Characterizing speech heterogeneity in schizophrenia-spectrum disorders

Oomen, Priscilla P; de Boer, Janna N; Brederoo, Sanne G; Voppel, Alban E; Brand, Bodyl A; Wijnen, Frank N K; Sommer, Iris E C

Published in:
Journal of psychopathology and clinical science

DOI:
10.1037/abn0000736

Publication date:
2022

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Psychopathology and Clinical Science

ISSN: 2769-7541

2022, Vol. 131, No. 2, 172–181
https://doi.org/10.1037/abn0000736

Characterizing Speech Heterogeneity in Schizophrenia-Spectrum Disorders

Priscilla P. Oomen[1, 2], Janna N. de Boer[1, 3], Sanne G. Brederoo[1, 2], Alban E. Voppel[1, 2], Bodyl A. Brand[1, 2], Frank N. K. Wijnen[4], and Iris E. C. Sommer[1, 2]

[1] Department of Biomedical Sciences of Cells and Systems, University Medical Center Groningen
[2] Department of Psychiatry, University Medical Center Groningen
[3] Department of Psychiatry, University Medical Center Utrecht
[4] Utrecht Institute of Linguistics OTS, Utrecht University

Schizophrenia-spectrum disorders (SSD) are highly heterogeneous in risk factors, symptom characteristics, and disease course outcome. Although speech anomalies have long been recognized as a core symptom of SSD, speech markers are an unexplored source of symptom heterogeneity that may be informative in recognizing relevant subtypes. This study investigated speech heterogeneity and its relation to clinical characteristics in a large sample of patients with SSD and healthy controls. Speech samples were obtained from 142 patients with SSD and 147 healthy controls by means of open-ended interviews. Speech was analyzed using standardized open-source acoustic speech software. Hierarchical clustering was conducted using acoustic speech markers. Symptom severity was rated with the Positive and Negative Syndrome Scale, and cognition was assessed with the Brief Assessment of Cognition for Schizophrenia. Three speech clusters could be distinguished in the patient group that differed regarding speech properties, independent of medication use. One cluster was characterized by mild speech disturbances, while two severely impaired clusters were recognized (fragmented speakers and prolonged pausers). Both clusters with severely impaired speech had more severe cognitive dysfunction than the mildly impaired speakers. Prolonged pausers specifically had difficulties with memory-related tasks. Prolonged pausing, as opposed to fragmented speaking, related to chronic active psychosis and refractory psychotic symptoms. Based on speech clustering, subtypes of patients emerged with distinct disease trajectories, symptomatology, and cognitive functioning. The identification of clinically relevant subgroups within SSD may help to characterize distinct profiles and benefit the tailoring of early intervention and improvement of long-term functional outcome.

General Scientific Summary

Speech anomalies have long been recognized as a core symptom of schizophrenia-spectrum disorders (SSD), yet speech markers are an unexplored source of symptom heterogeneity that may be informative in recognizing relevant subtypes of SSD. This study showed the existence of distinct speech subtypes with divergent disease trajectories, symptomatology, and cognitive functioning. This supports the notion that speech can provide valuable information about the patient and benefits the tailoring of early intervention and improvement of long-term functional outcome.

Keywords: language, speech, clustering, psychosis, schizophrenia

Supplemental materials: https://doi.org/10.1037/abn0000736.supp

Priscilla P. Oomen https://orcid.org/0000-0001-8430-0905
Janna N. de Boer https://orcid.org/0000-0003-1231-2733
Alban E. Voppel https://orcid.org/0000-0002-6768-668X
Frank N. K. Wijnen https://orcid.org/0000-0002-7196-6000

Iris E. C. Sommer was supported by a TOP Grant from The Netherlands Organization for Health Research and Development (ZonMW, Project 91213009). We are grateful to all research interns for their help with data collection and preparation. Iris E. C. Sommer is a consultant to Gabather and received research support from Janssen Pharmaceuticals Inc. and Sunovion Pharmaceuticals Inc.

Priscilla P. Oomen and Janna N. de Boer contributed equally to this work. Priscilla P. Oomen contributed to methodology, formal analysis, investigation, and writing—original draft. Janna N. de Boer contributed to conceptualization, methodology, investigation, and writing—original draft. Sanne G. Brederoo contributed to writing—review and editing. Alban E. Voppel contributed to writing—review and editing and investigation. Bodyl A. Brand contributed to writing—review and editing and investigation. Frank N. K. Wijnen contributed to writing—review and editing and supervision. Iris E. C. Sommer contributed to writing—review and editing, supervision, and funding acquisition.

Correspondence concerning this article should be addressed to Priscilla P. Oomen, Department of Biomedical Sciences of Cells and Systems, University Medical Center Groningen, Antonius Deusinglaan 2, 9713 AW Groningen, the Netherlands. Email: p.p.oomen@umcg.nl

Schizophrenia-spectrum disorders (SSD) have been characterized as highly heterogeneous. Heterogeneity is recognized in risk factors (e.g., drug use, comorbidities, early life brain trauma, gender; Owen et al., 2016; Voineskos et al., 2020), in symptom characteristics (e.g., hallucinations, delusions, speech abnormalities, motivation, cognitive dysfunction; Ahmed et al., 2018; Carruthers et al., 2019; Dickinson et al., 2018), as well as in disease course outcome (e.g., early/late onset, remission patterns, stability, refractory psychosis; Salagre et al., 2020; Suvisaari et al., 2018). Given that only two of the five diagnostic criteria for schizophrenia are required to make a diagnosis, there is room for substantial variability among individuals within the group of patients who receive the diagnosis schizophrenia (American Psychiatric Association, 2013). When individuals with schizophreniform disorder and schizoaffective disorder are also included, as is often the case, this variability increases even more. Acknowledging this high heterogeneity in symptoms has clinical importance since prognosis and required treatments and care may differ greatly among patients who share little to no symptoms (Cohen et al., 2014; Insel, 2014; Schnack, 2019). A largely unexplored source of symptom heterogeneity in SSD is speech.

Speech is a subset of the larger field of human language processing. The term speech is used for the oral output of language. Speech anomalies have long been recognized as a core symptom of SSD (Bleuler, 1911; Kraepelin et al., 1919), and clinicians often report these atypical speech patterns in their mental state examination, including descriptions of poverty of speech, slow or hesitant speech, and distinctive tone (Alpert et al., 2002). A recent meta-analysis of speech disturbances suggests that pitch variability, proportion of spoken time, speech rate, and pauses are abnormal in SSD (Parola et al., 2020). Recent developments in information technology and computational linguistics allow for the application of highly specialized language tools to spoken language, which makes speech analysis quick, objective, and reliable (Corcoran & Cecchi, 2020; de Boer et al., 2021).

Spoken language analysis is an important candidate for identifying heterogeneity for two main reasons. First, speech analysis fulfills the criteria for an ideal biomarker (Califf, 2018; Holland, 2016; Verma et al., 2011) because it is reliable (Eyben et al., 2016), consistent within individuals (Hasan et al., 2004; Ingram et al., 2013; Nolan & Grigoras, 2005), ecologically valid (Maryn et al., 2010; Sbordone, 1996; Schmuckler, 2001), easily measured, and inexpensive (de Boer et al., 2021). Moreover, the fact that speech anomalies are a characteristic symptom of SSD (American Psychiatric Association, 2013) points to the relevance of speech as a potential biomarker for these disorders. Second, speech disturbances are closely related to important predictors for clinical endpoints. Abnormalities in pausing have been associated with positive and negative symptoms, in both individuals at clinical high risk (Agurto et al., 2020; Sichlinger et al., 2019; Stanislawski et al., 2021) and patients with schizophrenia (Cohen et al., 2016). In addition, speech disturbances are related to cognitive function (Barker et al., 2020; R. W. Brown & Lenneberg, 1954; Carroll, 1964; Dunn, 2017). Features such as pauses, speech rate, and pitch variability are potential indicators of cognitive load both in healthy controls and individuals with SSD (Cohen et al., 2012, 2015; Khawaja et al., 2008), and speech abnormalities have been related to impairments in attention (Docherty et al., 2006). Furthermore, speech abnormalities are predictive of functional outcome (Bowie & Harvey, 2008; Dickinson et al., 2007), are related to social relations (Oliveira et al., 2015), and have a negative impact on quality of life in SSD (Tan et al., 2014). To date, little is known about the distribution of speech disturbances across individuals with an SSD. Traditionally dichotomizing a sample based on speech disturbances (e.g., "disturbed speech" vs. "normal speech") does not do justice to the different ways in which speech can be disturbed. To overcome this, we use a data-driven hierarchical clustering method to identify different patterns of speech anomalies in SSD patients. In such an approach, the speech data themselves are informative in recognizing clusters or subgroups of patients. We further assessed the association between speech clusters, cognition, and symptomatology in SSD patients to evaluate the quality of speech in identifying relevant and useful disease heterogeneity.

Method and Materials

Participants

The data of 142 patients with an SSD and 147 healthy controls were drawn from independent research studies examining cognition in SSD at the University Medical Center Utrecht (UMCU), the Netherlands. Approval from the UMCU Ethics Review Committee was obtained, and each trial is registered in the European Clinical Trials Database (EudraCT 2013-000834-36, 2015–004483-11, 2017–002406-12). Written informed consent was obtained prior to study participation. Psychiatric diagnoses were confirmed by the Structured Clinical Interview for *DSM–IV* (First, 2014), the Comprehensive Assessment of Symptoms and History (Andreasen et al., 1992), or the Mini-International Interview (Sheehan et al., 1998) depending on the study the participants originally enrolled in. Only Dutch native speakers were included in the present study. Exclusion criteria were the presence of uncorrected hearing difficulties or speaking dysfunction such as stuttering.

Procedure

Spoken Language

Open-ended, semistructured interviews were obtained from participants. To promote spontaneous speech by the participants, interviewers were instructed to refrain from speaking as much as possible without creating an unnatural interview setting. Interviews were recorded using headset cardioid microphones onto a TASCAM-DR40 steady state recorder, using two channels with 16.000-Hz sampling. Speech was elicited using a standard list of questions. All questions were deliberately neutral; topics that would have a different emotional valence for patients and healthy controls (such as health) were avoided. A question was skipped if the subject did not feel comfortable answering it. For a list of questions, we refer to Supplemental Table 3. Interviews lasted approximately 13 min for all participants. For more elaborate descriptions of the methodology, see previous reports by our group (de Boer, van Hoogdalem, et al., 2020; de Boer, Voppel, et al., 2020; de Boer et al., 2021; Voppel et al., 2021). Crosstalk (i.e., speech from the interviewer on the participants audio channel) was removed as follows: (a) silences were annotated on the interviewer's audio channel in PRAAT (Boersma & Weenink, 2013; function: annotate to text grid silences; settings: minimum pitch

100 Hz, time step .0, silence threshold −30.0 dB, minimum silence duration 1.0 s, minimum sounding duration 0.1 s), (b) the resulting regions (i.e., regions in which the interviewer was silent) were selected on the participants channel, and (c) these voiced (speech) regions were concatenated into a new audio file containing only the participant's speech.

Preprocessing of Speech Data

Based on a recent large systematic review and meta-analysis about acoustic speech patterns in patients with SSD and healthy controls (Parola et al., 2020), the following aspects of speech were assessed: the length and number of voiced (speech) regions, pauses, pitch variability, and proportion of spoken time. The Praat Script Syllable Nuclei v2 (Quené et al., 2011) was used to calculate proportion of spoken time (see Table 1). The GeMAPS parameter set was extracted using OpenSMILE (Eyben et al., 2013) to obtain the other three aspects of speech. GeMAPS provides arithmetic means and coefficients of variation (standard deviation normalized by the arithmetic mean) for each parameter. See Table 1 for an overview of the used variables. A high number of voiced regions indicates more fragmented speech, highly interrupted with pauses. For each aspect of speech, a z score was calculated relative to the healthy control participants.

Of note, the eGeMAPS parameter imposes no minimal length on voiced or unvoiced regions (Eyben et al., 2016), which means that all unvoiced frames are taken into the calculation of "unvoiced region length," even if they are only one frame long. Short silences in speech (< 200 ms) are often related to the articulation of particular sounds, notably plosives (e.g., the /p/, which introduces a short silence in the sound wave; Rosen, 1992). Therefore, we performed a second analysis in PRAAT to test the reliability of the eGeMAPS approximation of pauses. Average pause duration was calculated using the Praat Script Syllable Nuclei v2 (Quené et al., 2011) developed for Dutch. In this script, we defined pauses as silences longer than 200 ms, thereby excluding silences introduced by plosives. The resulting average pause duration from PRAAT was strongly correlated with the unvoiced region length parameter from eGeMAPS in all participants (n = 289, r = .809, p < .0001), suggesting that unvoiced region length is a reliable approximation of pause duration

Cognitive Functioning

Cognition was assessed in all patients and a subset of the healthy controls (n = 31) using the Brief Assessment of Cognition in Schizophrenia (BACS; Keefe et al., 2004), which consists of the following tasks:

1. List learning – Verbal memory
2. Digit sequencing – Working memory
3. Token motor task – Motor speed
4. Category Instances and Controlled Oral Word Association Test – Verbal fluency
5. Symbol coding – Attention and information processing speed
6. Tower of London – Executive function

Individual BACS scores were converted into standardized z scores that are corrected for age and gender based on previously published norm scores (Keefe et al., 2008). For demographic characteristics of the healthy controls with BACS scores, see Table S1.

Statistical Analysis

Statistical analyses were performed in SPSS Statistics Version 25.0. Subject characteristics were compared between healthy controls and patients with SSD using an analysis of variance (ANOVA) for continuous values and a chi-square test for categorical values. Next, for the sample of patients, z scores of each speech parameter as summarized in Table 1 were entered into the clustering analysis. Clustering analysis was conducted using an agglomerative hierarchical clustering approach. Case similarity was computed with squared Euclidean distance and Ward's (1963) linkage as agglomeration procedure specification. Collaborative examination of the dendrogram and the agglomeration schedule coefficients (see Figures S1 and S2) were used to establish the optimal number of clusters, following Carruthers et al. (2019). Emergent patient clusters and healthy controls were compared on demographic variables, cognitive function, and clinical variables using an ANOVA with Bonferroni post hoc correction for continuous values and a chi-square test or Fisher's exact test for categorical values. Speech features between the emergent clusters were compared using analysis of covariance (ANCOVA) corrected for age with

Table 1
Speech Parameters

Parameter	Definition	Description	Parameter type
Mean number of voiced regions per second	Number of speech regions	Average number of continuous voiced regions (F > 0) per second (more regions indicates more fragmented speech and thus speech interrupted with pauses)	Temporal
Mean voiced region length per second	Speech turn duration	The mean length of continuously voiced regions (F0 > 0)	Temporal
Mean unvoiced region length per second	Pause length	The mean length of unvoiced regions (F0 = 0)	Temporal
F0 semitone (*SD*)	Pitch variability	Pitch is the logarithmic F0 on a semitone frequency scale, starting at 27.5 Hz (semitone 0), coefficient of variation	Frequency related
Proportion of time articulating	Proportion of spoken time	Phonation time participant/full interview duration. Note: the full interview duration includes the speech of the interviewer.	Amount

Bonferroni post hoc correction. For all analyses, the alpha level was set at .05.

Results

Descriptive statistics and cognitive domain comparisons between healthy controls and the total sample of patients can be found in Table S2. Patients with SSD have significantly more interrupted speech, longer pause duration, lower pitch variability, and a lower proportion of time articulating compared to healthy controls (Table S2). Cluster analysis and inspection of the dendrogram and agglomeration schedule coefficients (see Figures S1 and S2) resulted in a three-cluster solution within the group of patients with SSD (see Table 2).

Speech features were normally distributed in both patients and healthy controls. Based on the five included aspects of speech, three clusters were observed that can be characterized as one mildly impaired cluster (the "mildly impaired speakers") and two severely impaired speakers (the "fragmented speakers" and the "prolonged pausers"); see Table 3 and Figure 1. In the following sections, we describe the speech characteristics and demographic information of these three groups in more detail. All the patient groups spoke a smaller percentage of time than the healthy controls, though no difference emerged between the speech groups (all ps $> .05$). This speech characteristic will therefore not be discussed in further detail. Significant age differences were shown. Fragmented speakers were significantly younger than healthy controls ($p < .001$) and mildly impaired speakers ($p = .001$) and prolonged pausers ($p < .001$). Therefore, additional group comparisons (ANCOVA) were performed corrected for age; see Table 2. Moreover, chlorpromazine equivalents did not significantly differ between speech groups ($p = .061$).

Mildly Impaired Speakers

Compared to healthy speakers, the mildly impaired speakers ($n = 58$) have more interrupted speech, increased pause duration, and decreased pitch variability (all ps $< .001$; see Table 3). The mildly impaired speakers can be considered mildly impaired since their average deviation from the controls is smaller than the deviation in the severely impaired groups ($p < .001$). Compared to both the fragmented speakers and prolonged pausers, the mildly impaired speakers have normal pause duration and pitch variability and have less fragmented speech ($p < .001$). They have an intermediate illness duration, and their PANSS score is on average 11 points lower than that of the prolonged pausers ($p < .001$). Mildly impaired speakers show less overall cognitive impairment than the other patients with SSD (see Figure 2; all ps $< .012$). Compared to healthy controls, significant cognitive impairment is present in list learning ($p = .011$) and symbol coding ($p < .001$; see Table 2).

Fragmented Speakers

Compared to both the healthy controls and the mildly impaired speakers, speech of fragmented speakers ($n = 64$) can be considered more severely impaired (see Table 3) and is most characterized by frequent use of short voiced regions (high number of voiced regions, z score 7.22, both ps $< .001$) indicating fragmented speech. They also spoke shorter periods of time compared to both the healthy controls and the other speech groups (low voiced region length, z score $-.69$, all ps $< .001$). Their PANSS scores are similar to those of the mildly impaired speakers and approximately 10 points lower than that of the prolonged pausers ($p < .001$). Their overall cognition was impaired compared to the mildly impaired speakers and the controls (both ps $< .001$). Cognitive impairment relative to healthy controls was evident in all

Table 2
Mean (SD) Scores for Patients With SSD and Healthy Controls

Sample characteristics	Healthy controls ($n = 147$)	Mildly impaired speakers ($n = 58$)	Fragmented speakers ($n = 64$)	Prolonged pausers ($n = 20$)	Statistic F, χ^2	df	p	Post hoc analyses	Post hoc analyses, age corrected
Gender, n (male/female)	86/61	46/12	44/20	16/4	$\chi^2 = 10.26$	3	$p = .016$	[a]	N/A
Age, M (SD)	34.95 (14.23)	34.97 (14.03)	26.09 (8.44)	40.10 (8.63)	$F = 9.79$	3	$p < .001$	[b, d, f]	N/A
Years of education, M (SD)	14.87 (1.92)	13.35 (2.28)	13.00 (2.46)	13.20 (2.04)	$F = 15.33$	3	$p < .001$	[a, b, c]	[a, b, c]
Years of education parents, M (SD)	12.77 (2.91)	13.04 (2.73)	12.46 (3.00)	11.09 (3.03)	$F = 2.01$	3	$p = .113$	—	[b, d]
Duration disease (years), M (SD)	N/A	7.00 (11.95)	2.19 (4.24)	15.14 (9.27)	$F = 16.96$	2	$p < .001$	[d, e, f]	[e, f]
Chlorpromazine equivalent, M (SD)	N/A	249.60 (229.48)	260.63 (220.60)	387.78 (269.29)	$F = 2.85$	2	$p = .061$	—	—
PANSS, M (SD)									
Total	N/A	45.14 (8.55)	46.55 (10.20)	56.80 (14.19)	$F = 10.09$	2	$p < .001$	[e, f]	[e, f]
Positive	N/A	10.33 (3.95)	9.69 (2.87)	13.60 (5.04)	$F = 8.69$	2	$p < .001$	[e, f]	[e, f]
Negative	N/A	11.10 (3.56)	12.95 (4.76)	13.95 (4.90)	$F = 4.39$	2	$p = .014$	[e]	[e]
General	N/A	23.71 (4.51)	23.91 (5.00)	29.25 (8.18)	$F = 8.84$	2	$p < .001$	[e, f]	[e, f]
BACS, M (SD) z score[g]									
Composite score	0.16 (1.32)	−0.82 (1.01)	−1.62 (1.13)	−1.76 (1.13)	$F = 20.41$	3	$p < .001$	[a, b, c, d, e]	[a, b, c, d, e]
Verbal memory	0.39 (1.17)	−0.40 (1.11)	−0.86 (1.07)	−1.23 (1.22)	$F = 11.65$	3	$p < .001$	[a, b, c, e]	[b, c]
Working memory	0.10 (1.09)	−0.56 (1.19)	−1.02 (1.12)	−1.81 (1.45)	$F = 12.21$	3	$p < .001$	[b, c, e]	[b, c, e]
Motor speed	−0.17 (1.12)	−0.61 (1.23)	−0.99 (1.18)	−1.44 (1.37)	$F = 5.61$	3	$p = .001$	[b, c]	[b, c]
Verbal fluency	0.13 (1.17)	−0.55 (1.13)	−1.01 (1.20)	−1.08 (1.18)	$F = 7.68$	3	$p < .001$	[b, c]	[a, b, c]
Attention and processing speed	−0.06 (1.28)	−0.96 (0.84)	−1.38 (0.93)	−1.42 (0.98)	$F = 13.76$	3	$p < .001$	[a, b, c]	[a, b, c]
Executive function	0.18 (0.87)	0.10 (0.95)	−0.38 (1.59)	−0.24 (1.37)	$F = 2.10$	3	$p = .102$	—	—

Note. Patients with SSD $n = 142$. PANSS = Positive and Negative Syndrome Scale; BACS = Brief Assessment of Cognition for Schizophrenia.
[a] HC significantly different from mildly impaired speakers. [b] HC significantly different from fragmented speakers. [c] HC significantly different from prolonged pausers. [d] Mildly impaired speakers significantly different from fragmented speakers. [e] Mildly impaired speakers significantly different from prolonged pausers. [f] Fragmented speakers significantly different from prolonged pausers. [g] BACS healthy controls $n = 31$.

Table 3
Mean (SD) Scores, Age Corrected, for Patients With SSD and Healthy Controls

Speech measures, *M* (*SD*) *z* score	Healthy controls (*n* = 147)		Mildly impaired speakers (*n* = 58)		Fragmented speakers (*n* = 64)		Prolonged pausers (*n* = 20)		Statistic	*df*	*p*	Post hoc analyses
	Z score	Raw score	*Z* score	Raw score	*Z* score	Raw score	*Z* score	Raw score				
Number of voiced regions (seconds)	0.00 (1.00)	1.25 (0.14)	1.77 (1.53)	1.50 (0.22)	7.22 (2.07)	2.27 (0.29)	−2.36 (1.75)	0.92 (0.25)	$F = 374.06$	3	$p < .001$	a, b, c, d, e, f
Voiced region length	0.00 (1.00)	0.25 (0.05)	0.33 (1.06)	0.27 (0.05)	−0.69 (0.76)	0.22 (0.04)	1.11 (1.79)	0.30 (0.09)	$F = 13.34$	3	$p < .001$	b, c, d, f
Pause length	0.00 (1.00)	0.20 (0.06)	1.44 (1.74)	0.28 (0.10)	0.39 (0.95)	0.22 (0.05)	3.00 (2.65)	0.37 (0.15)	$F = 39.07$	3	$p < .001$	a, c, d, e, f
Pitch variability	0.00 (1.00)	0.21 (0.04)	−0.64 (0.82)	0.18 (0.04)	0.15 (0.96)	0.22 (0.04)	−0.79 (0.96)	0.17 (0.04)	$F = 13.92$	3	$p < .001$	a, c, d, f
Proportion of spoken time	0.00 (1.00)	57.06 (8.65)	−1.44 (1.10)	44.70 (9.49)	−1.41 (1.11)	44.94 (9.55)	−1.65 (1.38)	42.81 (11.95)	$F = 42.83$	3	$p < .001$	a, b, c
Average deviation from healthy controls	N/A		1.12 (1.25)		1.97 (2.07)		1.78 (1.71)		$F = 34.58$	2	$p < .001$	d, e

Note. Patients with SSD $n = 142$.
[a] HC significantly different from mildly impaired speakers. [b] HC significantly different from fragmented speakers. [c] HC significantly different from prolonged pausers. [d] Mildly impaired speakers significantly different from fragmented speakers. [e] Mildly impaired speakers significantly different from prolonged pausers. [f] Fragmented speakers significantly different from prolonged pausers.

subdomains except for executive function (Tower of London; all *p*s $< .015$; see Table 2).

Prolonged Pausers

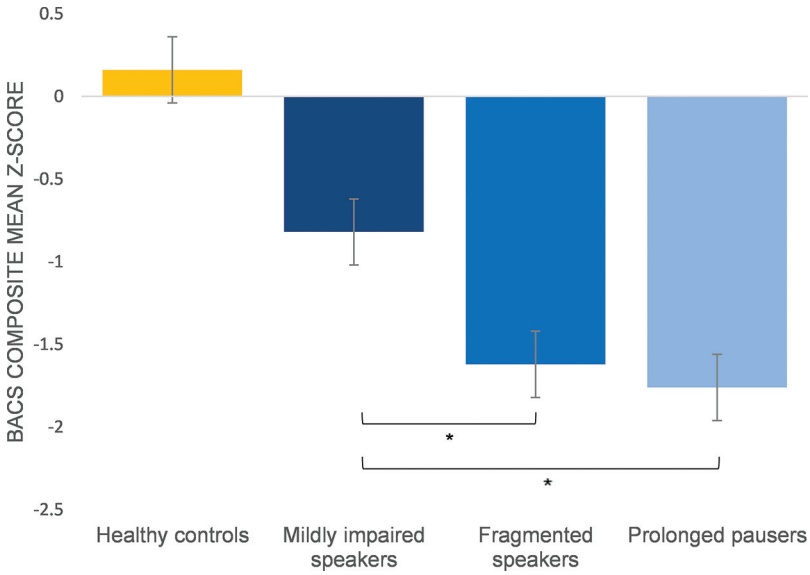
Compared to both the healthy controls and the other speech groups, speech of the prolonged pausers ($n = 20$) is more severely impaired. The most prominent characteristic is the length of their pauses (z score 3.00, $p < .001$) and their pause frequency indicated by a low number of voiced regions (z score −2.36, $p < .001$). Moreover, they have a higher pitch variability ($p = .001$) and decreased length of voiced regions compared to healthy speakers ($p < .001$). This group had the highest total, positive, and general PANSS scores compared to both the mildly impaired speakers and the fragmented speakers. Negative PANSS scores were higher compared to the mildly impaired speakers. Additionally, only 35% of the prolonged pausers were in remission, compared to 57% in the mildly impaired speakers and 58% in the fragmented speakers. Their overall cognition was impaired compared to the mildly impaired speakers and the healthy controls (both *p*s $< .001$). Compared to the mildly impaired speakers, prolonged pausers showed most impairments in verbal memory (list learning; $p = .055$) and working memory (digit sequencing; $p < .001$). Prolonged pausers showed impairments in all subdomains except for executive function (Tower of London) compared to the healthy controls ($p < .002$).

Figure 1
Variability in Speech Features Across the Three Speech Clusters and Healthy Controls

Note. Pentagons represent z scores for each speech feature, relative to the healthy control participants. See the online article for the color version of this figure.

Figure 2

Global Cognitive Functioning as Measured by the Composite BACS Score Across Speech Clusters and Healthy Controls



Note. BACS = Brief Assessment of Cognition in Schizophrenia. The healthy control group differed from all clusters on cognitive functioning. * $p < .05$. See the online article for the color version of this figure.

Discussion

This study examined speech heterogeneity in a large sample of schizophrenia-spectrum patients. A data-driven hierarchical clustering approach indicated the presence of three diverse speech subgroups in SSD. The three emergent clusters showed significant variability and differences across general disease characteristics, symptom profiles, and cognitive functioning. Based on speech heterogeneity, one mildly impaired speech group and two severely impaired speech groups could be distinguished. The severely impaired speech groups differed from each other in speech characteristics as well as symptom profiles, with one group being characterized by fragmented speech and the other group being characterized mostly by prolonged pauses. Our results show that individuals belonging to either of these severely impaired speech clusters have impaired cognition compared to those in the mildly impaired speech group. Moreover, the prolonged pausers have a longer illness duration and higher PANSS scores and are less likely to be in symptom remission indicating refractory symptoms, when compared to the other patients with SSD. These findings show that automatic analysis of just a few minutes of recording of natural speech can provide valuable information about the patient with regard to cognition and symptom remission. Given the ease and acceptability of the speech recording, it could find its place in clinical practice as a momentary biomarker to add objective and reliable information on mental status of the patient.

Our results are in line with a recent systematic review and meta-analysis on acoustic patterns in schizophrenia, demonstrating atypicalities in pitch variability, proportion of time spoken, and pauses compared to healthy controls (Parola et al., 2020). Furthermore, our findings are in line with previous research indicating an association between speech disturbances and cognition in SSD (Barch & Ceaser, 2012; Becker et al., 2012; M. Brown & Kuperberg, 2015; Cavelti et al., 2018; Hinzen & Rosselló, 2015; Kerns & Berenbaum, 2002; Liddle et al., 2002; Lundin et al., 2020; Sumner et al., 2018). Our results further indicate that increased pause time is specifically associated with memory since the group of prolonged pausers showed significantly more impairment on both verbal memory and working memory tasks compared to the mildly impaired speakers.

There are several ways in which poor cognition relates to speech properties. First, spontaneous speech can be described as the process of converting thoughts into temporal sequences of speech units. Pauses are an inherent feature of the normal speech process and can be interpreted as reflecting feedback loops in which a person processes what they just said, while the next unit of information is planned (Levelt, 1983; Lundholm Fors, 2015). In the case of cognitive dysfunction, such planning operations are not flawless, which affects the type of speech unit that is generated next. Pauses often occur before words that are rarely used because such words have a longer lexical retrieval time than frequently used words (Alario et al., 2002). Pauses at clausal boundaries are related to the complexity of the subsequent clausal structure as syntactically complex sentences require more planning time (Ferreira, 1991). In fact, pauses at clausal boundaries have been associated with the activation of the left temporal gyrus, which may be related to lexical retrieval (Kircher et al., 2004; Matsumoto et al., 2013). Thought-disordered patients with SSD have more sentence-initial pauses and more pauses before embedded sentences, indicating increased processing time of complex syntactic units (Çokal et al., 2019). This is substantiated by functional MRI studies showing a differential pattern of brain activity during pauses in patients with SSD, possibly reflecting impairments of lexical retrieval on a neurobiological level (Matsumoto et al., 2013). Previous research by our group performed in patients with SSD showed

an association between language production and white matter integrity in the language tracts (de Boer, van Hoogdalem, et al., 2020). Specifically, pause duration was a strong predictor for the integrity of the language tracts in patients with SSD (de Boer, van Hoogdalem, et al., 2020). As pause duration is thought to reflect speaking efficiency and/or processing speed (Deary et al., 2006; de Boer, van Hoogdalem, et al., 2020), increased white matter integrity appears to be associated with higher speech processing efficiency. Following from this, one could hypothesize that the prolonged pausers in the current study have reduced white matter tract integrity in comparison to the other speech clusters.

Second, speakers with poor attention might be easily distracted by competing thoughts or associations and forget their original discourse plan before completion. Indeed, sustained attention has been associated with communication failures in SSD (Docherty et al., 2006). Our results indicate that both severely impaired speech groups have more cognitive impairments than the mildly impaired speakers and controls. Given that attention is required for all cognitive tests, impaired attention might be a process underlying the more severely impaired speech in SSD. Remarkably, executive function assessed by the Tower of London was relatively spared in all three subgroups, although impairments of executive functioning have been frequently reported in SSD (Reichenberg & Harvey, 2007). Since the term executive functioning covers a large set of cognitive capacities, the Tower of London task may possibly be limited in the assessment of executive function as it primarily assesses planning but not, for example, inhibition, switching, and flexibility. Indeed, literature suggests that different executive functions show differential patterns of impairment in patients with schizophrenia (Thuaire et al., 2020). Significant differences in age were demonstrated between groups. Age is known to influence speech. For instance, more frequent pauses have been demonstrated in older speakers compared to younger speakers (Bóna, 2014). In addition, older age is associated with cognitive decline (Hedden & Gabrieli, 2004). However, additional analyses corrected for age showed similar patterns of cognitive performance, speech, and symptomatology across groups, indicating that group effects were not driven by differences in age. Interestingly, the amount of speech (proportion of spoken time) seems to be a generalized impairment in SSD as it does not differentiate between the three speech clusters.

The current study has some limitations. First, this study has a cross-sectional design. Although different patterns of disease progression emerged from our data, we do not know whether these speech subtypes are stable throughout the course of the disease. Second, possible influences of antipsychotic medication on speech and the formation of subgroups have not been assessed, while there is indication that antipsychotics affect speech (de Boer, Voppel, et al., 2020). Although chlorpromazine equivalents were not significantly different across speech clusters, the effects of cumulative dose could not be evaluated. Third, men were overrepresented in our sample. However, since percentages were about equal across clusters, this most likely did not influence the formation or characterization of clusters. Fourth, a limitation of cluster analyses in general is that they are influenced by the selected cluster algorithm and the criteria used to determine the number of clusters. However, we followed the recommended guidelines for reporting on cluster analysis by Carruthers et al. (2019). Of note, the current study focuses only on speech disturbances, which encompasses research focused on the acoustic properties of spoken language. Speech is a subset of the larger field of human language processing, which also includes the study of meaning, grammar, pragmatics, and language perception and acquisition. Further research should examine whether these clusters extend to other fields of language processing as well. A strength of the current study is the use of a semistructured interview with neutral prompts, many of which were memory related (e.g., "Tell us about your most recent birthday celebrations"). This line of questioning likely induced more pausing related to memory retrieval.

Over the past years, the use of natural language processing tools for the analysis and classification of psychosis has rapidly expanded. Acoustic measures of spoken language can be easily recorded and quantified through open-source software and are more objectively obtained compared to conventional speech methods that are assessed by clinicians. Recent literature has shown the value of speech analysis as a biomarker for psychosis (Corcoran et al., 2020; Corcoran & Cecchi, 2020; de Boer, Brederoo, et al., 2020; Hitczenko et al., 2021). Speech features have been proven to predict psychosis development with very high accuracy (Gutiérrez et al., 2017; Pietrowicz et al., 2019). Such biomarkers are of great clinical relevance in psychiatry. The current study adds to this line of research by acknowledging the heterogeneity of speech anomalies in SSD and showing that speech clusters have distinct clinical characteristics.

In conclusion, a cluster analysis was performed to investigate speech heterogeneity in a sample of patients with SSD. Three clusters emerged, with significant differences across disease characteristics, symptom profiles, and cognitive function. Defining the existence of subgroups within SSD may be useful in the characterization of heterogeneity. Further longitudinal studies are required to assess the possible predictive value of speech subgroups on the clinical and cognitive course of SSD. The identification of such clinically relevant subgroups within SSD may help to characterize distinct profiles and benefit the tailoring of early intervention and improvement of long-term functional outcome.

References

Agurto, C., Pietrowicz, M., Norel, R., Eyigoz, E. K., Stanislawski, E., Cecchi, G., & Corcoran, C. (2020). Analyzing acoustic and prosodic fluctuations in free speech to predict psychosis onset in high-risk youths. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 5575–5579). IEEE. https://doi.org/10.1109/EMBC44109.2020.9176841

Ahmed, A. O., Strauss, G. P., Buchanan, R. W., Kirkpatrick, B., & Carpenter, W. T. (2018). Schizophrenia heterogeneity revisited: Clinical, cognitive, and psychosocial correlates of statistically-derived negative symptoms subgroups. *Journal of Psychiatric Research*, *97*, 8–15. https://doi.org/10.1016/j.jpsychires.2017.11.004

Alario, F.-X., Costa, A., & Caramazza, A. (2002). Frequency effects in noun phrase production: Implications for models of lexical access. *Language and Cognitive Processes*, *17*(3), 299–319. https://doi.org/10.1080/01690960143000236

Alpert, M., Shaw, R. J., Pouget, E. R., & Lim, K. O. (2002). A comparison of clinical ratings with vocal acoustic measures of flat affect and alogia. *Journal of Psychiatric Research*, *36*(5), 347–353. https://doi.org/10.1016/S0022-3956(02)00016-X

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Publishing.

Andreasen, N. C., Flaum, M., & Arndt, S. (1992). The Comprehensive Assessment of Symptoms and History (CASH). An instrument for assessing diagnosis and psychopathology. *Archives of General Psychiatry*, *49*(8), 615–623. https://doi.org/10.1001/archpsyc.1992.01820080023004

Barch, D. M., & Ceaser, A. (2012). Cognition in schizophrenia: Core psychological and neural mechanisms. *Trends in Cognitive Sciences*, *16*(1), 27–34. https://doi.org/10.1016/j.tics.2011.11.015

Barker, M. S., Nelson, N. L., & Robinson, G. A. (2020). Idea formulation for spoken language production: The interface of cognition and language. *Journal of the International Neuropsychological Society*, *26*(2), 226–240. https://doi.org/10.1017/S1355617719001097

Becker, T. M., Cicero, D. C., Cowan, N., & Kerns, J. G. (2012). Cognitive control components and speech symptoms in people with schizophrenia. *Psychiatry Research*, *196*(1), 20–26. https://doi.org/10.1016/j.psychres.2011.10.003

Bleuler, E. (1911). *Dementia praecox oder Gruppe der Schizophrenien* [Dementia praecox or the group of schizophrenias] (1st ed.). Diskord.

Boersma, P., & Weenink, D. (2013). Praat [Computer software]. University of Amsterdam.

Bóna, J. (2014). Temporal characteristics of speech: The effect of age and speech style. *The Journal of the Acoustical Society of America*, *136*(2), EL116–EL121. https://doi.org/10.1121/1.4885482

Bowie, C. R., & Harvey, P. D. (2008). Communication abnormalities predict functional outcomes in chronic schizophrenia: Differential associations with social and adaptive functions. *Schizophrenia Research*, *103*(1–3), 240–247. https://doi.org/10.1016/j.schres.2008.05.006

Brown, M., & Kuperberg, G. R. (2015). A hierarchical generative framework of language processing: Linking language perception, interpretation, and production abnormalities in schizophrenia. *Frontiers in Human Neuroscience*, *9*, Article 643. https://doi.org/10.3389/fnhum.2015.00643

Brown, R. W., & Lenneberg, E. H. (1954). A study in language and cognition. *Journal of Abnormal and Social Psychology*, *49*(3), 454–462. https://doi.org/10.1037/h0057814

Califf, R. M. (2018). Biomarker definitions and their applications. *Experimental Biology and Medicine*, *243*(3), 213–221. https://doi.org/10.1177/1535370217750088

Carroll, J. B. (1964). *Language and thought*. Prentice-Hall.

Carruthers, S. P., Van Rheenen, T. E., Gurvich, C., Sumner, P. J., & Rossell, S. L. (2019). Characterising the structure of cognitive heterogeneity in schizophrenia spectrum disorders. A systematic review and narrative synthesis. *Neuroscience and Biobehavioral Reviews*, *107*, 252–278. https://doi.org/10.1016/j.neubiorev.2019.09.006

Cavelti, M., Kircher, T., Nagels, A., Strik, W., & Homan, P. (2018). Is formal thought disorder in schizophrenia related to structural and functional aberrations in the language network? A systematic review of neuroimaging findings. *Schizophrenia Research*, *199*, 2–16. https://doi.org/10.1016/j.schres.2018.02.051

Cohen, A. S., Dinzeo, T. J., Donovan, N. J., Brown, C. E., & Morrison, S. C. (2015). Vocal acoustic analysis as a biometric indicator of information processing: Implications for neurological and psychiatric disorders. *Psychiatry Research*, *226*(1), 235–241. https://doi.org/10.1016/j.psychres.2014.12.054

Cohen, A. S., Mitchell, K. R., Docherty, N. M., & Horan, W. P. (2016). Vocal expression in schizophrenia: Less than meets the ear. *Journal of Abnormal Psychology*, *125*(2), 299–309. https://doi.org/10.1037/abn0000136

Cohen, A. S., Mitchell, K. R., & Elvevåg, B. (2014). What do we really know about blunted vocal affect and alogia? A meta-analysis of objective assessments. *Schizophrenia Research*, *159*(2–3), 533–538. https://doi.org/10.1016/j.schres.2014.09.013

Cohen, A. S., Morrison, S. C., Brown, L. A., & Minor, K. S. (2012). Towards a cognitive resource limitations model of diminished expression in schizotypy. *Journal of Abnormal Psychology*, *121*(1), 109–118. https://doi.org/10.1037/a0023599

Çokal, D., Zimmerer, V., Turkington, D., Ferrier, N., Varley, R., Watson, S., & Hinzen, W. (2019). Disturbing the rhythm of thought: Speech pausing patterns in schizophrenia, with and without formal thought disorder. *PLoS ONE*, *14*(5), Article e0217404. https://doi.org/10.1371/journal.pone.0217404

Corcoran, C. M., & Cecchi, G. A. (2020). Using language processing and speech analysis for the identification of psychosis and other disorders. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *5*(8), 770–779. https://doi.org/10.1016/j.bpsc.2020.06.004

Corcoran, C. M., Mittal, V. A., Bearden, C. E., E Gur, R., Hitczenko, K., Bilgrami, Z., Savic, A., Cecchi, G. A., & Wolff, P. (2020). Language as a biomarker for psychosis: A natural language processing approach. *Schizophrenia Research*, *226*, 158–166. https://doi.org/10.1016/j.schres.2020.04.032

Deary, I. J., Bastin, M. E., Pattie, A., Clayden, J. D., Whalley, L. J., Starr, J. M., & Wardlaw, J. M. (2006). White matter integrity and cognition in childhood and old age. *Neurology*, *66*(4), 505–512. https://doi.org/10.1212/01.wnl.0000199954.81900.e2

de Boer, J. N., Brederoo, S. G., Voppel, A. E., & Sommer, I. E. C. (2020). Anomalies in language as a biomarker for schizophrenia. *Current Opinion in Psychiatry*, *33*(3), 212–218. https://doi.org/10.1097/YCO.0000000000000595

de Boer, J. N., van Hoogdalem, M., Mandl, R. C. W., Brummelman, J., Voppel, A. E., Begemann, M. J. H., van Dellen, E., Wijnen, F. N. K., & Sommer, I. E. C. (2020). Language in schizophrenia: Relation with diagnosis, symptomatology and white matter tracts. *NPJ Schizophrenia*, *6*(1), Article 10. https://doi.org/10.1038/s41537-020-0099-3

de Boer, J. N., Voppel, A. E., Brederoo, S. G., Schnack, H. G., Truong, K. P., Wijnen, F. N. K., & Sommer, I. E. C. (2021). Acoustic speech markers for schizophrenia-spectrum disorders: A diagnostic and symptom-recognition tool. *Psychological Medicine*. Advance online publication. https://doi.org/10.1017/S0033291721002804

de Boer, J. N., Voppel, A. E., Brederoo, S. G., Wijnen, F. N. K., & Sommer, I. E. C. (2020). Language disturbances in schizophrenia: The relation with antipsychotic medication. *NPJ Schizophrenia*, *6*(1), Article 24. https://doi.org/10.1038/s41537-020-00114-3

Dickinson, D., Bellack, A. S., & Gold, J. M. (2007). Social/communication skills, cognition, and vocational functioning in schizophrenia. *Schizophrenia Bulletin*, *33*(5), 1213–1220. https://doi.org/10.1093/schbul/sbl067

Dickinson, D., Pratt, D. N., Giangrande, E. J., Grunnagle, MLin., Orel, J., Weinberger, D. R., Callicott, J. H., & Berman, K. F. (2018). Attacking heterogeneity in schizophrenia by deriving clinical subgroups from widely available symptom data. *Schizophrenia Bulletin*, *44*(1), 101–113. https://doi.org/10.1093/schbul/sbx039

Docherty, N. M., Strauss, M. E., Dinzeo, T. J., & St-Hilaire, A. (2006). The cognitive origins of specific types of schizophrenic speech disturbances. *The American Journal of Psychiatry*, *163*(12), 2111–2118. https://doi.org/10.1176/ajp.2006.163.12.2111

Dunn, J. (2017). Computational learning of construction grammars. *Language and Cognition*, *9*(2), 254–292. https://doi.org/10.1017/langcog.2016.7

Eyben, F., Scherer, K. R., Schuller, B. W., Sundberg, J., Andre, E., Busso, C., Devillers, L. Y., Epps, J., Laukka, P., Narayanan, S. S., & Truong, K. P. (2016). The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for voice research and affective computing. *IEEE Transactions on Affective Computing*, *7*(2), 190–202. https://doi.org/10.1109/TAFFC.2015.2457417

Eyben, F., Weninger, F., Gross, F., & Schuller, B. (2013). Recent developments in OpenSMILE, the Munich open-source multimedia feature extractor. In A. Jaimes, N. Sebe, & N. Boujemaa (Eds.), *MM '13: Proceedings of the 21st ACM International Conference on Multimedia* (pp. 835–838). Association for Computing Machinery.

Ferreira, F. (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language*, *30*(2), 210–233. https://doi.org/10.1016/0749-596X(91)90004-4

First, M. B. (2014). Structured Clinical Interview for the *DSM (SCID)*. In R. L. Cautin and S. O. Lilienfeld (Eds.), *The encyclopedia of clinical psychology* (pp. 1–6). Wiley. https://doi.org/10.1002/9781118625392.wbecp351

Gutiérrez, E. D., Corlett, P. R., Corcoran, C. M., & Cecchi, G. A. (2017). Using automated metaphor identification to aid in detection and

prediction of first-episode schizophrenia. In M. Palmer, R. Hwa, & S. Riedel (Eds.), *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing* (pp. 2923–2930). Association for Computational Linguistics. https://doi.org/10.18653/v1/d17-1316

Hasan, M. R., Jamil, M., & Rahman, M. G. R. M. S. (2004). Speaker identification using mel frequency cepstral coefficients. *Variations*, *1*(4). 565–568. https://scholar.google.com/scholar_lookup?title=Speaker+identification+using+Mel+frequency+Cepstral+coefficients&author=Hasan+M.+R.&author=Jamil+M.&author=Rahman+M.&publication+year=2004&journal=Variations&volume=1&pages=565-568

Hedden, T., & Gabrieli, J. D. E. (2004). Insights into the ageing mind: A view from cognitive neuroscience. *Nature Reviews Neuroscience*, *5*(2), 87–96. https://doi.org/10.1038/nrn1323

Hinzen, W., & Rosselló, J. (2015). The linguistics of schizophrenia: Thought disturbance as language pathology across positive symptoms. *Frontiers in Psychology*, *6*, Article 971. https://doi.org/10.3389/fpsyg.2015.00971

Hitczenko, K., Mittal, V. A., & Goldrick, M. (2021). Understanding language abnormalities and associated clinical markers in psychosis: The promise of computational methods. *Schizophrenia Bulletin*, *47*(2), 344–362. https://doi.org/10.1093/schbul/sbaa141

Holland, R. L. (2016). What makes a good biomarker? *Advances in Precision Medicine*, *1*(1), 66–77. https://doi.org/10.18063/APM.2016.01.007

Ingram, J. C. L., Prandolini, R., & Ong, S. (2013). Formant trajectories as indices of phonetic variation for speaker identification. *International Journal of Speech Language and the Law*, *3*(1), 129–145. https://doi.org/10.1558/ijsll.v3i1.129

Insel, T. R. (2014). The NIMH research domain criteria (RDoC) project: Precision medicine for psychiatry. *The American Journal of Psychiatry*, *171*(4), 395–397. https://doi.org/10.1176/appi.ajp.2014.14020138

Keefe, R. S. E., Goldberg, T. E., Harvey, P. D., Gold, J. M., Poe, M. P., & Coughenour, L. (2004). The Brief Assessment of Cognition in Schizophrenia: Reliability, sensitivity, and comparison with a standard neurocognitive battery. *Schizophrenia Research*, *68*(2–3), 283–297. https://doi.org/10.1016/j.schres.2003.09.011

Keefe, R. S. E., Harvey, P. D., Goldberg, T. E., Gold, J. M., Walker, T. M., Kennel, C., & Hawkins, K. (2008). Norms and standardization of the Brief Assessment of Cognition in Schizophrenia (BACS). *Schizophrenia Research*, *102*(1–3), 108–115. https://doi.org/10.1016/j.schres.2008.03.024

Kerns, J. G., & Berenbaum, H. (2002). Cognitive impairments associated with formal thought disorder in people with schizophrenia. *Journal of Abnormal Psychology*, *111*(2), 211–224. https://doi.org/10.1037/0021-843X.111.2.211

Khawaja, M. A., Ruiz, N., & Chen, F. (2008). Think before you talk: An empirical study of relationship between speech pauses and cognitive load. In N. Bidwell (Ed.), *OZCHI '08: Proceedings of the 20th Australasian Conference on Computer-Human Interaction: Designing for Habitus and Habitat* (pp. 335–338). Association for Computing Machinery. https://doi.org/10.1145/1517744.1517814

Kircher, T. T. J., Brammer, M. J., Levelt, W., Bartels, M., & McGuire, P. K. (2004). Pausing for thought: Engagement of left temporal cortex during pauses in speech. *NeuroImage*, *21*(1), 84–90. https://doi.org/10.1016/j.neuroimage.2003.09.041

Kraepelin, E., & Robertson, G. M. (1919). Dementia praecox. In G. M. Robertson (Ed.), *Textbook of psychiatry* (8th ed.). E. & S. Livingstone.

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*, *14*(1), 41–104. https://doi.org/10.1016/0010-0277(83)90026-4

Liddle, P. F., Ngan, E. T. C., Caissie, S. L., Anderson, C. M., Bates, A. T., Quested, D. J., White, R., & Weg, R. (2002). Thought and language index: An instrument for assessing thought and language in schizophrenia. *The British Journal of Psychiatry*, *181*(4), 326–330. https://doi.org/10.1192/bjp.181.4.326

Lundholm Fors, K. (2015). *Production and perception of pauses in speech* [Doctoral dissertation, University of Gothenburg]. https://gupea.ub.gu.se/handle/2077/39346

Lundin, N. B., Hochheiser, J., Minor, K. S., Hetrick, W. P., & Lysaker, P. H. (2020). Piecing together fragments: linguistic cohesion mediates the relationship between executive function and metacognition in schizophrenia. *Schizophrenia Research*, *215*, 54–60. https://doi.org/10.1016/j.schres.2019.11.032

Maryn, Y., Corthals, P., Van Cauwenberge, P., Roy, N., & De Bodt, M. (2010). Toward improved ecological validity in the acoustic measurement of overall voice quality: Combining continuous speech and sustained vowels. *Journal of Voice*, *24*(5), 540–555. https://doi.org/10.1016/j.jvoice.2008.12.014

Matsumoto, K., Kircher, T. T. J., Stokes, P. R. A., Brammer, M. J., Liddle, P. F., & McGuire, P. K. (2013). Frequency and neural correlates of pauses in patients with formal thought disorder. *Frontiers in Psychiatry*, *4*(4), Article 127. https://doi.org/10.3389/fpsyt.2013.00127

Nolan, F., & Grigoras, C. (2005). A case for formant analysis in forensic speaker identification. *International Journal of Speech Language and the Law*, *12*(2), 143–173. https://doi.org/10.1558/sll.2005.12.2.143

Oliveira, S. E. H., Esteves, F., & Carvalho, H. (2015). Clinical profiles of stigma experiences, self-esteem and social relationships among people with schizophrenia, depressive, and bipolar disorders. *Psychiatry Research*, *229*(1–2), 167–173. https://doi.org/10.1016/j.psychres.2015.07.047

Owen, M. J., Sawa, A., & Mortensen, P. B. (2016). Schizophrenia. *The Lancet*, *388*(10039), 86–97. https://doi.org/10.1016/S0140-6736(15)01121-6

Parola, A., Simonsen, A., Bliksted, V., & Fusaroli, R. (2020). Voice patterns in schizophrenia: A systematic review and Bayesian meta-analysis. *Schizophrenia Research*, *216*, 24–40. https://doi.org/10.1016/j.schres.2019.11.031

Pietrowicz, M., Agurto, C., Norel, R., Eyigoz, E., Cecchi, G., Bilgrami, Z. R., & Corcoran, C. (2019). *A new approach for automating analysis of responses on verbal fluency tests from subjects at-risk for schizophrenia.* https://www.isca-speech.org/archive_v0/Interspeech_2019/pdfs/2987.pdf

Quené, H., Persoon, I., & de Jong, N. (2011). *Praat script syllable nuclei v2* [Praat script]. https://sites.google.com/site/speechrate/Home/praat-script-syllable-nuclei-v2

Reichenberg, A., & Harvey, P. D. (2007). Neuropsychological impairments in schizophrenia: Integration of performance-based and brain imaging findings. *Psychological Bulletin*, *133*(5), 833–858. https://doi.org/10.1037/0033-2909.133.5.833

Rosen, S. (1992). Temporal information in speech: Acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *336*(1278), 367–373. https://doi.org/10.1098/rstb.1992.0070

Salagre, E., Grande, I., Solé, B., Mezquida, G., Cuesta, M., Díaz-Caneja, C., Amoretti, S., Lobo, A., González-Pinto, A., Moreno, C., Pina-Camacho, L., Corripio, I., Baeza, I., Bergé, D., Verdolini, N., Carvalho, A., Vieta, E., & Bernardo, M., & the PEPs Group. (2020). Exploring risk and resilient profiles for functional impairment and baseline predictors in a 2-year follow-up first-episode psychosis cohort using latent class growth analysis. *Journal of Clinical Medicine*, *10*(1), Article 73. https://doi.org/10.3390/jcm10010073

Sbordone, R. J. (1996). Ecological validity: Some critical issues for the neuropsychologist. In R. J. Sbordone & C. J. Long (Eds.), *Ecological validity of neuropsychological testing* (pp. 15–41). Gr Press/St Lucie Press, Inc.

Schmuckler, M. A. (2001). What is ecological validity? A dimensional analysis. *Infancy*, *2*(4), 419–436. https://doi.org/10.1207/S15327078IN0204_02

Schnack, H. G. (2019). Improving individual predictions: Machine learning approaches for detecting and attacking heterogeneity in schizophrenia (and other psychiatric diseases). *Schizophrenia Research*, *214*, 34–42. https://doi.org/10.1016/j.schres.2017.10.023

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., & Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for *DSM–IV* and ICD-10. *The Journal of Clinical Psychiatry*, *59*(Suppl. 20), 22–57. https://www.psychiatrist.com/jcp/neurologic/neurology/mini-international-neuropsychiatric-interview-mini/

Sichlinger, L., Cibelli, E., Goldrick, M., & Mittal, V. A. (2019). Clinical correlates of aberrant conversational turn-taking in youth at clinical high-risk for psychosis. *Schizophrenia Research*, *204*, 419–420. https://doi.org/10.1016/j.schres.2018.08.009

Stanislawski, E. R., Bilgrami, Z. R., Sarac, C., Garg, S., Heisig, S., Cecchi, G. A., Agurto, C., & Corcoran, C. M. (2021). Negative symptoms and speech pauses in youths at clinical high risk for psychosis. *NPJ Schizophrenia*, *7*(1), Article 3. https://doi.org/10.1038/s41537-020-00132-1

Sumner, P. J., Bell, I. H., & Rossell, S. L. (2018). A systematic review of task-based functional neuroimaging studies investigating language, semantic and executive processes in thought disorder. *Neuroscience and Biobehavioral Reviews*, *94*, 59–75. https://doi.org/10.1016/j.neubiorev.2018.08.005

Suvisaari, J., Mantere, O., Keinänen, J., Mäntylä, T., Rikandi, E., Lindgren, M., Kieseppä, T., & Raij, T. T. (2018). Is it possible to predict the future in first-episode psychosis? *Frontiers in Psychiatry*, *9*, Article 580. https://doi.org/10.3389/fpsyt.2018.00580

Tan, E. J., Thomas, N., & Rossell, S. L. (2014). Speech disturbances and quality of life in schizophrenia: Differential impacts on functioning and life satisfaction. *Comprehensive Psychiatry*, *55*(3), 693–698. https://doi.org/10.1016/j.comppsych.2013.10.016

Thuaire, F., Rondepierre, F., Bacon, E., Vallet, G. T., Jalenques, I., & Izaute, M. (2020). Executive functions in schizophrenia aging: Differential effects of age within specific executive functions. *Cortex*, *125*, 109–121. https://doi.org/10.1016/j.cortex.2019.12.003

Verma, M., Patel, P., & Verma, M. (2011). Biomarkers in prostate cancer epidemiology. *Cancers*, *3*(4), 3773–3798. https://doi.org/10.3390/cancers3043773

Voineskos, A. N., Jacobs, G. R., & Ameis, S. H. (2020). Neuroimaging heterogeneity in psychosis: neurobiological underpinnings and opportunities for prognostic and therapeutic innovation. *Biological Psychiatry*, *88*(1), 95–102. https://doi.org/10.1016/j.biopsych.2019.09.004

Voppel, A. E., de Boer, J. N., Brederoo, S. G., Schnack, H. G., & Sommer, I. (2021). Quantified language connectedness in schizophrenia-spectrum disorders. *Psychiatry Research*, *304*, Article 114130. https://doi.org/10.1016/j.psychres.2021.114130

Ward, J. H., Jr. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, *58*(301), 236–244. https://doi.org/10.1080/01621459.1963.10500845

Received January 6, 2021
Revision received October 5, 2021
Accepted October 21, 2021 ■