizophrenia* **11**, 83 (2025).
25. Boersma, P. & Weenink, D. PRAAT: doing phonetics by computer [computer program]. http://www.praat. Available via https://scholar.google.com/citations?view_op=view_citation&hl=en&citation_for_view=9v4hT2kAAAAJ:vfT5ieZw1WcC.
26. Eyben, F., Wöllmer, M. & Schuller, B. Opensmile: the Munich versatile and fast open-source audio feature extractor. In: *Proceedings of the 18th ACM International Conference on Multimedia. MM '10* 1459–1462 (Association for Computing Machinery, 2010).
27. Schuller, B. et al. The INTERSPEECH 2016 computational paralinguistics challenge: deception, sincerity & native language. In: *Interspeech* 2001–2005 (ISCA, 2016).
28. Eyben, F. et al. The Geneva minimalistic acoustic parameter set (GeMAPS) for voice research and affective computing. *IEEE Transac. Affective Comput.* **7**, 190–202 (2016).
29. de Jong, N. H. & Wempe, T. Praat script to detect syllable nuclei and measure speech rate automatically. *Behav. Res. Methods* **41**, 385–390 (2009).
30. Bazarevsky, V., Kartynnik, Y., Vakunov, A., Raveendran, K. & Grundmann, M. BlazeFace: sub-millisecond neural face detection on mobile GPUs. Preprint at https://api.semanticscholar.org/CorpusID:195886588 (2019).
31. Kartynnik, Y., Ablavatski, A., Grishchenko, I. & Grundmann, M. Real-time facial surface geometry from monocular video on mobile GPUs. Preprint at https://api.semanticscholar.org/CorpusID:196831662 (2019).
32. Ekman, P. & Friesen, W. V. *Facial Action Coding System: A Technique for the Measurement of Facial Movement* (Consulting Psychologists Press, 1978).
33. Brendan McMahan H., Ramage D., Talwar K. & Zhang L. Learning differentially private recurrent language models. Preprint at http://arxiv.org/abs/1710.06963 (2017).
34. Kaissis, G. et al. End-to-end privacy preserving deep learning on multi-institutional medical imaging. *Nat. Mach. Intell.* **3**, 473–484 (2021).
35. Asgari-Targhi, A. et al. Bridging science and hope: integrating and communicating lived experience in accelerating medicines partnership® schizophrenia program. *Schizophrenia* **11**, 57 (2025).
36. Covington, M. A. et al. Schizophrenia and the structure of language: the linguist's view. *Schizophr. Res.* **77**, 85–98 (2005).
37. Hinzen, W. & Rosselló, J. The linguistics of schizophrenia: thought disturbance as language pathology across positive symptoms. *Front. Psychol.* **6**, 971 (2015).
38. Schneider, K. et al. Syntactic complexity and diversity of spontaneous speech production in schizophrenia spectrum and major depressive disorders. *Schizophrenia* **9**, 35 (2023).
39. Elleuch, D., Chen, Y., Luo, Q. & Palaniyappan, L. Speaking of yourself: a meta-analysis of 80 years of research on pronoun use in schizophrenia. *Schizophr. Res.* **279**, 22–30 (2025).
40. Elleuch, D., Chen, Y., Luo, Q. & Palaniyappan, L. Relationship between grammar and schizophrenia: a systematic review and meta-analysis. *Commun. Med.* **5**, 235 (2025).
41. de Boer, J. N., Voppel, A. E., Brederoo, S. G., Wijnen, F. N. K. & Sommer, I. E. C. Language disturbances in schizophrenia: the relation with antipsychotic medication. *NPJ Schizophr.* https://doi.org/10.1038/S41537-020-00114-3 (2020).
42. Fairbanks, H. II. The quantitative differentiation of samples of spoken language. *Psychol. Monogr.* **56**, 17–38 (1944).
43. Buck, B. & Penn, D. L. Lexical characteristics of emotional narratives in schizophrenia: relationships with symptoms, functioning, and social cognition. *J. Nerv. Ment. Dis.* **203**, 702–708 (2015).
44. Chaves, M. F., Rodrigues, C., Ribeiro, S., Mota, N. B. & Copelli, M. Grammatical impairment in schizophrenia: an exploratory study of the pronominal and sentential domains. *PLoS One* **18**, e0291446 (2023).
45. Tang, S. X. et al. Natural language processing methods are sensitive to sub-clinical linguistic differences in schizophrenia spectrum disorders. *NPJ Schizophr.* **7**, 25 (2021).
46. Mackinley, M., Chan, J., Ke, H., Dempster, K. & Palaniyappan, L. Linguistic determinants of formal thought disorder in first episode psychosis. *Early Interv. Psychiatry.* **15**, 344–351 (2021).
47. Figueroa-Barra, A. et al. Automatic language analysis identifies and predicts schizophrenia in first-episode of psychosis. *Schizophrenia* **8**, 53 (2022).
48. Morgan, S. E. et al. Natural language processing markers in first episode psychosis and people at clinical high-risk. *Transl. Psychiatry* **11**, 630 (2021).
49. Çokal, D. et al. Referential noun phrases distribute differently in Turkish speakers with schizophrenia. *Schizophr. Res.* **259**, 104–110 (2023).
50. Thomas, P. et al. Speech and language in first onset psychosis differences between people with schizophrenia, mania, and controls. *Br. J. Psychiatry* **168**, 337–343 (1996).
51. Fraser, W. I., King, K. M., Thomas, P. & Kendell, R. E. The diagnosis of schizophrenia by language analysis. *Br. J. Psychiatry.* **148**, 275–278 (1986).
52. Morice, R. & McNicol, D. The comprehension and production of complex syntax in schizophrenia. *Cortex.* **21**, 567–580 (1985).
53. Silva, A. M. et al. Syntactic complexity of spoken language in the diagnosis of schizophrenia: a probabilistic Bayes network model. *Schizophr. Res.* **259**, 88–96 (2023).
54. Çokal, D. et al. The language profile of formal thought disorder. *NPJ Schizophr.* **4**, 18 (2018).
55. Sevilla, G. et al. Deficits in nominal reference identify thought disordered speech in a narrative production task. *PLoS One* **13**, e0201545 (2018).
56. Arslan, B. et al. Computational analysis of linguistic features in speech samples of first-episode bipolar disorder and psychosis. *J. Affect. Disord.* **363**, 340–347 (2024).
57. Fineberg, S. K. et al. Self-reference in psychosis and depression: a language marker of illness. *Psychol. Med.* **46**, 2605–2615 (2016).
58. Lundin, N. B., Cowan, H. R., Singh, D. K. & Moe, A. M. Lower cohesion and altered first-person pronoun usage in the spoken life narratives of individuals with schizophrenia. *Schizophr. Res.* **259**, 140–149 (2023).
59. Dalal, T. C. et al. Speech based natural language profile before, during and after the onset of psychosis: a cluster analysis. *Acta Psychiatr. Scand.* **151**, 332–347 (2025).
60. Panikratova, Y. R. et al. Executive regulation of speech production in schizophrenia: a pilot neuropsychological study. *Neurosci. Behav. Physiol.* **51**, 415–422 (2021).
61. Touvron, H. et al. Llama: open and efficient foundation language models. Preprint at https://arxiv.org/abs/2302.13971 (2023).
62. Qi, P., Zhang, Y., Zhang, Y., Bolton, J. & Manning, C. D. Stanza: a Python natural language processing toolkit for many human languages. In: *Proc the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations* 101–108 (Association for Computational Linguistics, 2020).
63. de Marneffe, M. C., Manning, C. D., Nivre, J. & Zeman, D. Universal dependencies. *Comput. Linguist.* **47**, 1–54 (2021).
64. Marcus, M. P., Santorini, B. &, Marcinkiewicz, M. A. *Building a Large Annotated Corpus of English: The Penn Treebank* (ed. Hirschberg, J.) 313–330 (MIT Press, 1993).
65. Shan, G. & Gerstenberger, S. Fisher's exact approach for post hoc analysis of a chi-squared test. *PLoS One* **12**, e0188709 (2017).

## ACKNOWLEDGEMENTS

A full list of AMP SCZ members and affiliations can be found at https://www.ampscz.org/members/ and within the Supplementary File. The Accelerating Medicines Partnership® in Schizophrenia (AMP SCZ) is a public-private partnership managed by the Foundation for the National Institutes of Health. The AMP SCZ research program is funded through contributions from both public and private partners, including the National Institute of Mental Health (NIMH) under grants U24MH124629, U01MH124631, and U01MH124639, as well as the Wellcome Trust under grants 220664/Z/20/Z and 220664/A/20/Z.

## AUTHOR CONTRIBUTIONS

Z.R.B. and P.M.W. conceptualized and wrote the first draft of the manuscript. All other authors reviewed and edited the draft. P.M.W. and Z.R.B. conducted the analysis and developed the tables and figures. All authors have substantially contributed to the conception of the work described in the paper, or the data acquisition, quality control, and data curation. All authors have approved the submitted version and have agreed to be listed and accountable for their own contributions.

## COMPETING INTERESTS

The authors declare the following competing interests: K.A. is on the Australian Cognitive Impairment Associated with Schizophrenia Advisory Board for Boehringer Ingelheim and receives honorary funds. D.D. has received honorary funds for one educational seminar for CSL Sequiris. A.A. is a cofounder, serves as a member of the Board of Directors, as a scientific adviser, and holds equity in Manifest Technologies, Inc., and is a coinventor on the following patent: Anticevic A, Murray JD, and Ji JL:

Systems and Methods for NeuroBehavioral Relationships in Dimensional Geometric Embedding, PCT International Application No. PCT/US2119/022110, filed Mar 13, 2019. E.C. has received speaker fees at non-promotional educational events. P.F.-P. has received research funds or personal fees from Lundbeck, Angelini, Menarini, Sunovion, Boehringer Ingelheim, Proxymm Science, Otsuka, outside the current study. J.K. has received speaking or consulting fees from Janssen, Boehringer Ingelheim, ROVI, and Lundbeck. C.M.D.-C. has received grant support from Instituto de Salud Carlos III, Spanish Ministry of Science and Innovation, and honoraria or travel support from Angelini, Janssen, and Viatris; RU has received speaker fees at a non-promotional educational event: Otsuka: Consultancy for Viatris and Springer Healthcare. Honorary General Secretary, British Association for Psychopharmacology (unpaid). J.M.K. is a consultant to or receives honoraria and/or travel support and/or speakers bureau: Alkermes, Allergan, Boehringer-Ingelheim, Cerevel, Dainippon Sumitomo, H. Lundbeck, HealthRhythms, HLS Therapeutics, Indivior, Intracellular Therapies, Janssen Pharmaceutical, Johnson & Johnson, Karuna Therapeutics/Bristol Meyer-Squibb, LB Pharmaceuticals, Mapi, Maplight, Merck, Minerva, Neurocrine, Newron, Novartis, NW PharmaTech, Otsuka, Roche, Saladax, Sunovion, Teva; RSK provides consulting to Alkermes, Boehringer-Ingelheim. S.W.W. has received speaking fees from the American Psychiatric Association and from Medscape Features. He has been granted a US patent no. 8492418 B2 for a method of treating prodromal schizophrenia with glycine agonizts. He owns stock in NW PharmaTech. C.A. has been a consultant to or has received honoraria or grants from Acadia, Angelini, Biogen, Boehringer, Gedeon Richter, Janssen Cilag, Lundbeck, Medscape, Menarini, Minerva, Otsuka, Pfizer, Roche, Sage, Servier, Shire, Schering Plough, Sumitomo Dainippon Pharma, Sunovion, and Takeda; GDH has been a consultant for Bristol Myers Squibb. P.J.M. has been a consultant for Otsuka and TEVA; and Z.T. has been a consultant for Manifest Technologies. C.M.C. is an Associate Editor of Schizophrenia. All other authors report no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s41537-025-00669-z.

**Correspondence** and requests for materials should be addressed to Phillip M. Wolff.

**Reprints and permission information** is available at http://www.nature.com/reprints

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



[1]Department of Psychology, Emory University, Atlanta, GA, USA. [2]IBM TJ Watson Research Center, Yorktown Heights, NY, USA. [3]Institute for Technology in Psychiatry, McLean Hospital, Belmont, MA, USA. [4]Department of Psychiatry, Harvard Medical School, Boston, MA, USA. [5]Orygen Youth Health Research Centre, University of Melbourne, Parkville, VIC, Australia. [6]Centre for Youth Mental Health, The University of Melbourne, Parkville, VIC, Australia. [7]Department of Psychiatry, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. [8]Department of Psychiatry, Yale School of Medicine, New Haven, CT, USA. [9]Department of Psychiatry, Hotchkiss Brain Institute, University of Calgary, Calgary, AB, Canada. [10]General Psychiatry Service, Treatment and Early Intervention in Psychosis Program (TIPP–Lausanne), Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland. [11]Department of Psychosis Studies, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK. [12]Department of Child and Adolescent Psychiatry, Institute of Psychiatry and Mental Health, Hospital General Universitario Gregorio Marañón, Instituto de Investigación Sanitaria Gregorio Marañón (IiSGM), CIBERSAM, ISCIII, School of Medicine, Universidad Complutense, Madrid, Spain. [13]Department of Psychiatry and Behavioral Health, Ohio State University, Columbus, OH, USA. [14]Department of Psychology, Ohio State University, Columbus, OH, USA. [15]Institute for Mental Health, School of Psychology, University of Birmingham, Birmingham, UK. [16]Birmingham Women's and Children's NHS Foundation Trust, Birmingham, UK. [17]University of California, San Diego, CA, USA. [18]Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [19]Department of Psychiatry, School of Clinical Medicine, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong, Hong Kong. [20]Olin Neuropsychiatry Research Center, Hartford HealthCare Behavioral Health Network, Hartford, CT, USA. [21]Service de Psychiatrie Générale Dép. de Psychiatrie CHUV Lausanne, Lausanne, Switzerland. [22]Department of Psychiatry, Donald and Barbara Zucker School of Medicine, Hempstead, NY, USA. [23]Feinstein Institute for Medical Research, Manhasset, NY, USA. [24]Department of Psychology & Neuroscience, Temple University, Philadelphia, PA, USA. [25]Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy. [26]Department of Psychiatry, IMHAY, University of Chile, Santiago, Chile. [27]Behavioral Health Services, PeaceHealth Medical Group, Eugene, OR, USA. [28]Copenhagen Research Centre for Mental Health, Mental Health Copenhagen, Department of Psychology, University of Copenhagen, Copenhagen, Denmark. [29]Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA. [30]Department of Psychiatry, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany. [31]Department of Psychiatry, Chonnam National University Medical School, Gwangju, Korea. [32]Mindlink, Gwangju Bukgu Mental Health Center, Gwangju, Korea. [33]Department of Public Mental Health, Medical Faculty Mannheim, Central Institute of Mental Health, Heidelberg University, Heidelberg, Germany. [34]Department of Psychiatry, Hanyang University Hospital, Seoul, South Korea. [35]Department of Psychiatry, Hanyang University College of Medicine, Seoul, South Korea. [36]Department of Psychiatry and Psychotherapy, Jena University Hospital, Jena, Germany. [37]Department of Psychiatry, Washington University School of Medicine, St. Louis, MO, USA. [38]Department of Psychiatry and Behavioral Sciences and Weill Institute for Neurosciences, University of California, San Francisco, San Francisco, CA, USA. [39]Mental Health Service 116D, Veterans Affairs San Francisco Health Care System, San Francisco, CA, USA. [40]Department of Psychology, Northwestern University, Evanston, Il, USA. [41]Mental Health Services in the Capital Region, Copenhagen, Denmark. [42]Department of Clinical Medicine, Copenhagen University Hospital, Copenhagen, Denmark. [43]CAMEO, Early Intervention in Psychosis Service, Cambridgeshire and Peterborough NHS Foundation Trust, Cambridge, UK. [44]Department of Medicine, Institute of Biomedical Research (IBSAL), Universidad de Salamanca, Salamanca, Spain. [45]Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [46]Connecticut Mental Health Center, New Haven, CT, USA. [47]Prevention Science Institute, University of Oregon, Eugene, OR, USA. [48]Department of Psychological Science, University of California, Irvine, CA, USA. [49]PEPP-Montreal, Douglas Research Centre, Montreal, QC, Canada. [50]Department of Psychiatry, McGill University, Montreal, QC, Canada. [51]Department of Psychiatry, University of Rochester Medical Center, Rochester, NY, USA. [52]Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA. [53]Department of Psychology, University of Georgia, Athens, GA, USA. [54]Department of Psychiatry, University of Rochester Medical Center, Rochester, New York, USA. [55]Department of Psychiatric Rehabilitation and Counseling Professions, Rutgers University, Piscataway, NJ, USA. [56]Institute of Mental Health, Singapore, Singapore. [57]Shanghai Mental Health Center, Shanghai Jiaotong University School of Medicine, Shanghai, China. [58]Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [59]Department of Psychiatry, Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Hempstead, NY, USA. [60]Department of Psychiatry, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, Los Angeles, CA, USA. [61]Department Biobehavioral Sciences and Psychology, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, Los Angeles, CA, USA. [62]Department of Software Engineering and Information Technology, École de technologie supérieure, Montréal, QC, Canada. [63]These authors contributed equally: Cheryl M. Corcoran, Phillip M. Wolff. *A list of authors and their affiliations appears at the end of the paper. ✉email: pwolff@emory.edu

**ACCELERATING MEDICINES PARTNERSHIP® SCHIZOPHRENIA (AMP® SCZ)**

Zarina R. Bilgrami[1], Eduardo Castro[2], Carla Agurto[2], Einat Liebenthal[3,4], Michaela Ennis[4], Justin T. Baker[3,4], Isabelle Scott[5,6], Beau-Luke Colton[5,6], Kang Ik K. Cho[7], Linying Li[1], Zailyn Tamayo[8], Mara Henecks[8], Habiballah Rahimi Eichi[3,4], Tae'lar Henry[1], Jean Addington[9], Luis K. Alameda[10,11], Celso Arango[12], Nicholas J. K. Breitborde[13,14], Matthew R. Broome[15,16], Kristin S. Cadenhead[17], Monica E. Calkins[18], Eric Yu Hai Chen[19], Jimmy Choi[20], Philippe Conus[10,21], Barbara A. Cornblatt[22,23], Lauren M. Ellman[24], Paolo Fusar-Poli[11,25], Pablo A. Gaspar[26], Carla Gerber[27], Louise Birkedal Glenthøj[28], Leslie E. Horton[29], Christy Hui[19], Joseph Kambeitz[30], Lana Kambeitz-Ilankovic[30], Matcheri S. Keshavan[3], Sung-Wan Kim[31,32], Nikolaos Koutsouleris[11,33], Jun Soo Kwon[34,35], Kerstin Langbein[36], Daniel Mamah[37], Covadonga M. Diaz-Caneja[12], Daniel H. Mathalon[38,39], Vijay A. Mittal[40], Merete Nordentoft[41,42], Godfrey D. Pearlson[1,8], Jesus Perez[43,44], Diana O. Perkins[45], Albert R. Powers III[8,46], Jack Rogers[15], Fred W. Sabb[47], Jason Schiffman[48], Jai L. Shah[49,50], Steven M. Silverstein[51], Stefan Smesny[36], William S. Stone[4,52], Walid Yassin[4,52], Gregory P. Strauss[53], Judy L. Thompson[54,55], Rachel Upthegrove[15], Swapna Verma[56], Jijun Wang[57], Daniel H. Wolf[18], Patrick D. McGorry[5,6], Rene S. Kahn[58], John M. Kane[23,59], Alan Anticevic[8,46], Carrie E. Bearden[60,61], Dominic Dwyer[5,6], Tashrif Billah[7], Sylvain Bouix[7,62], Ofer Pasternak[7], Martha E. Shenton[4,7], Scott W. Woods[8], Barnaby Nelson[5,6], Guillermo A. Cecchi[2], Cheryl M. Corcoran[58,63] and Phillip M. Wolff[1,63] ✉

A full list of members and their affiliations appears in the Supplementary Information.