Voice Patterns in Schizophrenia: A systematic Review and Bayesian Meta-Analysis

Parola Alberto^a, Simonsen Arndis^{b,c}, Bliksted Vibeke^{b,c}, Fusaroli Riccardo^{c,d}

^a Department of Psychology, University of Turin, Italy

^b Psychosis Research Unit - Department of Clinical Medicine, Aarhus University, Denmark

^c The Interacting Minds Center - School of Culture and Society, Aarhus University, Denmark

^d Department of Linguistics, Semiotics and Cognitive Science – School of Communication and Culture, Aarhus University, Denmark

Correspondence concerning this article should be addressed to:

Alberto Parola, Dipartimento di Psicologia, Via Verdi 10, 10123 Turin (Italy).

Phone number: +390116703065

Fax number: +39 0118159039

E-mail: alparola@unito.it

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

Abstract

Voice atypicalities have been a characteristic feature of schizophrenia since its first definitions. They are often associated with core negative symptoms such as flat affect and alogia, and with the social impairments seen in the disorder. This suggests that voice atypicalities may represent a marker of clinical features and social functioning in schizophrenia. We systematically reviewed and meta-analyzed the evidence for distinctive acoustic patterns in schizophrenia, as well as their relation to clinical features. We identified 46 articles, including 55 studies with a total of 1254 patients with schizophrenia and 699 healthy controls. Summary effect sizes (Hedges' g and Pearson's r) estimates were calculated using multilevel Bayesian modeling. We identified weak atypicalities in pitch variability ($g = -0.55$) related to flat affect, and stronger atypicalities in proportion of spoken time, speech rate, and pauses (g 's between -0.75 and -1.89) related to alogia and flat affect. However, the effects were mostly modest (with the important exception of pause duration) compared to perceptual and clinical judgments, and characterized by large heterogeneity between studies. Moderator analyses revealed that tasks with a more demanding cognitive and social component showed larger effects both in contrasting patients and controls and in assessing symptomatology. In conclusion, studies of acoustic patterns are a promising but, yet unsystematic avenue for establishing markers of schizophrenia. We outline recommendations towards more cumulative, open, and theory-driven research.

Keywords:

Acoustic analysis, social communication, machine learning, biomarker, negative symptoms, speech signal.

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

Introduction

Individuals with schizophrenia display atypical voice patterns, qualitatively described in terms of poverty of speech, increased pauses, distinctive tone and intensity of voice (Alpert et al., 2000; Andreasen et al., 1985; Cohen et al., 2016, 2012b; Galynker et al., 2000; Hoekert et al., 2007; Trémeau et al., 2005). Voice atypicalities have been reported since the first definitions of the disorder (Bleuler, 1911; Kraepelin, 1919), are used in the clinical assessment process, and assume an even stronger relevance in the light of growing findings associating voice patterns to cognitive function, emotional states, and social engagement (Cohen and Elvevåg, 2014; Cohen and Hong, 2011; Faurholt-Jepsen et al., 2016; Nevler et al., 2017; Pisanski et al., 2016; Trigeorgis et al., 2016; Tsanas et al., 2011; Wang et al., 2015; Williams et al., 2014; Yin et al., 2007).

Voice atypicalities may thus constitute a window into the underlying clinical and cognitive features of the disorder. Indeed, they have been associated with core negative symptoms of schizophrenia such as blunted affect (e.g. diminished emotional expression, lack of vocal intonation), and alogia (e.g. poverty of speech, latency of speech and blocking) (Murray Alpert et al., 2000; Andreasen, 1984; Cohen et al., 2012; Millan, Fone, Steckler, & Horan, 2014; Ross, Orbelo, Cartwright, et al., 2001). Negative symptoms are included among the primary diagnostic criteria of schizophrenia (DSM-V), and are associated with early age of onset, poor social and functional outcome, reduced quality of life, and poor response to medication and treatment (Couture et al., 2011; Häfner et al., 1999; Rabinowitz et al., 2012; Tandon et al., 2008). Vocal expression also reflects a key component of social communication, a domain frequently impaired in individuals with schizophrenia (Bambini et al., 2016; Bosco et al., 2019; Brüne and Bodenstein, 2005; Champagne-Lavau and Stip, 2010; Colle et al., 2013; Meilijson et al., 2004; Parola et al., 2018). Difficulties in controlling voice

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

to express affective states, e.g. emotional, mood or motivational states, or to mark relevant information may dramatically reduce the ability of these individuals to communicate effectively in social context. Impairments in social communication may in turn lead to experience of failure in social situations, and atypical voice qualities may generate more negative attitudes and social judgments on the part of others than toward speakers with typical voice qualities (e.g. Altenberg & Ferrand, 2006). This, in turn, may result in social withdrawal and the further aggravation of social cognitive impairments (Bambini et al., 2016; Bowie and Harvey, 2008; Del-Monte et al., 2013; Sparks et al., 2010; Tan et al., 2014; Thoma et al., 2009). Voice atypicalities may thus represent an important marker that parallels both clinical features and social cognitive functioning of individuals with schizophrenia over time (Cohen et al., 2016; Rapcan et al., 2010; Tahir et al., 2019).

By understanding the implications of voice abnormalities, we can provide a first step to develop more effective tools to assist clinicians in assessing this heterogeneous disorder. For example, different speech and voice features such as pause number and duration, speech rate and pitch and intensity variability, have been proved to be potential indicators of cognitive load both in healthy individuals (Berthold and Jameson, 1999; Cohen et al., 2015; Khawaja et al., 2008; Yin et al., 2007), and individual with schizotypal disorder (Cohen et al., 2012a). In addition, voice analysis may potentially allow to assess the response to psychosocial or pharmacological treatment over longer periods using objective and quantitative indices, and enhance the capability of clinicians to capture the complex relationship between emotion regulation, expressive behavior, social perception and cognitive and clinical features of the disorder (e.g. Ben-Zeev et al., 2017; Dahlgren et al., 2018; Tahir et al., 2019). Finally, vocal abnormalities may be related to neuromotor disorders frequently associated with schizophrenia and its neurodevelopmental pathophysiology (Cannizzaro et al., 2005; Konopka and Roberts, 2016; Matsumoto et al., 2013; Walther, 2015; Walther and Strik, 2012), or secondary induced by antipsychotic medications (Peluso et al., 2012; Tenback et al.,

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

2010). Clarifying the vocal abnormalities associated with schizophrenia may thus provide tools to support the assessment of quantitative measures of cognitive and clinical features related to voice, such as cognitive load, negative symptomatology and affective states, but also contribute toward a better understanding of the nature and pathophysiology of the disorder.

Despite the importance of studying vocal expression in schizophrenia, and the routine assessments performed using interview-based clinical rating scales, our understanding of voice abnormalities in schizophrenia is limited. Previous work on voice atypicalities can be organized into three categories: qualitative perceptual ratings, quantitative acoustic analyses, and multivariate machine learning (ML) investigations. Most previous studies employing qualitative ratings reported robust differences between patients with schizophrenia and healthy controls (HC) across several perceptual features of their voice, often described as lack of inflection, poverty of speech and latency of speech (Cohen et al., 2014; Emmerson et al., 2009; Hoekert et al., 2007; Mueser et al., 1994). However informative, qualitative rating scales have serious limitations. They rely on raters' expertise and intuition, thus lacking scalability to large corpora, that is, they are expensive and time-consuming to apply to a large number of recordings and require specialized training of the clinicians. Further, they display low sensitivity to complex and multivariate acoustic patterns and variations in context and time (Alpert, Pouget, & Silva, 1995; Cohen & Elvevåg, 2014; Cohen, Mitchell, Docherty, & Horan, 2016a; Cohen et al., 2012). A different approach involves the use of automated analysis of speech to identify acoustic features of vocal production, arguably with a greater reliability, sensitivity and validity. However, such studies have so far reported smaller and seemingly more contradictory findings: some indicate slower speech (Tavano et al., 2008), more pronounced pauses (Cannizzaro et al., 2005; Martínez-Sánchez et al., 2015; Rapcan et al., 2010) and reduced prosodic variability (Compton et al., 2018; Martínez-Sánchez et al., 2015; Ross et al., 2001); while others indicate no reliable acoustic differences between

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

individuals with schizophrenia and HC (Cohen et al., 2016; Docherty, 2012; Meaux et al., 2018). A meta-analysis of 13 studies (Cohen et al., 2014) suggests large differences between individuals with schizophrenia and HC on pause and speech duration, and more modest on intensity and pitch variability. While informative, the number of studies included in the meta-analysis was small compared to the currently available literature and, given the high heterogeneity of patients with schizophrenia, a more systematic review accounting for the potential sources of heterogeneity in the effects is required: individual differences (e.g. gender, age and education), contextual factors (e.g. type of task) and clinical features (e.g. symptomatology and medication). This is crucial if we want to understand the mechanisms underlying voice atypicalities. Different mechanisms may lead to different acoustic patterns in different contexts, e.g. cognitive impairment leading to longer pauses, and even more so when the cognitive and social demands of the task are increased. On the contrary, voice atypicalities due to differences in fine motor control of the vocal folds should be similar across tasks. Further, a few studies have adopted a more fine-grained perspective, and assessed the relationship between acoustic measures and clinical features with some promise; however, the findings are still very sparse (Alpert, Shaw, Pouget, & Lim, 2002; Alpert et al., 2000; Cohen et al., 2016a; Meaux et al., 2018; Zhang, Pan, Gui, Zhu, & Cui, 2016).

Finally, more recent studies have tried to capitalize on the technological advancements in speech signal processing, and the application of multivariate ML techniques to better capture the complex, multivariate and often non-linear nature of acoustic patterns (Bone et al., 2017; Huys et al., 2016; for an introduction to ML techniques in the context of voice analysis see also the appendix to Fusaroli, Lambrechts, Bang, Bowler, & Gaigg, 2017). These studies extract more nuanced acoustic measures, e.g. spectral and glottal features, and assess how accurately the diagnosis can be identified only relying on acoustic measures. The results are promising (Bonneh, Levanon, Dean-Pardo, Lossos, & Adini, 2011; Faurholt-Jepsen et al., 2016; Martínez-Sánchez et al., 2015; Rapcan et al., 2010; Tsanas et al., 2011;

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

Williams et al., 2014), but a complete and comparative overview of the findings in schizophrenia is currently missing. Crucially, the reliability of ML results has been shown to be strongly dependent on the availability of large datasets and the validation of the findings across datasets (Bone et al., 2016; Chekroud, 2018; Foody, 2017; James et al., 2013; Van Der Ploeg et al., 2014), which presence we wanted to assess in the literature on voice in schizophrenia.

Despite the promise of acoustic markers of clinical features in schizophrenia, it is yet unclear how to quantify them, that is, which acoustic features we should focus on, and the evidence for their relation to specific clinical features of the disorder. The aim of the present study was to fill this gap by systematically reviewing and meta-analyzing the current state of evidence for acoustic atypicalities in schizophrenia as a whole as well as their relation to the specific clinical features. Further, we evaluated the size and availability of previous datasets, and the attitudes towards data sharing of the authors of the studies reviewed to assess whether a more cumulative science of voice atypicalities in schizophrenia can be attempted. Note that the aim of this meta-analysis is less to provide a more accurate estimation of the voice atypicalities in schizophrenia than it is to provide the basis for more effective future studies, by identifying current practices, issues and promising venues.

Methods

Inclusion criteria for literature search

We adopted the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Guidelines (PRISMA, Stewart et al., 2015) for transparent reporting of a systematic review. We pre-registered our protocol by specifying a priori the study rationale, eligibility criteria, search strategy, moderator variables, and statistical analyses (see <https://bit.ly/2EEFeQZ>). The literature search was conducted on Pubmed and Google Scholar, the latter including dissertations and unpublished manuscripts. The search terms used were (prosody OR

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

inflection OR intensity OR pitch OR fundamental frequency OR speech rate OR voice quality OR acoustic OR intonation OR vocal) AND (schizo*). The search was conducted on August 21 2017, and updated on April 12 2018. We complemented the list by performing a backward and forward literature search: we screened the bibliography of the papers found, and the papers citing them as identified by Google Scholar.

Articles were screened for eligibility by two authors (AP and AS). Study selection was conducted according to the following inclusion criteria: (a) empirical study, (b) quantification of acoustic features in the vocal production of participants with schizophrenia or schizoaffective disorder¹ (c) sample including at least two individuals with schizophrenia or schizoaffective disorder, (d) inclusion of a non-clinical comparison group, or an assessment of variation in acoustic features in relation to severity of clinical features. Clinical comparison groups (e.g. with depression) were excluded because the limited number of studies did not permit meta-analytic estimations. Fig. 1 shows the flow-diagram of study selection. We report the assessment of the risk of bias in the Supplementary Materials.

¹ We included schizotypy in literature search to better cover schizophrenia spectrum disorder. However, given schizotypy is only included in the schizophrenia spectrum in the ICD and is mentioned in the personality disorders in the DSM classification, we only included schizotypy in additional analysis in the supplementary material. Studies without a control group, i.e. only including patients with schizophrenia, were included in the qualitative synthesis and data extraction as possible additional datapoints for future studies.

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

Figure 1

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

Data extraction

For all the studies we reported the available clinical and demographic data, including pre-registered potential moderations. In particular we report: sample sizes, matching criteria, presence of a non-clinical control group, diagnosis, demographical variables (age, education, gender, language and ethnicity), clinical information (symptom clinical ratings, duration of illness, age of onset, hospitalization), level of intelligence (IQ), cognitive screening, medication. Further we extracted information about the speech production task, group-level acoustic estimates (mean and standard deviation), and correlation coefficients between acoustic measures and clinical ratings. While we extracted details about the speech production tasks (e.g. picture description, video description, biographical episodes), this fine-grained categorization involved too few studies in each category to be useful in a meta-analytic perspective. We therefore grouped speech production tasks into three broader categories: 1) *Constrained production* includes highly structured monological tasks such as reading aloud or repeating sequence of numbers. 2) *Free monological production* includes less constrained monological tasks such as description of pictures or videos, or providing narrative accounts (e.g. of a happy event, or of one's life). Compared to constrained production, free production is more challenging, as the linguistic materials are less pre-defined by the task. 3) *Social interaction* includes structured and semi-structured interviews, as well as spontaneous conversations. The production is dialogical and involves interpersonal factors and dynamics. Compared to the other categories, social interaction is more challenging, as it both involves less pre-defined linguistic materials and the cognitive and social load of having to interact with another person. One could of course envision other categorization typologies; we therefore included all extracted task details in our publicly available dataset. Selected characteristics of included studies are available in Table 1.

Running title: VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

Table 1. Selected characteristics of the studies included in the meta-analysis.

Study	Authors	Included in the meta-analysis	Control group	Sample size and matching criteria	Clinical features	Task	Medication	Findings
1	Pinheiro, Rezaii, Rauber, & Niznikiewicz, 2016	YES	YES	17 SCZ 18 CT Age, sex, education, parental SES	PANNS, SANS, SAPS	Constrained production (reading single words)	NR	Duration of utterance: (words duration/ms): NS Pitch mean (Hz): NS Intensity mean (db): NS
2	Zhang et al., 2016	YES	YES	26 SCZ 30 CT Age, sex, education	PANNS, SANS, CGI-S	Social interaction (phone conversation)	Not medicated	Formants (F1, F2, F3, F4, F5, F6) (Hz/db): NS Formant bandwidth: NS Formants intensity variability (entropy): NS Spectral features: Mel-frequency cepstral coefficient (MFCC): $p < .001$ (Lower in SCZ) Linear prediction coding (LPC): $p < .001$ (Higher in SCZ)
3	Bernardini et al., 2016	YES	NO	20 NA	PANNS, SANS	Spontaneous production (narrative)	NR	Correlations: PANSS TOTAL: Percent time talking: NS; Pitch mean: NS; Pitch variability: NS. SANS TOTAL: Percent time talking: NS; Pitch mean: NS; Pitch variability: NS. PANSS NEGATIVE: Percent time talking: NS; Pitch mean: NS; Pitch variability: NS. SANS FLAT AFFECT: Percent time talking: NS; Pitch mean: NS; Pitch variability: NS; SANS ALOGIA: Percent time talking: NS; Pitch mean: NS; Pitch variability: NS.
3	Bernardini et al., 2016	YES	NO	20 NA	PANNS, SANS	Spontaneous production (narrative)	NR	Correlations: PANSS TOTAL: Percent time talking: NS; Pitch mean: NS; Pitch variability: NS; SANS TOTAL: Percent time talking: NS; Pitch mean: NS; Pitch variability: NS. PANSS NEGATIVE: Percent time talking: NS; Pitch mean: $p = .04$ (Positive); Pitch variability: NS. SANS FLAT AFFECT: Percent time talking: NS; Pitch mean: NS; F0 SD: NS; SANS ALOGIA: Percent time talking: NS; Pitch mean: $p = .005$ (Positive); Pitch variability: $p < .05$ (Negative).
4	Martínez-Sánchez et al., 2015	YES	YES	45 SCZ 35 CT Age, sex, education	BPRS	Constrained interaction (reading)	Medicated	Pause percentage (>300ms): $p < .001$ (Higher in SCZ) Pitch mean (Hz): NS Pitch variability (Hz): NS Intensity variability (Hz): NS Intensity mean (db): $p < .001$ (Lower in SCZ) Correlations: BPRS TOTAL: Proportion of pauses: NS; Pitch mean: NS; Pitch variability: NS; Intensity mean: NS.

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

								BPRS NEGATIVE: Proportion of pauses: NS; Pitch mean: NS; Pitch variability: NS; Intensity mean: NS BPRS POSITIVE: Proportion of pauses: $p = .021$ (Negative); Pitch mean: NS; Pitch variability: NS; Intensity mean: NS
5	Alpert et al., 2002	YES	NO	30 SCZ NA	SANS	Social interaction (interview)	Medicated	Correlations: SANS FLAT AFFECT: Percent time talking: $p < .01$ (Negative); Speech latency: NS; Pitch variability: $p < .05$ (Negative); Intensity variability: NS; SANS ALOGIA: Percent time talking: NS; Speech latency: NS; Pitch variability: NS; Intensity variability: NS
6	Rapcan et al., 2010	YES	YES	39 SCZ 18 CT NR	SANS, BPRS	Constrained interaction (reading brief text)	Medicated	Duration of utterance (s.): $p < .001$ (Shorter in SCZ) Percent time talking: NS Duration of pauses (s.): $p = .011$ Percentage of silence: $p < .001$ (Higher in SCZ) Number of pauses (>250ms): $p < .001$ (Higher in SCZ) Pitch variability (cv): NS Intensity variability: $p < .04$ (Higher in SCZ) Correlations: BPRS TOTAL: Duration of utterance: $p < .01$ (Positive); Proportion of silence: $p < .05$ (Negative); Duration of pauses: NS; Number of pauses: NS; Pitch variability: NS; Intensity variability: $p < .05$ (Positive). NEGATIVE SANS: Duration of utterance: $p < .05$ (Positive); Duration of pauses: NS; Proportion of silence: $p < .05$ (Negative); Number of pauses: NS; Pitch variability: NS; Intensity variability: $p < .05$ (Positive).
7	Cannizzaro, Cohen, Rappard, & Snyder, 2005b	YES	YES	13 SCZ 6 CT NR	PANSS	Constrained (count) + Spontaneous (narrative elicitation)	Medicated	Duration of pauses (>200msec): TASK 1 (Constrained): NS; TASK 2 (Free): $p < .001$ (Higher in SCZ). Percentage of pauses: TASK 1: NS; TASK 2: $p < .001$ (Higher in SCZ) Number of pauses: TASK 1: NS; TASK 2: NS Pause variability: TASK 1: NS; TASK 2: $p < .001$ (Higher in SCZ)
8	Graux, Courtine, Bruneau, Camus, & El-Hage, 2015	YES	YES	26 SCZ 26 CT NR	NR	Constrained (read letters)	Medicated	Pitch mean (Hz): $p = .021$ (Higher in SCZ)
9	McGilloway, Cooper, & Douglas-Cowie, 2003	YES	YES	72 SCZ 40 CT Age, sex	SANS	Constrained (read passage)	Medicated	Duration of pauses: NS Intensity mean: NS
10	Sison, Alpert, Fudge, & Stern, 1996	YES	NO	24 SCZ NA	SANS	Social interaction (interview)	Medicated	Correlations: SANS FLAT AFFECT: Duration of utterance: NS; Duration of pauses: $p = .008$ (Positive); Pitch variability: NS; Intensity variability: NS
11	Cohen, Alpert, Nienow, Dinzeo, & Docherty, 2008	YES	YES	60 SCZ 19 CT Age, sex and parental	SANS, SAPS	Spontaneous production (narrative)	Medicated	Speech rate (words/sec): NS Pitch variability (Hz): $p = .041$ (Lower in SCZ) Correlations: SANS FLAT AFFECT: Speech rate $p < .01$ (Negative);

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

				SEI.				
								Pitch variability: NS; SANS ALOGIA: Speech rate: $p < .01$ (Negative); Pitch variability: NS;
12	Cohen, Kim, & Najolia, 2013	YES	NO	26 SCZ NA	BPRS	Spontaneous production (picture elicitation)	Medicated	<u>Correlations:</u> BPRS TOTAL PSYCHO: Duration of pauses: NS. BPRS NEGATIVE PSYCHO: Duration of pauses: NS BPRS POSITIVE: Duration of pauses: NS
13	Alpert, Kotsaftis, & Pouget, 1997	YES	YES	19 SCZ 20 CT NR	SANS	Social interaction (interview) + Spontaneous production (monologue)	NR	Duration of pauses: $p < .001$ (Higher in SCZ). Response latency: $p < .001$ (Higher in SCZ).
14	Alpert et al., 2000	YES	YES	46 SCZ 20 CT Age, education	SANS	Social interaction (semi-structured interview)	Medicated	Speech rate (words/s): $p < .001$ (Lower in SCZ) Percent time talking: $p = .012$ (Lower in SCZ) Pitch variability (SEMITONES): $p = .001$ (Lower in SCZ) Intensity variability (db): NS <u>Correlations:</u> SANS FLAT AFFECT: Percent time talking: $p < .01$ (Negative); Pitch variability: $p < .01$ (Negative); Intensity variability: NS; SANS ALOGIA: Percent time talking: $p < .01$ (Negative); Pitch variability: NS; Intensity variability: NS
15	Covington et al., 2012	YES	NO	25 SCZ	PANSS	Social interaction (PANSS interview)	Medicated	<u>Correlations:</u> SANS TOTAL: Pitch mean: NS; Pitch variability: NS; PANSS NEGATIVE: Pitch mean: NS; Pitch variability: NS; PANSS POSITIVE: Pitch mean: NS; Pitch variability: NS; SANS FLAT AFFECT: Pitch mean: NS; Pitch variability: NS; SANS ALOGIA: Pitch mean: NS; Pitch variability: NS;
16	Matsumoto et al., 2013	YES	YES	6 SCZ 6 CT Age, education, IQ.	SANS, SAPS	Spontaneous production (picture elicitation)	Medicated	Pause duration (>250ms): NS. Number of pauses: $p = .04$ (Lower in SCZ).
17	Alpert, Clark, & Pouget, 1994	YES	NO	17 SCZ NA	SANS	Social interaction (semi-structured interview) + Spontaneous production (narrative elicitation)	Medicated	<u>Correlations:</u> SANS ALOGIA: Speech rate: $p < .01$ (Negative); Duration of pauses: within-clauses: NS; between-clauses: $p < .01$ (Positive); switching-clauses $p < .01$ (Positive); Number of pauses: within-clauses: NS; between-clauses: NS; switching-clauses: NS; filled pause within-clauses: NS; filled pause between-clauses: NS
18	Kring, Alpert, Neale, & Harvey, 1994	YES	NO	23 SCZ NA	SANS, BPRS	Social interaction (semi-structured interview)	Unmedicated	<u>Correlations:</u> BPRS TOTAL: Percent time talking: NS; BPRS POSITIVE: Percent time talking: NS BPRS NEGATIVE: Percent time talking: NS BPRS BLUNTED: Percent time talking: NS SANS TOTAL: Percent time talking: NS
19	Pinheiro et al., 2017	YES	YES	15 SCZ 16 CT Age. sex. and parental SES	PANNS, SANS, SAPS	Constrained interaction (reading single words)	Medicated	Duration of utterance (words/msec): $p = .028$ (Higher in SCZ). Pitch mean (hz): NS Intensity mean (db): NS

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

20	Resnick & Oltmanns, 1984	YES	YES	10 SCZ 20 CT Age	NR	Social interaction (clinical interview)	Medicated	Duration of pauses (> 250ms): $p = .013$ (Higher in SCZ). Percent time talking : NS
21	Mandal, Srivastava, & Singh, 1990	YES	YES	40 SCZ 60 CT NR	NR	Spontaneous production (facial expression picture elicitation)	Medicated	Speech rate (words/sec): $p < .001$ (Lower in SCZ). Percent time talking : $p < .001$ (Lower in SCZ). Response latency : $p < .001$ (Lower in SCZ).
22	Tavano et al., 2008	YES	YES	37 SCZ 37 CT Age and sex.	BPRS	Spontaneous production (narrative) + Social interaction interview)	Medicated	TASK 1 (Free): Speech rate (words/sec): $p = .027$ (Lower in SCZ). TASK 2 (Social): Speech rate : $p < .001$ (Lower in SCZ).
23	Perlini et al., 2012	YES	YES	30 SCZ 30 CT Age, sex and education	BPRS	Spontaneous production (narrative) + Social interaction (interview)	Medicated	Speech rate : $p = .009$ (Lower in SCZ).
24	Rutter, 1977	YES	YES	12 SCZ 12 CT Sex	NR	Social interaction (conversation)	NR	Speech rate : Task 1: NS; Task 2: NS; Percent time talking : Task 1: $p = .019$ (Lower in SCZ); Tasks 2: NS
25	St-Hilaire, Cohen, & Docherty, 2008	YES	YES	48 SCZ 48 CT Age, sex, parental SES and ethnicity	BPRS	Social interaction (semi-structured interview)	Medicated	Speech rate (words/sec): $p < .001$ (Lower in SCZ).
26	Shaw et al., 1999	YES	NO	30 SCZ	SANS	Social interaction (interview)	Medicated	<u>Correlations</u> : SANS FLAT AFFECT: Duration of pauses : $p < .01$ (Positive); Pitch variability : NS
27	Docherty, 2012	YES	YES	53 SCZ 23 CT Age, sex, parent education, ethnicity	PANSS, PSYRATS	Social interaction (interview)	NR	Speech rate : NS
28	Rochester, Martin, & Thurston, 1977	YES	YES	40 SCZ 20 CT Sex. age	NR	Social interaction (interviews)	Medicated	Percent time talking : $p = .001$ (Lower in SCZ). Duration of pauses : $p < .001$ (Higher in SCZ).
29	Compton et al., 2018	YES	YES	94 SCZ 101 CT Age, ethnicity, and marital status	PANSS, SANS, CAINS	Spontaneous production (narrative) + Constrained interaction (reading)	Medicated	Pitch variability : Task 1 (Free): NS; Task 2 (Constrained): NS <u>Correlations</u> : TOTAL SANS: Pitch variability : $P = .002$ (Positive) NEGATIVE PANSS: Pitch variability : NS SANS FLAT AFFECT: Pitch variability : NS SANS ALOGIA: Pitch variability : $p < .001$ (Positive).
30	Salomé, Boyer, & Fayol, 2002	YES	YES	10 SCZ 10 CT Sex and education	NR	Spontaneous production (narrative)	Medicated	Speech rate : $p = .014$ (Lower in SCZ). Number of pauses : NS
31	Kliper, Vaizman, Weinshall, & Portuguese, 2010	YES	YES	22 SCZ 20 CT NR	SANS	Social interaction (interview) + Constrained interaction (reading)	NR	<u>Correlations</u> : SANS TOTAL: Duration of utterance : $p < .001$ (Negative); Percent time talking : $p < .01$ (Negative); Duration of pauses : $p < .05$ (Positive); Intensity variability : $p < .01$ (Negative)

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

32	Kliper, Portuguese, & Weinshall, 2016	YES	YES	22 SCZ 20 CT Age, sex, and education.	PANSS, SANS	Social interaction (clinical interviews)	NR	<p>Duration of utterance: $p < .001$ (Lower in SCZ). Percent time talking: $p < .001$ (Lower in SCZ). Duration of pauses: $p < .001$ (Higher in SCZ). Pitch variability: $p < .001$ (Lower in SCZ). Intensity variability: $p < .001$ (Higher in SCZ). Correlations: TOTAL PANSS: Duration of utterance: NS; Percent time talking: NS; Duration of pauses: NS; Pitch variability: NS; Intensity variability: NS. SANS TOTAL: Duration of utterance: $p = .04$ (Negative); Percent time talking: $p < .01$ (Negative); Duration of pauses: $p < .01$ (Positive); Pitch variability: NS; Intensity variability: NS. SANS FLAT AFFECT: Duration of utterance: $p < .01$ (Negative); Percent time talking: $p < .01$ (Negative); Duration of pauses: $p < .01$ (Positive); Pitch variability: NS; Intensity variability: NS. SANS ALOGIA: Duration of utterance: $p < .01$ (Negative); Percent time talking: $p < .01$ (Negative); Duration of pauses: $p < .01$ (Positive); Pitch variability: NS; Intensity variability: NS.</p>
33	Ross, Orbelo, Testa, et al., 2001	YES	YES	45 SCZ 19 CT Not matched	SANS, SAPS, BPRS	Constrained production (repetition) + social interaction (interview)	Medication stabilized	<p>TASK 1 (Constrained): Pitch variability: $p < .0001$ (Lower in SCZ). TASK 2 (Social): Pitch variability: $p < .0001$ (Lower in SCZ).</p>
34	Meaux et al., 2018	YES	YES	36 SCZ 25 CT Sex and education	BPRS	Spontaneous production (emotional picture elicitation)	NR	<p>Pitch variability: Task 1 (Free): NS; Task 2: NS. Intensity variability: Task 1 (Free): NS; Task 2: NS. Correlations: BPRS BLUNTED AFFECT: Pitch variability: Task 1 (Free): NS; Task 2 (Free): NS; Intensity variability: Task 1 (Free): NS; Task 2 (Free): NS</p>
35	Püschel, Stassen, Bomben, Scharfetter, & Hell, 1998	NO	YES	45 SCZ 45 CT Sex and age	SANS, PANSS, INSKA	Constrained production (counting and reading passage)	Medicated	<p>Duration of pauses: NS Number of pauses: $p = .0001$ Silence percentage: $p = .0002$ Duration of utterance: $p = .0001$ Total length of pauses: $p = .0001$ Total length of utterances: $p = .0001$ Pitch mean: NS Pitch variability: $p = .0137$ Intensity variability: $p = .0001$</p>
36	Cohen et al., 2016b	NO	YES	309 SCZ 117 CT Not matched	BPRS	Spontaneous production (narrative) + Social interaction (conversation)	NR	<p>Duration of pauses: $p = .002$ Number of pauses: NS Pitch variability: NS Intensity variability: NS Pitch instability: NS Volume instability: NS</p>

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

37	Michelas et al., 2014	NO	YES	10 SCZ 10 CT	PANSS	Constrained production	Medicated	NR.
38	Alpert, Pouget, Welkowitz, & Cohen, 1993	NO	NO	61 SCZ	SANS	Social interaction (neutral interviews) + Spontaneous production	Medicated	Correlations: SANS ALOGIA: Pause duration: $p < .05$; Duration of utterance: NS. SANS FLAT AFFECT: Pause duration: NS; Duration of utterance: $p < .01$.
39	Glaister, Feldstein, & Pollack, 1980	NO	NO	5 SCZ	BPRS	Social interaction (interview)	NR	Correlations: BPRS TOTAL: Proportion of spoken time: $p < .05$.
40	Clemmer, 1980	NO	YES	40 SCZ 40 CT Sex, age, education and ethnicity.	NR	Constrained production (reading passage) --+ spontaneous production (retell passage)	Medicated	NR
41	Stassen et al., 1995	NO	YES	42 SCZ 42 CT Sex and age.	SANS, PANSS, INSKA	Constrained production (counting and reading passage)	Medicated	See Table_S3
42	Leudar, Thomas, & Johnston, 1994	NO	YES	30 SCZ 17 CT Sex, age, and parental social class	SANS	Spontaneous production (description of a sequence of actions)	NR	NR
43	Alpert, Welkowitz, Sobin, Borod, & Rosen, 1989	NO	YES	20 SCZ 21 CT Not matched	SANS	Social interaction (structured interview) + Spontaneous production (monologue)	NR	NR
44	Todder, Avissar, & Schreiber, 2013	NO	YES	15 SCZ 15 CT Sex	PANSS	Social interaction (structured interview)	Not medicated	Shannon entropy: $p < .001$ (Higher in SCZ) Bi-spectral analysis: $p < .01$ (Higher in SCZ) Correlations: POSITIVE PANSS: Shannon entropy: $p < .05$; Bi-spectral analysis: NS. NEGATIVE PANSS: Shannon entropy: NS; Bi-spectral analysis: $p = .026$ TOTAL PANSS: Shannon entropy: NS; Bi-spectral analysis: $p = .01$
45	Cohen et al., 2017	YES	NO	40 SCZ NA	SANS	Social interaction (role-play conversation)	NR	
46	Cohen et al., 2012	YES	NO	26 SCZ NA	BPRS	Spontaneous production (emotional picture elicitation)	Medicated	

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

47	Alpert, 1977	YES	NO	30 SCZ NA	NR	Constrained production (reading brief text)	Medicated
48	Wan, Penketh, Salmon, & Abel, 2008	YES	NO	15 SCZ NA	NR	Social interaction (conversation)	NR
49	Tolkmitt, Helfrich, Standke, & Scherer, 1982	YES	NO	11 SCZ NR	NR	Social interaction (clinical interview)	NR
50	Martínez, Donado, Alberto, Eslava, & Diana, 2015	YES	YES	6 SCZ NR	NR	Spontaneous production (narrative elicitation)	NR
51	Alpert et al., 1995	YES	NO	6 SCZ NA	SANS, SAPS	Social interaction (interview)	NR

Note: We included in the table those studies that reported: 1) descriptive statistics for schizophrenia (SCZ) and healthy controls (HC) groups, correlation coefficients, or statistical tests for these measures 2) estimates for acoustic measures only for group of individuals with schizophrenia, and 3) studies considered as eligible for the meta-analysis, but not included due to missing group-level estimates. P-values of statistical tests comparing individuals with SCZ and HC and correlation coefficients have been extracted from original articles. When estimates for acoustic measures were reported for subgroups of patients, or for different task conditions within the same speech task, we averaged across them weighting the values by sample size, and we then computed independent samples t-test between individuals with SCZ and HC groups. When in the original articles were provided estimates for acoustic measures but not p-value of the comparisons between groups, we computed independent samples t-test (or correlation coefficient) between individuals with SCZ and HC control groups. When the authors provided us original data we recomputed independent samples t-tests (or correlation coefficients) using the original data. Clinical features: PANSS, The Positive And Negative Symptoms Scale; SANS, Scale for the Assessment of Negative Symptoms; SAPS, Scale for the Assessment of Positive symptoms. CGI-S, The Clinical Global Impression Scale; BPRS, Brief Psychiatric Rating Scale; PSYRATS, The Psychotic Symptom Rating Scales; CAINS, The Clinical Assessment Interview for Negative Symptoms; InSka, The Intentionality Scale.

When more than four studies reported statistical estimates for an acoustic measure, or correlation with symptomatology, we performed meta-analysis of the effects. When estimates for acoustic measures were reported for subgroups of patients, or for different task conditions within the same speech task, we averaged across them weighting the values by sample size.

In case of multivariate ML studies, we used a different focus. Multivariate ML approaches differ from the studies previously described in two main ways. While more traditional studies focus on a single feature at time, multivariate ML studies analyze multiple acoustic features simultaneously. While more traditional studies focus on best explaining all the current samples (minimizing within sample error), multivariate ML studies focus on generalizability of the results to new samples (minimizing out-of-sample error), e.g. by using validation and cross-validation techniques. In reviewing ML studies, we focused on reporting the algorithms adopted, the acoustic features considered and the performance of the algorithms in either discriminating individuals with schizophrenia from HC with respect to the acoustic measures considered or predicting the severity of clinical features (e.g. negative symptoms) from acoustic measures (see Table S3 in appendix).

We contacted all authors to obtain missing group-level estimates and individual-level data. Statistics on authors' contact availability, propensity to respond and self-reported barriers to data sharing are also reported.

Statistical analysis

Meta-analyses were performed following well-established procedures (Doebler and Holling, 2015; Field and Gillett, 2010; Quintana, 2015; Viechtbauer, 2010) and complemented by a Bayesian framework (Bürkner et al., 2017; Williams, Rast, & Bürkner, 2018). To estimate the differences in vocal patterns between individuals with schizophrenia and HC we extracted the standardized mean difference (SMD; also known as Hedges' g). To estimate relations

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

between vocal patterns and clinical features we extracted the raw correlation coefficient (Pearson's r). These effects were analyzed using 2-level hierarchical Bayesian regression models to estimate the pooled effect sizes and corresponding credible (i.e., Bayesian confidence) intervals. The multilevel structure allowed us to explicitly model the heterogeneity (or σ^2) in the results of the studies analyzed. By including a random effect by study, we assumed that the variability in experimental design, acoustic analyses and population samples could generate heterogeneous findings and allowed the model to estimate such heterogeneity. We then measured and tested for heterogeneity of the studies using the Cochran's Q statistic (Cochran, 1954), which reveals how much of the overall variance can be attributed to true between-study variance. To analyze the influence of potential moderators explaining between study heterogeneity, meta-regression models were applied separately. Note that only speech task presented enough data points to be analyzed as moderator. Other pre-registered moderators were not sufficiently reported and would have required access to individual level data for adequate treatment.

Priors were chosen to be only weakly informative so that their influence on the meta-analytic estimates were small, only discounting extreme values: a normal distribution centered at 0 (no effect), with a standard deviation of 0.5 for the overall effect, and a positive truncated normal distribution centered at 0, with a standard deviation of 0.5 for the heterogeneity of effects (standard deviation of random effects). We report 95% credible intervals (CIs), i.e. the intervals within which there is a 95% probability that the true value of the parameter (e.g. effect size) is contained, given the assumptions of the model. We provide evidence ratios (ER) and credibility scores. ERs quantify the evidence provided by the posterior estimates in favor of the effect of diagnosis or of clinical feature (e.g. longer pauses in schizophrenia compared to HC) against the alternatives (e.g. same length or shorter pauses in schizophrenia). An ER equal to 3 indicates the hypothesis is 3 times more likely than the alternative. A credibility score indicates the percentage of posterior estimates falling above 0.

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

Because Bayesian methods are less commonly used and understood, we also report p-values in order to reach a broader audience. Note that the p-values are calculated on the same 2-level hierarchical model as the Bayesian inference, with the difference that p-value statistics rely on completely flat priors and assume Gaussian distributions for all estimated parameters.

To assess the potential role of speech production task in explaining the patterns observed, we compared the baseline model with a second multilevel Bayesian model including task as predictor of difference in vocal patterns. We used Leave-One-Out-Information-Criterion (LOOIC) and stacking weights indicating the probability that the model including task is better able to predict new data than baseline (Vehtari et al., 2017).

To explore the possibility of publication bias, potential for funnel plot asymmetry was examined visually and tested using the rank correlation test (Begg and Mazumdar, 1994). The raw data and analysis scripts are available at <https://osf.io/qdkt4/>. The supplementary materials report an additional analysis including schizotypy. All computation was done in R (Core R Team, 2013) relying on metafor, brms and Stan (Burkner, 2013; Carpenter et al., 2017; Viechtbauer, 2010).

Results

3.1 Study selection

See Fig. 1 for full details on the selection. We were able to retrieve relevant statistical estimates from 46 articles (55 studies) from the texts or the authors, with a total of 1254 patients (466 F) with schizophrenia and 699 controls (323 F). We contacted a total of 57 authors – including those of studies that were later deemed ineligible due to lack of statistical estimates – requesting additional information and individual level acoustic estimates for each participant: 40 (70.2%) responded and 10 (18%) provided at least some of the requested data. Chief reasons to decline sharing data were: i) effort required ($n = 15$, 50 %), ii) data loss ($n = 14$, 43.3% of respondents), iii) ethical concerns with data sharing ($n = 3$, 3 %), iv) skepticism

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

towards quantitative meta-analyses ($n = 1$, 3.3%). For more details on the email to the authors and their answers, see Supplementary Materials.

3.2 Differences in acoustic patterns between individuals with schizophrenia and healthy controls

Detailed results are reported in Table 2. Hierarchical Bayesian meta-analyses revealed significant effects of diagnosis (in terms of Hedges' g) on pitch variability (-0.55, 95% CIs: -1.06, 0.09), proportion of spoken time (-1.26, 95% CIs: -2.26, 0.25), speech rate (-0.75, 95% CIs: -1.51, 0.04), and duration of pauses (1.89, 95% CIs: 0.72, 3.21), see Fig. 2. No significant effect was found for pitch mean (0.25, 95% CIs: -0.72, 1.30), intensity variability (0.739, 95% CIs: -2.01, 3.39), duration of utterance (-0.155, 95% CIs: -2.56, 2.26) and number of pauses (0.05, 95% CIs: -1.23, 1.13). We generally found high heterogeneity between studies, indicating a likely high diversity in samples and methods, and publication bias, indicating a tendency to publish only significant results, thus making the published literature not fully representative of the actual population of study (see Table 2).

Figure 2

Table 2. Main Results of the meta-analysis for the effect of diagnosis on acoustic measures, and for the correlations between acoustic measures and symptoms ratings.

Effect of diagnosis on acoustic measures									
Acoustic Features	Participants (female) and median	Number of studies (articles)	Influential study	Estimates -Hedges'g [95% CI]	P - value	ER (Credibility)	Sigma squared [95% CI]	Q- stats (p-value)	Publication bias
Pitch Mean	103 SCZ (22 F) 95 CT (23 F)	4 (4)	Yes, Martinez Sanchez et al. (2015)	0.253 [-0.724, 1.304]	.273	3.467 (78%)	1.131 [0.005, 7.19]	8.282 (p = .041)	No, K = 0.0, p = 1.0
			Removing influential	0.505 [-0.859, 1.994]	.273	9.471 (90%)	2.143 [0.0, 16.934]		
Pitch variability	387 SCZ (92F) 257 CT (106 F)	11 (8)	No	-0.546 [-1.059, -0.095]	.005	99.0 (99%)	0.566 [0.152, 1.64]	58.895 (p < .001)	Yes, K = -0.709, p = .002
Intensity variability	104 SCZ (22 F) 65 CT (20 F)	4 (3)	Yes, Kliper et al. (2015)	0.739 [-2.01, 3.388]	.164	3.648 (78%)	7.888 [0.56, 47.934]	35.89 (p < .001)	No, K = 1.0, p = .083
			Removing influential	0.152 [-1.005, 1.311]	.164	2.62 (72%)	1.932 [0.00, 13.486]		
Proportion of spoken time	267 SCZ (106 F) 211 (98 F)	11 (9)	No	-1.26 [-2.257, -0.245]	.001	149.943 (99%)	2.538 [0.787, 7.224]	113.308 (p < .001)	No, K = -0.236 p = .359
Duration of utterance	93 SCZ (30 F) 72 CT (30 F)	4 (4)	No	-0.155 [-2.556, 2.26]	.739	1.475 (60%)	6.045 [0.32, 37.49]	27.78 (p < .001)	No, K = 0.333, p = .75
Speech rate	336 SCZ (111 F) 259 (107 F)	11 (9)	No	-0.75 [-1.514, 0.036]	.015	32.473 (97%)	1.447 [0.467, 3.915]	104.414 (p < .001)	No, K = -0.055, p = .879
Duration of pauses	221 SCZ (128 F) 150 CT (92 F)	9 (8)	No	1.891 [0.721, 3.213]	< .001	234.294 (100%)	3.129 [0.754, 10.086]	75.624 (p < .001)	Yes, K = 0.667 p = .013
Number of pauses	68 SCZ (23 F) 40 CT (13 F)	5 (4)	Yes, Matsumoto et al. (2013)	0.046 [-1.225, 1.131]	.782	1.321 (57%)	1.531 [0.017, 8.496]	11.61 (p = .02)	Yes, K = -1.0, p = .017
			Removing influential	0.355 [-0.991, 1.615]	.782	4.739 (83%)	1.481 [0.001, 10.692]		

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

Correlations between acoustic measures and clinical symptoms ratings

Acoustic Features	Clinical features	Number of studies	Participants	Influential study	Pearson's r 95% CI	P - value	ER (credibility)	Sigma squared 95% CI	Q- stats	Publication bias
Pitch Mean	Negative symptoms	7 (3)	107 SCZ (33 F) 21	No	0.096 [-0.158, 0.346]	.136	4.198 (81%)	0.071 [0.00, 0.345]	9.976 (p = .126)	No, K = -0.619, p = .069
	Positive symptoms	4 (3)	107 (33 F) 21	No	-0.185 [-0.691, 0.316]	.04	6.89 (87%)	0.245 [0, 1.714]	3.586 (p = .31)	No, K = 0.333, p = .75
Pitch variability	General psychopathology	5 (4)	146 (48 F) 22	No	-0.091 [-0.34, 0.15]	.3	3.84 (79%)	0.057 [0, 0.354]	2.283 (p = .684)	No, K = 0, p = 1.0
	Negative symptoms	11 (6)	261 (77 F) 22		-0.01 [-0.196, 0.144]	.836	1.117 (53%)	0.041 [0.002, 0.135]	19.292 (p = .037)	No, K = -0.2, p = 0.445
	Positive symptoms	4 (3)	107 (33 F) 21	Yes, Covington et al., (2012)	-0.027 [-0.686, 0.763]	.698	1.509 (60%)	0.525 [0.001, 4.294]	7.248 (p = .064)	No, K = 0.333, p = .75
				Removing influential	-0.05 [-0.715, 0.62]	.755	1.517 (60%)	0.422 [0.001, 2.944]		
	Alogia rating	9 (7)	313 (68 F) 26	No	-0.035 [-0.317, 0.22]	.465	1.5 (60%)	0.135 [0.032, 0.421]	45.478 (p < .001)	No, K = -0.314, p = .246
	Flat affect rating	13 (10)	403 (81 F) 30	No	-0.106 [-0.262, -0.047]	.044	11.719 (92%)	0.053 [0.009, 0.153]	32.763 (p = .001)	No, K = -0.117 p = .582
Intensity variability	Flat affect rating	6 (5)	158 (22 F) 30	No	-0.005 [-0.324, 0.308]	.745	1.03 (51%)	0.117 [0.001, 0.658]	10.219 (p = .069)	No, K = -0.067, P = 1.0
Proportion of spoken time	General psychopathology	5(4)	124 (35 F) 22	Yes, Rapcan et al. (2010)	-0.026 [-0.53, 0.375]	.85	1.714 (63%)	0.268 [0.005, 1.411]	11.475 (p = .022)	No, K = -0.4, p = .483
				Removing influential	-0.069 [-0.536, 0.335]	.662	1.816 (65%)	0.235 [0.005, 1.335]		
	Negative psychopathology	9 (5)	146 (35 F) 22	No	-0.229 [-0.499, 0.035]	.198	23.29 (96%)	0.131 [0.027, 0.405]	35.506 (p < .001)	No, K = 0.333, p = .26
	Alogia rating	5 (4)	138 (23 F) 22	No	-0.413 [-0.723, -0.07]	< .001	58.259 (98%)	0.127 [0, 0.805]	7.344 (p = .119)	No, k = 0.333, p = .435
	Flat affect rating	6 (5)	161 (23 F) 22.5	No	-0.384 [-0.612, -0.082]	< .001	83.211 (99%)	0.08 [0, 0.456]	7.901 (p = .162)	Yes, K = .867, p = .017

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

Duration of pauses	Negative psychopathology	4 (4)	109 (30 F) 24	Yes, Rapcan et al. (2010)	0.302 [-0.199, 0.783]	.003	15.667 (94%)	0.246 [0, 1.754]	4.971 (p = .174)	No, K = 0.333, p = .75
				Removing influential	0.295 [-0.211, 0.757]	= .008	14.267 (93%)	0.37 [0, 2.129]		

Note: CI, credible interval; P values are 2-tailed; Evidence ratio (ER) quantify the evidence provided by the data in favor of the effect of associations between clinical features and acoustic measures (e.g. longer pauses associated to higher rating of alogia) against the alternatives (e.g. no association). An ER equal to 3 indicates the hypothesis is 3 times more likely than the alternative. A credibility score indicates the percentage of posterior estimates falling above 0.

Moderator analysis

For detailed results, see Table S1 (in appendix). Adding the speech production task employed systematically increased the explained variability (stacking weights: 100%) in schizophrenia atypicalities for pitch variability (effects in Constrained Monologue, CM: -0.54; Free Monologue, FM: -0.11; Dialogue, D: -1.38), proportion of spoken time (CM: -0.38; FM: -2.7; D: -1.48), speech rate (CM: NA; FM: -0.98; D: -0.77), number of pauses (CM: 0.17; FM: -0.09; D: NA), and intensity variability (CM: NA; FM: -0.19; D: 1.45). In general, we observe that dialogical and free speech show the biggest differences, while constrained monologue displays the smallest schizophrenia atypicalities in vocal patterns, except for pitch variability.

3.3 Correlation between acoustic measures and clinical ratings

For detailed results, see Table 2. Hierarchical Bayesian meta-analysis revealed significant overall correlation between flat affect and pitch variability (-0.11, 95% CIs: -0.26, 0.05) and proportion of spoken time (-0.38, 95% CIs: -0.61, -0.08), alogia and proportion of spoken time (-0.41, 95% CIs: -0.72, 0.07), positive symptoms and pitch mean (-0.19, 95% CIs: -0.69, 0.32), negative symptoms and pause duration (0.30, 95% CIs: -0.20, 0.78), see Fig. 3. No significant correlation was found between flat affect and intensity variability (-0.01, 95% CIs: -0.32, 0.31), alogia and pitch variability (-0.04, 95% CIs: -0.32, 0.22), general psychopathology and proportion of spoken time (-0.03, 95% CIs: -0.53, 0.375) and pitch variability (-0.09, 95% CIs: -0.34, 0.15), positive symptoms and pitch variability (-0.03, 95% CIs: -0.68, .076), negative symptoms and pitch mean (0.01, 95% CIs: -0.16, 0.35), pitch variability (-0.01, 95% CIs: -0.20, 0.14), and proportion of spoken time (-0.23, 95% CIs: -0.50, 0.04) (see Table 2). We generally found very high heterogeneity between studies, and publication bias, which suggests caution in interpreting these results and the need for more systematic replications (see Table 2).

Figure 3

Moderator analysis

For detailed results, see Table S2 (in appendix). Adding speech task to the model credibly improved it (stacking weights: 100%) for correlations between *pitch variability* and positive symptom severity (effects in Constrained Monologue, CM: -0.001; Free Monologue, FM: -0.40; Dialogue, D: 0.79), negative symptom severity (CM: 0.01; FM: 0.17; D: -0.43), alogia (FM: 0.10; D: -0.24) and flat affect (FM: 0.03; D: -0.30), between *intensity variability* and flat affect (FM: 0.28; D: -0.12), and between *proportion of spoken time* and total psychopathology (CM: 0.37; FM: -0.29; D: -0.15), negative symptom severity (CM: 0.49; FM: -0.34; D: -0.46), alogia (FM: -0.35; D: -0.46) and flat affect (FM: -0.21; D: -0.44). In general, we see that dialogic speech shows the strongest correlations with symptomatology, and constrained monological speech the weakest ones.

3.4. Multivariate machine learning (ML) studies

We found 5 ML articles fitting our criteria, all focused on identifying acoustic markers of the disorder (Klipper et al., 2016; Martínez-Sánchez et al., 2015; Püschel et al., 1998; Rapcan et al., 2010; Stassen et al., 1995) and 1 including the prediction of severity of clinical features from acoustic measures (Püschel et al., 1998). All acoustic features employed are reported in Table S3. Four studies employed linear discriminant analysis (LDA) and one employed support vector machines (SVM) to classify individuals with schizophrenia vs. HC. LDA linearly combines the acoustic features to generate a distribution of probability of a new voice belonging to an individual with schizophrenia. SVMs, on the contrary, construct multi-dimensional hyperplanes defined by the features and their interactions and there identify the regions that best separate the two groups. SVMs better account for non-linear relations between features and diagnosis. All studies reported accuracy beyond 75% and up to 87.5%. Results of four studies were cross-validated (Martínez-Sánchez et al., 2015; Rapcan et al.,

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

2010; Stassen et al., 1995). Only one study (Rapcan et al., 2010) reported additional performance indices such as specificity, sensitivity, and area under the curve (AUC).

Two studies (Püschel et al., 1998; Stassen et al., 1995) attempted to predict the symptomatology (negative symptoms severity) from acoustic measures. The studies relied on LDA and reported an accuracy of 78.6% and 75.9% in classifying individuals with schizophrenia with higher vs lower scores of negative symptoms (PANSS negative < 11 or < 21 and SANS < 13 and < 37), and 71.4 % accuracy in predicting a future (14 days after) measurement of negative symptoms.

Discussion

Overview

Early descriptions of schizophrenia point to atypical voice patterns and studies relying on perceptual judgments and clinical ratings of voice patterns have indeed found large differences between patients and controls (Cohen et al., 2014). This suggests the existence of acoustic markers of the disorder. We set out to systematically review and meta-analyze the literature on the topic to assess the evidence for atypical acoustic patterns as markers of the disorder and to better inform future research. We were able to analyze the aggregated data from 46 unique articles including 1254 individuals with schizophrenia and 699 HC. The univariate studies identified several null results, as well as weak atypicalities in pitch variability (perhaps in relation to flat affect), and stronger atypicalities in duration (possibly related to alogia and flat affect). The effect sizes suggest a within-sample discriminative accuracy between 66% and 80%, likely less if assessing new data. The multivariate ML studies paint a more promising picture, with overall out-of-sample accuracies between 76.5% and 87.5%. When assessing the relation between acoustic features and symptomatology, we found that specific symptoms that are more directly related to voice, e.g. in their description

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

in the clinical scales, yield slightly stronger results, with flat affect being related to speech variability and proportion of spoken time; and alogia being related to proportion of spoken time. Further, the results across all analyses suggest that dialogical productions, that is, tasks with a perhaps higher cognitive load and a more demanding social component, tend to involve larger effect sizes both in contrasting patients and controls and in assessing symptomatology. Free monological production follows and constrained production produces generally the smallest effects. Crucially, the studies analyzed mostly used widely different methods for sample selection, acoustic pre-preprocessing, feature extraction and selection. Indeed, we find large heterogeneity in the findings of the analyzed studies, and a large uncertainty in all our meta-analytic estimates.

What have we then learned? In line with a previous non-systematic meta-analysis (13 studies, Cohen et al., 2014), we do indeed find evidence for acoustic markers of schizophrenia, further supporting the relation between clinical features of schizophrenia and voice patterns. However, the effect sizes are mostly too small for practical applications, not comparable to those of perceptual and clinical judgments, and in any case plagued by large between-studies variability. The only effect size which revealed to be large and robust was pause duration, specifically in relation with negative symptomatology, suggesting the potential of pause as a marker of schizophrenia symptomatology, which should be further investigated. While good progress has been made in the field, the review highlights a number of issues to be overcome to more satisfactorily understand acoustic patterns in schizophrenia and their potential. In particular we identified the following obstacles to the scientific understanding of acoustic features in schizophrenia: i) small sample sizes in terms of both participants and repeated measures, ii) heterogeneous, not fully up-to-date and underspecified methods in data collection and analysis, leading to scarce comparability between studies; iii) very limited

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

attempts at theory driven research directly tackling the mechanisms underlying atypical vocal patterns in schizophrenia. These are discussed below.

Sample size. Schizophrenia is a heterogeneous disorder, and indeed several studies attempted to more specifically investigate the relation of acoustic features with the symptomatology of the disorder. However, given the limited meta-analytic effect sizes and the awareness that replications tend to show a marked shrinkage of effect sizes (Klein et al., 2018), we need to move beyond small heterogeneous studies. The majority of the studies analyzed include between 20 and 30 patients, plausibly due to the difficulty in accessing clinical populations. However, an expected Cohen's d of 0.6 (pitch variability) would require at least 74 participants per group to reach a 95% power (calculations relying on G*Power (Faul et al., 2009)) at which effect size estimates are reliable (Schönbrodt and Perugini, 2013). If we considered the more conservative possibility of a smaller true effect size of 0.3, the required sample size would be 290 participants per group. While including as varied a sample as possible is an unavoidable concern, there are strategies to reduce the sample size needed. For instance, one could employ repeated measures, that is, collecting repeated voice samples over time. Using 10 repeated measures per participant brings the required sample from 290 participants per group to 82 (assuming that they are still representative of the full population). Repeated measures are also very useful to better understand the reliability of the acoustic patterns over re-testing and potentially across different contexts. In particular, we have seen that dialogical speech production tasks might yield stronger vocal differences, but without a controlled within-subject contrast it is difficult to assess whether this is due to the nature of the task or to other confounds in the sample and study design. We therefore recommend the use of multiple tasks systematically varying the cognitive and social demands involved in speech production.

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

Bigger sample sizes enable also a more fine-grained investigation of individual differences, such as the role of demographical (age, education, gender, language and ethnicity), cognitive and clinical features of the participants on their voice patterns and relatedly on the vocal markers of schizophrenia. Analyzing how acoustic features vary with fine-grained individual differences and context of speech production can help uncover the mechanisms behind atypical vocal patterns and provide an additional insight into schizophrenia. Indeed, we observe that acoustic features are more strongly related to specific symptoms (alogia, flat affect) than to global scores of psychopathologies.

Methods. We found that the field predominantly focuses on traditional acoustic features: pitch, intensity and duration measures. Even in these cases, the processing of the voice recordings and extraction of the features is poorly documented and arguably widely heterogeneous. Previous studies have found that different assumptions and settings in the feature extraction process might significantly affect the results (e.g. Kiss, van Santen, Prud'Hommeaux, & Black (2012) show different results for different choice of ceiling in pitch extraction). Further, speech pathology and speech signal processing research has developed a wide array of acoustic features more directly relatable to production mechanisms like fine-grained muscle control, or clarity of articulation (for some examples see Cummins, Sethu, Epps, Schnieder, & Krajewski, 2015), which are almost completely ignored in schizophrenia research. To overcome these barriers, we recommend the use of freely available open source software solutions providing standard procedures in the extraction of acoustic features and the documentation of the settings chosen (Boersma and Weenink, 2018; Degottex et al., 2014; Eyben et al., 2010)². Use of new features should be compared against this baseline to facilitate comparability between studies.

² Note that while the bibliographic references date to 2010 and 2014, the referred software has been frequently updated since.

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

Further, the vast majority of the studies focused on one acoustic feature at a time failing to produce effects comparable to those found in perceptual judgment studies. This supports the idea that perception is a complex process, non-linearly combining multiple acoustic cues. Multivariate techniques may thus allow to better capture vocal atypicalities. Indeed, the four ML studies we were able to identify provide promising out-of-sample accuracies, indicating that voice of individuals with schizophrenia may contain enough information to reliably distinguish between the two populations. However, the almost complete lack of overlap in features and methods employed in these studies makes it hard to assess how reliable the findings are across samples and whether there are more promising features and algorithms we should focus on.

Theory-driven research. A common feature of many of the studies reviewed is the lack of theoretical background. For example, limited attention is paid to clinical features and their severity and the choice of the speech-production task and acoustic measures used is often under-motivated. On the contrary, by putting hypothesized mechanisms to the test, more theory-driven research on vocal production in schizophrenia would improve our understanding of the disorder itself. For instance, social cognitive impairments (Green et al., 2015; Penn et al., 2008; Sergi et al., 2007) would motivate hypotheses on prosodic patterns when speaking to an interlocutor enabling the investigation of the impact of socio-cognitive mechanisms on vocal production, often hypothesized to be impaired in neuropsychiatric disorders. Comparison of narrative description and reading allows testing for the impact of word search (alogia) against a more general lack of motivation and energy (Frith, 1992; Lysaker and Bell, 1995; Trémeau et al., 2013), and adding a sustained phonation (say “aaaaa”) for the role of fine motor control of the vocal fold. Thus, using several tasks might help testing mechanistic hypotheses, and relatedly allow to assess whether speakers with a different clinical profile show differential vocal patterns across tasks. By including different

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

tasks with diverse cognitive and social constraints, it would be possible to produce more robust results not specifically bound to a specific context, and to investigate the mechanisms and contextual factors responsible for voice abnormalities. Further, increased attention on the mechanisms should enable a more cross-diagnostic perspective, assessing the presence of atypicalities specific to schizophrenia or to more general cognitive and clinical features (Borsboom, 2017). For example, in other work we show that reduced pitch variability might be specific to schizophrenia, when compared to autism spectrum disorder and right hemisphere damage (Fusaroli et al., 2017; Weed and Fusaroli, 2019). However, we would expect to find reduced pitch variability also in affective disorders (Cummins et al., 2015b, 2015a, 2014). An approach based on clinical and cognitive profiles would provide more informative support to the clinicians than just the identification of the diagnostic class.

Further, little work has been done on the physiological mechanisms underlying atypical voice production in schizophrenia, e.g. whether related to auditory processing, pitch control, neuromotor disorders, or antipsychotic medications (Cannizzaro et al., 2005; Konopka and Roberts, 2016; Matsumoto et al., 2013; Peluso et al., 2012; Tenback et al., 2010; Walther, 2015; Walther and Strik, 2012). This is an important venue for further investigations.

Finally, speech involves not only voice but also more linguistic aspects: for instance, lexical choices, syntactic and semantic structure are all mentioned in the symptomatology of the disorder (e.g. poverty of content and tangentiality). Investigating these additional features of speech complementing acoustic analyses could yield further insight. For example, recent research (Çokal et al., 2019) showed that only syntactically motivated and not general pause patterns reliably distinguished participants with schizophrenia from non-clinical controls. Further, automated measures of semantic coherence are being developed to assess symptoms like tangentiality (Corcoran et al., 2018). It is an open question how they relate to vocal patterns and how they can complement each other.

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

Open Science. The recommendations to rely on large sample sizes, include individual differences, and cumulatively employ acoustic features from previous studies might seem too cumbersome, or even unreasonable, given the high costs of research, ethical and practical constraints in accessing clinical populations and proliferation of acoustic measures. This is why we recommend open science practices to be included already in the research design. Releasing in controlled and ethically sound ways one's datasets enables the construction of large collective samples and re-analysis of the data to replicate and extend previous findings. Accessing previous datasets is currently unfeasible, due to lack of answers from corresponding authors, data loss and the practical and time-consuming hurdle of finding, preparing and sharing the data years after the study has been published. This suggests that planning data-sharing from the onset of the study is necessary to ensure a more open, collective and nuanced science of acoustic markers in schizophrenia, conscious of the individual differences and diverse symptomatology. Sharing identifiable (voice) data related to clinical populations requires serious ethical considerations and careful sharing systems, but there are available datasets of voice recordings in e.g. people with Parkinson's, bipolar disorder, depression and autism spectrum disorder (Ambite et al., 2015; Cummins et al., 2015b; Gratch et al., 2014; Schuller et al., 2013; Tsanas et al., 2014), thus suggesting that these hurdles can be overcome. In line with these recommendations, all the data and the codes used in this manuscript are available at <https://osf.io/qdkt4/>.

Conclusion

We have systematically reviewed the evidence for acoustic markers of schizophrenia and its symptomatology, as well as the research practices employed. We did not find conclusive evidence for clear acoustic markers of schizophrenia, although pitch variability and duration are potential candidates, with strong but heterogeneous evidence in favor of the use of pause

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

duration. Multivariate studies are more promising, but their generalizability across samples could not be assessed. To advance the study of vocal markers of schizophrenia we outlined a series of recommendations towards more cumulative, open, and theory-driven research.

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

Reference List

- Alpert, M., 1977. Imagery Mediation of Vocal Emphasis in Flat Affect. *Arch. Gen. Psychiatry* 34, 208. <https://doi.org/10.1001/archpsyc.1977.01770140098011>
- Alpert, M., Clark, A., Pouget, E.R., 1994. The syntactic role of pauses in the speech of schizophrenic patients with alogia. *J. Abnorm. Psychol.* 103, 750–757. <https://doi.org/http://dx.doi.org/10.1037/0021-843X.103.4.750>
- Alpert, M., Kotsaftis, A., Pouget, E.R., 1997. At issue: Speech fluency and schizophrenic negative signs. *Schizophr. Bull.* 23, 171–177. <https://doi.org/10.1093/schbul/23.2.171>
- Alpert, M., Pouget, E.R., Silva, R., 1995. Cues to the assessment of affects and moods: Speech fluency and pausing. *Psychopharmacol. Bull.* 31, 421–424. <https://doi.org/10.1016/j.joms.2006.09.018>
- Alpert, M., Pouget, E.R., Welkowitz, J., Cohen, J., 1993. Mapping schizophrenic negative symptoms onto measures of the patient's speech: Set correlational analysis. *Psychiatry Res.* 48, 181–190. [https://doi.org/10.1016/0165-1781\(93\)90070-W](https://doi.org/10.1016/0165-1781(93)90070-W)
- Alpert, M., Rosenberg, S.D., Pouget, E.R., Shaw, R.J., 2000. Prosody and lexical accuracy in flat affect schizophrenia. *Psychiatry Res.* 97, 107–118. [https://doi.org/10.1016/S0165-1781\(00\)00231-6](https://doi.org/10.1016/S0165-1781(00)00231-6)
- 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. *J. Psychiatr. Res.* 36, 347–353. [https://doi.org/10.1016/S0022-3956\(02\)00016-X](https://doi.org/10.1016/S0022-3956(02)00016-X)
- Alpert, M., Welkowitz, R.J., Sobin, C., Borod, J.C., Rosen, A., 1989. Vocal acoustic correlates of flat affect in schizophrenia. Similarity to Parkinson's disease and right hemisphere disease and contrast with depression. *Br. J. Psychiatry* 154, 51–56.
- Ambite, J.L., Tallis, M., Alpert, K., Keator, D.B., King, M., Landis, D., Konstantinidis, G., Calhoun, V.D., Potkin, S.G., Turner, J.A., Wang, L., 2015. SchizConnect: Virtual Data integration in neuroimaging, in: *Lecture Notes in Computer Science (Including Subseries*

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).

https://doi.org/10.1007/978-3-319-21843-4_4

Andreasen, N.C., 1984. Scale for the assessment of negative symptom. University of Iowa, Iowa City.

Andreasen, N.C., Hoffman, R., Grove, W., 1985. Mapping abnormalities in Language and cognition, in: Alpert, A. (Ed.), *Controversies in Schizophrenia*. Guilford Press, New York, NY, pp. 199–226.

Bambini, V., Arcara, G., Bechi, M., Buonocore, M., Cavallaro, R., Bosia, M., 2016. The communicative impairment as a core feature of schizophrenia: Frequency of pragmatic deficit, cognitive substrates, and relation with quality of life. *Compr. Psychiatry* 71, 106–120. <https://doi.org/10.1016/j.comppsy.2016.08.012>

Begg, C.B., Mazumdar, M., 1994. Operating Characteristics of a Rank Correlation Test for Publication Bias. *Biometrics* 50, 1088. <https://doi.org/10.2307/2533446>

Ben-Zeev, D., Brian, R., Wang, R., Wang, W., Campbell, A.T., Aung, M.S.H., Merrill, M., Tseng, V.W.S., Choudhury, T., Hauser, M., Kane, J.M., Scherer, E.A., 2017. CrossCheck: Integrating self-report, behavioral sensing, and smartphone use to identify digital indicators of psychotic relapse. *Psychiatr. Rehabil. J.* 40, 266–275. <https://doi.org/10.1037/prj0000243>

Bernardini, F., Lunden, A., Covington, M., Broussard, B., Halpern, B., Alolayan, Y., Crisafio, A., Pauselli, L., Balducci, P.M., Capulong, L., Attademo, L., Lucarini, E., Salierno, G., Natalicchi, L., Quartesan, R., Compton, M.T., 2016. Associations of acoustically measured tongue/jaw movements and portion of time speaking with negative symptom severity in patients with schizophrenia in Italy and the United States. *Psychiatry Res.* 239, 253–258. <https://doi.org/10.1016/j.psychres.2016.03.037>

Berthold, A., Jameson, A., 1999. Interpreting Symptoms of Cognitive Load in Speech Input. Springer, Vienna, pp. 235–244. https://doi.org/10.1007/978-3-7091-2490-1_23

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

Bleuler, E., 1911. *Dementia Praecox or the Group of Schizophrenias*. International University Press, New York, NY.

Boersma, P., Weenink, D., 2018. Praat: doing phonetics by computer [Computer program]. Version 6.0.43 [WWW Document]. retrieved 8 Sept. 2018.

Bone, D., Bishop, S.L., Black, M.P., Goodwin, M.S., Lord, C., Narayanan, S.S., 2016. Use of machine learning to improve autism screening and diagnostic instruments: effectiveness, efficiency, and multi-instrument fusion. *J. Child Psychol. Psychiatry* 57, 927–937. <https://doi.org/10.1111/jcpp.12559>

Bone, D., Lee, C.-C., Chaspari, T., Gibson, J., Narayanan, S., 2017. Signal Processing and Machine Learning for Mental Health Research and Clinical Applications [Perspectives]. *IEEE Signal Process. Mag.* 34, 196–195. <https://doi.org/10.1109/MSP.2017.2718581>

Bonneh, Y.S., Levanon, Y., Dean-Pardo, O., Lossos, L., Adini, Y., 2011. Abnormal Speech Spectrum and Increased Pitch Variability in Young Autistic Children. *Front. Hum. Neurosci.* 4, 1–7. <https://doi.org/10.3389/fnhum.2010.00237>

Borsboom, D., 2017. A network theory of mental disorders. *World Psychiatry* 16, 5–13. <https://doi.org/10.1002/wps.20375>

Bosco, F., Berardinelli, L., Parola, A., 2019. The ability of patients with schizophrenia to comprehend and produce sincere, deceitful and ironic communicative intentions: The role of theory of mind and executive functions. *Front. Psychol.* <https://doi.org/10.3389/fpsyg.2019.00827>

Bowie, C.R., Harvey, P.D., 2008. Communication abnormalities predict functional outcomes in chronic schizophrenia: Differential associations with social and adaptive functions. *Schizophr. Res.* 103, 240–247. <https://doi.org/10.1016/j.schres.2008.05.006>

Brüne, M., Bodenstein, L., 2005. Proverb comprehension reconsidered—‘theory of mind’ and the pragmatic use of language in schizophrenia. *Schizophr. Res.* 75, 233–239. <https://doi.org/10.1016/j.schres.2004.11.006>

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

Bürkner, P.C., Williams, D.R., Simmons, T.C., Woolley, J.D., Cohen, A.S., Mitchell, K.R.,

Elvevåg, B., Fusaroli, R., Lambrechts, A., Bang, D., Bowler, D.M., Gaigg, S.B., CNX

Anatomy and Physiology, O.C., Pickering, M.J., Garrod, S., Quintana, D.S., 2017.

Intranasal Oxytocin May Improve High-Level Social Cognition in Schizophrenia, but

Not Social Cognition or Neurocognition in General: A Multilevel Bayesian Meta-

analysis. *Schizophr. Bull.* 43, 1291–1303. <https://doi.org/10.1093/schbul/sbx053>

Bürkner PC, 2013. brms : An R Package for Bayesian Multilevel Models using Stan. *J. Stat.*

Softw. 80, 1–28. <https://doi.org/10.18637/jss.v080.i01>

Cannizzaro, M.S., Cohen, H., Rappard, F., Snyder, P.J., 2005. Bradyphrenia and bradykinesia

both contribute to altered speech in schizophrenia: A quantitative acoustic study. *Cogn.*

Behav. Neurol. 18, 206–210. <https://doi.org/10.1097/01.wnn.0000185278.21352.e5>

Carpenter, B., Guo, J., Hoffman, M.D., Brubaker, M., Gelman, A., Lee, D., Goodrich, B., Li,

P., Riddell, A., Betancourt, M., 2017. Stan : A Probabilistic Programming Language. *J.*

Stat. Softw. 76. <https://doi.org/10.18637/jss.v076.i01>

Champagne-Lavau, M., Stip, E., 2010. Pragmatic and executive dysfunction in schizophrenia.

J. Neurolinguistics 23, 285–296. <https://doi.org/10.1016/j.jneuroling.2009.08.009>

Chekroud, A., 2018. T107. Why Validation Matters: A Demonstration Predicting

Antipsychotic Response Using 5 Rcts. *Schizophr. Bull.* 44(Suppl 1).

Clemmer, E.J., 1980. Psycholinguistic aspects of pauses and temporal patterns in

schizophrenic speech. *J. Psycholinguist. Res.* 9, 161–185.

<https://doi.org/10.1007/BF01067469>

Cochran, W.G., 1954. The Combination of Estimates from Different Experiments. *Biometrics*

10, 101. <https://doi.org/10.2307/3001666>

Cohen, A.S., Alpert, M., Nienow, T.M., Dinzeo, T.J., Docherty, N.M., 2008. Computerized

measurement of negative symptoms in schizophrenia. *J. Psychiatr. Res.* 42, 827–836.

<https://doi.org/10.1016/j.jpsychires.2007.08.008>

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

- 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 Res.* 226, 235–241.
<https://doi.org/10.1016/j.psychres.2014.12.054>
- Cohen, A.S., Elvevåg, B., 2014. Automated computerized analysis of speech in psychiatric disorders. *Curr. Opin. Psychiatry.* <https://doi.org/10.1097/YCO.0000000000000056>
- Cohen, A.S., Hong, S.L., 2011. Understanding Constricted Affect in Schizotypy Through Computerized Prosodic Analysis. *J. Pers. Disord.* 25, 478–491.
<https://doi.org/10.1521/pedi.2011.25.4.478>
- Cohen, A.S., Kim, Y., Najolia, G.M., 2013. Psychiatric symptom versus neurocognitive correlates of diminished expressivity in schizophrenia and mood disorders. *Schizophr. Res.* 146, 249–253. <https://doi.org/10.1016/j.schres.2013.02.002>
- Cohen, A.S., Mitchell, K.R., Docherty, N.M., Horan, W.P., 2016. Vocal expression in schizophrenia: Less than meets the ear. *J. Abnorm. Psychol.* 125, 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. *Schizophr. Res.* 159, 533–538. <https://doi.org/10.1016/j.schres.2014.09.013>
- Cohen, A.S., Mitchell, K.R., Strauss, G.P., Blanchard, J.J., Buchanan, R.W., Kelly, D.L., Gold, J., McMahon, R.P., Adams, H.A., Carpenter, W.T., 2017. The effects of oxytocin and galantamine on objectively-defined vocal and facial expression: Data from the CIDAR study. *Schizophr. Res.* 188, 141–143.
<https://doi.org/10.1016/j.schres.2017.01.028>
- Cohen, A.S., Morrison, S.C., Brown, L.A., Minor, K.S., 2012a. Towards a cognitive resource limitations model of diminished expression in schizotypy. *J. Abnorm. Psychol.* 121, 109–118. <https://doi.org/10.1037/a0023599>

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

- Cohen, A.S., Najolia, G.M., Kim, Y., Dinzeo, T.J., 2012b. On the boundaries of blunt affect/alogia across severe mental illness: Implications for Research Domain Criteria. *Schizophr. Res.* 140, 41–45. <https://doi.org/10.1016/j.schres.2012.07.001>
- 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, 1–14. <https://doi.org/10.1371/journal.pone.0217404>
- Colle, L., Angeleri, R., Vallana, M., Sacco, K., Bara, B.G., Bosco, F.M., 2013. Understanding the communicative impairments in schizophrenia: A preliminary study. *J. Commun. Disord.* 46, 294–308. <https://doi.org/10.1016/j.jcomdis.2013.01.003>
- Compton, M.T., Lunden, A., Cleary, S.D., Pauselli, L., Alolayan, Y., Halpern, B., Broussard, B., Crisafio, A., Capulong, L., Balducci, P.M., Bernardini, F., Covington, M.A., 2018. The aprosody of schizophrenia: Computationally derived acoustic phonetic underpinnings of monotone speech. *Schizophr. Res.* 197, 392–399. <https://doi.org/10.1016/j.schres.2018.01.007>
- Corcoran, C.M., Carrillo, F., Fernández-Slezak, D., Bedi, G., Klim, C., Javitt, D.C., Bearden, C.E., Cecchi, G.A., 2018. Prediction of psychosis across protocols and risk cohorts using automated language analysis. *World Psychiatry* 17, 67–75. <https://doi.org/10.1002/wps.20491>
- Core R Team, 2013. R: A language and environment for statistical computing. R Found. Stat. Comput. Vienna, Austria.
- Couture, S.M., Granholm, E.L., Fish, S.C., 2011. A path model investigation of neurocognition, theory of mind, social competence, negative symptoms and real-world functioning in schizophrenia. *Schizophr. Res.* 125, 152–160. <https://doi.org/10.1016/j.schres.2010.09.020>
- Covington, M.A., Lunden, S.L.A., Cristofaro, S.L., Wan, C.R., Bailey, C.T., Broussard, B.,

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

- Fogarty, R., Johnson, S., Zhang, S., Compton, M.T., 2012. Phonetic measures of reduced tongue movement correlate with negative symptom severity in hospitalized patients with first-episode schizophrenia-spectrum disorders. *Schizophr. Res.* 142, 93–95.
<https://doi.org/10.1016/j.schres.2012.10.005>
- Cummins, N., Scherer, S., Krajewski, J., Schnieder, S., Epps, J., Quatieri, T.F., 2015a. A review of depression and suicide risk assessment using speech analysis. *Speech Commun.* 71, 10–49. <https://doi.org/10.1016/J.SPECOM.2015.03.004>
- Cummins, N., Sethu, V., Epps, J., Krajewski, J., 2014. Probabilistic acoustic volume analysis for speech affected by depression, in: *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*.
- Cummins, N., Sethu, V., Epps, J., Schnieder, S., Krajewski, J., 2015b. Analysis of acoustic space variability in speech affected by depression. *Speech Commun.* 75, 27–49.
<https://doi.org/10.1016/j.specom.2015.09.003>
- Dahlgren, S., Sandberg, A.D., Strömbergsson, S., Wenhov, L., Råstam, M., Nettelbladt, U., 2018. Prosodic traits in speech produced by children with autism spectrum disorders – Perceptual and acoustic measurements. *Autism Dev. Lang. Impair.* 3, 239694151876452.
<https://doi.org/10.1177/2396941518764527>
- Degottex, G., Kane, J., Drugman, T., Raitio, T., Scherer, S., 2014. COVAREP - A collaborative voice analysis repository for speech technologies, in: *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*. pp. 960–964. <https://doi.org/10.1109/ICASSP.2014.6853739>
- Del-Monte, J., Raffard, S., Salesse, R.N., Marin, L., Schmidt, R.C., Varlet, M., Bardy, B.G., Philippe Boulenger, J., Christine Gély-Nargeot, M., Capdevielle, D., 2013. Nonverbal expressive behaviour in schizophrenia and social phobia. *Psychiatry Res.* 210, 29–35.
<https://doi.org/10.1016/j.psychres.2013.05.034>
- Docherty, N.M., 2012. Missing referents, psychotic symptoms, and discriminating the internal

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

from the externalized. *J. Abnorm. Psychol.* 121, 416–423.

<https://doi.org/10.1037/a0026348>

Doebler, P., Holling, H., 2015. Meta-Analysis of Diagnostic Accuracy with mada. *R Packag.* 1–15.

Emmerson, L.C., Ben-zeev, D., Granholm, E., Tiffany, M., Golshan, S., Jeste, D. V., 2009.

Prevalence and longitudinal stability of negative symptoms in healthy participants 1438–1444. <https://doi.org/10.1002/gps>

Eyben, F., Wöllmer, M., Schuller, B., 2010. Opensmile: the munich versatile and fast open-source audio feature extractor. *Proc. ACM Multimed.* 1459–1462.

<https://doi.org/10.1145/1873951.1874246>

Faul, F., Erdfelder, E., Lang, A., Buchner, A., 2009. Statistical power analyses using

G*Power 3.1: Tests for correlation and regression analyses. *Behav. Res. Methods* 41, 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>

Faurholt-Jepsen, M., Busk, J., Frost, M., Vinberg, M., Christensen, E.M., Winther, O.,

Bardram, J.E., Kessing, L. V., 2016. Voice analysis as an objective state marker in bipolar disorder. *Transl. Psychiatry* 6, e856. <https://doi.org/10.1038/tp.2016.123>

Field, A.P., Gillett, R., 2010. How to do a meta-analysis. *Br. J. Math. Stat. Psychol.* 63, 665–694.

Foody, G., 2017. Impacts of Sample Design for Validation Data on the Accuracy of

Feedforward Neural Network Classification. *Appl. Sci.* 7, 888.

<https://doi.org/10.3390/app7090888>

Frith, C.D., 1992. *The Cognitive Neuropsychology of Schizophrenia*. Lawrence Erlbaum, Hove.

Fusaroli, R., Lambrechts, A., Bang, D., Bowler, D.M., Gaigg, S.B., 2017. “Is voice a marker for Autism spectrum disorder? A systematic review and meta-analysis.” *Autism Res.* 10, 384–407. <https://doi.org/10.1002/aur.1678>

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

- Fusaroli, Riccardo, Lambrechts, A., Bang, D., Bowler, D.M., Gaigg, S.B., 2017. “Is voice a marker for Autism spectrum disorder? A systematic review and meta-analysis.” *Autism Res.* 10, 384–407. <https://doi.org/10.1002/aur.1678>
- Galynker, I., Cohen, L., Cai, J., 2000. Negative Symptoms in Patients With Major Depressive Disorder: A Preliminary Report. *Neuropsychiatry Neuropsychol Behav Neurol* 13, 171–176.
- Glaister, J., Feldstein, S., Pollack, H., 1980. Chronographic speech patterns of acutely psychotic patients: A preliminary note. *J. Nerv. Ment. Dis.* 168, 219–223. <https://doi.org/10.1097/00005053-198004000-00005>
- Gratch, J., Artstein, R., Lucas, G., Stratou, G., Scherer, S., Nazarian, A., Wood, R., Boberg, J., Devault, D., Marsella, S., Traum, D., Rizzo, S., Morency, L.-P., 2014. The Distress Analysis Interview Corpus of human and computer interviews, in: *LREC*. pp. 3123–3128.
- Graux, J., Courtine, J.B., Bruneau, N., Camus, V., El-Hage, W., 2015. Higher fundamental voice frequency is related to extrapyramidal symptoms in schizophrenia. *Schizophr. Res.* 161, 517–518. <https://doi.org/10.1016/j.schres.2014.10.045>
- Green, M.F., Horan, W.P., Lee, J., 2015. Social cognition in schizophrenia. *Nat. Rev. Neurosci.* <https://doi.org/10.1038/nrn4005>
- Häfner, H., Löffler, W., Maurer, K., Hambrecht, M., An Der Heiden, W., 1999. Depression, negative symptoms, social stagnation and social decline in the early course of schizophrenia. *Acta Psychiatr. Scand.* 100, 105–118. <https://doi.org/10.1111/j.1600-0447.1999.tb10831.x>
- Hoekert, M., Kahn, R.S., Pijnenborg, M., Aleman, A., 2007. Impaired recognition and expression of emotional prosody in schizophrenia: Review and meta-analysis. *Schizophr. Res.* 96, 135–145. <https://doi.org/10.1016/j.schres.2007.07.023>
- Huys, Q.J.M., Maia, T. V., Frank, M.J., 2016. Computational psychiatry as a bridge from

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

- neuroscience to clinical applications. *Nat. Neurosci.* <https://doi.org/10.1038/nn.4238>
- James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An Introduction to Statistical Learning. <https://doi.org/10.1007/978-1-4614-7138-7>
- Khawaja, M.A., Ruiz, N., Chen, F., 2008. Think before you talk, in: Proceedings of the 20th Australasian Conference on Computer-Human Interaction Designing for Habitus and Habitat - OZCHI '08. ACM Press, New York, New York, USA, p. 335. <https://doi.org/10.1145/1517744.1517814>
- Kiss, G., van Santen, J.P.H., Prud'Hommeaux, E.T., Black, L.M., 2012. Quantitative Analysis of Pitch in Speech of Children with Neurodevelopmental Disorders. *Interspeech* 1343–1346.
- Klein, R.A., Vianello, M., Hasselman, F., Adams, B.G., Adams Jr, R.B., Alper, S., ... &, Batra, R., 2018. Many Labs 2: Investigating variation in replicability across samples and settings. *Adv. Methods Pract. Psychol. Sci.* 1, 443–490.
- Klipper, R., Portuguese, S., Weinshall, D., 2016. Prosodic analysis of speech and the underlying mental state. *Commun. Comput. Inf. Sci.* 604, 52–62. https://doi.org/10.1007/978-3-319-32270-4_6
- Klipper, R., Vaizman, Y., Weinshall, D., Portuguese, S., 2010. Evidence for Depression and Schizophrenia in speech Prosody. *Isca* 25–27.
- Konopka, G., Roberts, T.F., 2016. Insights into the Neural and Genetic Basis of Vocal Communication. *Cell* 164, 1269–1276. <https://doi.org/10.1016/J.CELL.2016.02.039>
- Kraepelin, E., 1919. Dementia Praecox and Paraphrenia. University of Edinburgh, Edinburgh, UK.
- Kring, A.M., Alpert, M., Neale, J.M., Harvey, P.D., 1994. A multimethod, multichannel assessment of affective flattening in schizophrenia. *Psychiatry Res.* 54, 211–222. [https://doi.org/10.1016/0165-1781\(94\)90008-6](https://doi.org/10.1016/0165-1781(94)90008-6)
- Leudar, I., Thomas, P., Johnston, M., 1994. Self-monitoring in speech production: effects of

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

verbal hallucinations and negative symptoms. *Psychol. Med.* 24, 749–761.

<https://doi.org/10.1017/S0033291700027902>

Lysaker, P., Bell, M., 1995. Work and Meaning: Disturbance of Volition and Vocational

Dysfunction in Schizophrenia. *Psychiatry* 58, 392–400.

<https://doi.org/10.1080/00332747.1995.11024743>

Mandal, M.K., Srivastava, P., Singh, S.K., 1990. Paralinguistic characteristics of speech in

schizophrenics and depressives. *J. Psychiatr. Res.* 24, 191–196.

[https://doi.org/10.1016/0022-3956\(90\)90059-Y](https://doi.org/10.1016/0022-3956(90)90059-Y)

Martínez-Sánchez, F., Muela-Martínez, J.A., Cortés-Soto, P., García Meilán, J.J., Vera

Ferrándiz, J.A., Egea Caparrós, A., Pujante Valverde, I.M., 2015. Can the Acoustic

Analysis of Expressive Prosody Discriminate Schizophrenia? *Span. J. Psychol.* 18, E86.

<https://doi.org/10.1017/sjp.2015.85>

Martínez, A., Donado, F., Alberto, C., Eslava, M., Diana, L., 2015. Valoración de prosodia

espontánea afectiva y análisis de discurso en pacientes con esquizofrenia y demencia

frontotemporal (DFT) variante lingüística. *Rev. Colomb. Psiquiatr.* 44, 13–19.

<https://doi.org/10.1016/j.rcp.2015.01.003>

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. *Front. Psychiatry* 4, 1–9. <https://doi.org/10.3389/fpsy.2013.00127>

McGilloway, S., Cooper, S.J., Douglas-Cowie, E., 2003. Can patients with chronic

schizophrenia express emotion? A speech analysis. *Schizophr. Res.* 64, 189–190.

[https://doi.org/10.1016/S0920-9964\(03\)00012-4](https://doi.org/10.1016/S0920-9964(03)00012-4)

Meaux, L.T., Mitchell, K.R., Cohen, A.S., 2018. Blunted vocal affect and expression is not

associated with schizophrenia: A computerized acoustic analysis of speech under

ambiguous conditions. *Compr. Psychiatry* 83, 84–88.

<https://doi.org/10.1016/j.comppsy.2018.03.009>

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

- Meilijson, S.R., Kasher, A., Elizur, A., 2004. Language performance in chronic schizophrenia: a pragmatic approach. *J. Speech. Lang. Hear. Res.* 47, 695–713.
[https://doi.org/10.1044/1092-4388\(2004/053\)](https://doi.org/10.1044/1092-4388(2004/053))
- Michelas, A., Faget, C., Portes, C., Lienhart, A.S., Boyer, L., Lançon, C., Champagne-Lavau, M., 2014. Do patients with schizophrenia use prosody to encode contrastive discourse status? *Front. Psychol.* 5, 1–13. <https://doi.org/10.3389/fpsyg.2014.00755>
- Millan, M.J., Fone, K., Steckler, T., Horan, W.P., 2014. Negative symptoms of schizophrenia: Clinical characteristics, pathophysiological substrates, experimental models and prospects for improved treatment. *Eur. Neuropsychopharmacol.* 24, 645–692.
<https://doi.org/10.1016/j.euroneuro.2014.03.008>
- Mueser, K.T., Sayers, S.L., Schooler, N.R., Mance, R.M., Haas, G.L., 1994. A multisite investigation of the reliability of the scale for the assessment of negative symptoms. *Am. J. Psychiatry* 151, 1453–1462. <https://doi.org/10.1176/ajp.151.10.1453>
- Nevler, N., Ash, S., Jester, C., Irwin, D.J., Liberman, M., Grossman, M., 2017. Automatic measurement of prosody in behavioral variant FTD. *Neurology* 89, 650–656.
<https://doi.org/10.1212/WNL.0000000000004236>
- Parola, A., Berardinelli, L., Bosco, F.M., 2018. Cognitive abilities and theory of mind in explaining communicative-pragmatic disorders in patients with schizophrenia. *Psychiatry Res.* 260, 144–151. <https://doi.org/10.1016/j.psychres.2017.11.051>
- Peluso, M.J., Lewis, S.W., Barnes, T.R.E., Jones, P.B., 2012. Extrapyramidal motor side-effects of first- and second-generation antipsychotic drugs. *Br. J. Psychiatry* 200, 387–392. <https://doi.org/10.1192/bjp.bp.111.101485>
- Penn, D.L., Sanna, L.J., Roberts, D.L., 2008. Social cognition in schizophrenia: an overview. *Schizophr. Bull.* 34, 408–11. <https://doi.org/10.1093/schbul/sbn014>
- Perlini, C., Marini, A., Garzitto, M., Isola, M., Cerruti, S., Marinelli, V., Rambaldelli, G., Ferro, A., Tomelleri, L., Dusi, N., Bellani, M., Tansella, M., Fabbro, F., Brambilla, P.,

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

2012. Linguistic production and syntactic comprehension in schizophrenia and bipolar disorder. *Acta Psychiatr. Scand.* 126, 363–376. <https://doi.org/10.1111/j.1600-0447.2012.01864.x>
- Pinheiro, A.P., Rezaii, N., Rauber, A., Nestor, P.G., Spencer, K.M., Niznikiewicz, M., 2017. Emotional self–other voice processing in schizophrenia and its relationship with hallucinations: ERP evidence. *Psychophysiology* 54, 1252–1265. <https://doi.org/10.1111/psyp.12880>
- Pinheiro, A.P., Rezaii, N., Rauber, A., Niznikiewicz, M., 2016. Is this my voice or yours? The role of emotion and acoustic quality in self-other voice discrimination in schizophrenia. *Cogn. Neuropsychiatry* 21, 335–353. <https://doi.org/10.1080/13546805.2016.1208611>
- Pisanski, K., Cartei, V., McGettigan, C., Raine, J., Reby, D., 2016. Voice Modulation: A Window into the Origins of Human Vocal Control? *Trends Cogn. Sci.* 20, 304–318. <https://doi.org/10.1016/j.tics.2016.01.002>
- Püschel, J., Stassen, H.H., Bomben, G., Scharfetter, C., Hell, D., 1998. Speaking behavior and speech sound characteristics in acute schizophrenia. *J. Psychiatr. Res.* 32, 89–97. [https://doi.org/10.1016/S0022-3956\(98\)00046-6](https://doi.org/10.1016/S0022-3956(98)00046-6)
- Quintana, D.S., 2015. From pre-registration to publication: A non-technical primer for conducting a meta-analysis to synthesize correlational data. *Front. Psychol.* 6, 1–9. <https://doi.org/10.3389/fpsyg.2015.01549>
- Rabinowitz, J., Levine, S.Z., Garibaldi, G., Bugarski-Kirola, D., Berardo, C.G., Kapur, S., 2012. Negative symptoms have greater impact on functioning than positive symptoms in schizophrenia: Analysis of CATIE data. *Schizophr. Res.* 137, 147–150. <https://doi.org/10.1016/j.schres.2012.01.015>
- Rapcan, V., D’Arcy, S., Yeap, S., Afzal, N., Thakore, J., Reilly, R.B., 2010. Acoustic and temporal analysis of speech: A potential biomarker for schizophrenia. *Med. Eng. Phys.* 32, 1074–1079. <https://doi.org/10.1016/j.medengphy.2010.07.013>

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

- Resnick, H.S., Oltmanns, T.F., 1984. Hesitation patterns in the speech of thought-disordered schizophrenic and manic patients. *J. Abnorm. Psychol.* 93, 80.
<https://doi.org/10.1037/0021-843X.93.1.80>
- Rochester, S.R., Martin, J.R., Thurston, S., 1977. Thought-process disorder in schizophrenia: The listener's task. *Brain Lang.* 4, 95–114. [https://doi.org/10.1016/0093-934X\(77\)90009-8](https://doi.org/10.1016/0093-934X(77)90009-8)
- Ross, E D, Orbelo, D.M., Cartwright, J., Hansel, S., Burgard, M., Testa, J.A., Buck, R., 2001. Affective-prosodic deficits in schizophrenia: comparison to patients with brain damage and relation to schizophrenic symptoms. *J. Neurol. Neurosurg. Psychiatry.* 70, 597–604.
- Ross, E. D., Orbelo, D.M., Testa, J.A., Hansel, S., Burgard, M., Cartwright, J., Buck, R., 2001. Affective-prosodic deficits in schizophrenia: Profiles of patients with brain damage and comparison with relation to schizophrenic symptoms. *J. Neurol. Neurosurg. Psychiatry* 70, 597–604. <https://doi.org/10.1136/jnnp.70.5.597>
- Rutter, D.R., 1977. Speech patterning in recently admitted and chronic long stay schizophrenic patients. *Br. J. Soc. Clin. Psychol.* 16, 47–55.
- Salomé, F., Boyer, P., Fayol, M., 2002. Written but not oral verbal production is preserved in young schizophrenic patients. *Psychiatry Res.* 111, 137–145.
[https://doi.org/10.1016/S0165-1781\(02\)00142-7](https://doi.org/10.1016/S0165-1781(02)00142-7)
- Schönbrodt, F.D., Perugini, M., 2013. At what sample size do correlations stabilize? *J. Res. Pers.* 47, 609–612. <https://doi.org/10.1016/j.jrp.2013.05.009>
- Schuller, B., Steidl, S., Batliner, A., Vinciarelli, A., Scherer, K., Ringeval, F., Chetouani, M., Weninger, F., Eyben, F., Marchi, E., Mortillaro, M., Salamin, H., Polychroniou, A., Valente, F., Kim, S., 2013. The INTERSPEECH 2013 computational paralinguistics challenge: Social signals, conflict, emotion, autism, in: *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH.*
- Sergi, M.J., Rassovsky, Y., Widmark, C., Reist, C., Erhart, S., Braff, D.L., Marder, S.R.,

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

- Green, M.F., 2007. Social cognition in schizophrenia: Relationships with neurocognition and negative symptoms. *Schizophr. Res.* <https://doi.org/10.1016/j.schres.2006.09.028>
- Shaw, R.J., Dong, M., Lim, K.O., Faustman, W.O., Pouget, E.R., Alpert, M., 1999. The relationship between affect expression and affect recognition in schizophrenia. *Schizophr. Res.* 37, 245–250. [https://doi.org/10.1016/S0920-9964\(98\)00172-8](https://doi.org/10.1016/S0920-9964(98)00172-8)
- Sison, C.E., Alpert, M., Fudge, R., Stern, R.M., 1996. Constricted expressiveness and psychophysiological reactivity in schizophrenia. *J. Nerv. Ment. Dis.* 184, 589–597. <https://doi.org/10.1097/00005053-199610000-00002>
- Sparks, A., McDonald, S., Lino, B., O'Donnell, M., Green, M.J., 2010. Social cognition, empathy and functional outcome in schizophrenia. *Schizophr. Res.* 122, 172–178. <https://doi.org/10.1016/j.schres.2010.06.011>
- St-Hilaire, A., Cohen, A.S., Docherty, N.M., 2008. Emotion word use in the conversational speech of schizophrenia patients. *Cogn. Neuropsychiatry* 13, 343–356. <https://doi.org/10.1080/13546800802250560>
- Stassen, H.H., Albers, M., Tewesmeier, M., Woggon, B., Püschel, J., Scharfetter, C., Tewesmeier, M., Woggon, B., 1995. Speaking behavior and voice sound characteristics associated with negative schizophrenia. *J. Psychiatr. Res.* 29, 277–96. [https://doi.org/10.1016/0022-3956\(95\)00004-o](https://doi.org/10.1016/0022-3956(95)00004-o)
- Stewart, L.A., Clarke, M., Rovers, M., Riley, R.D., Simmonds, M., Stewart, G., Tierney, J.F., 2015. Preferred reporting items for a systematic review and meta-analysis of individual participant data: The PRISMA-IPD statement. *JAMA - J. Am. Med. Assoc.* 313, 1657–1665. <https://doi.org/10.1001/jama.2015.3656>
- Tahir, Y., Yang, Z., Id, D.C., Thalmann, N., Thalmann, D., Maniam, Y., Amirah, N., Tan, L., Lee, J., Keong, C., Dauwels, J., 2019. Non-verbal speech cues as objective measures for negative symptoms in patients with schizophrenia 1–17.
- Tan, E.J., Thomas, N., Rossell, S.L., 2014. Speech disturbances and quality of life in

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

- schizophrenia: Differential impacts on functioning and life satisfaction. *Compr. Psychiatry* 55, 693–698. <https://doi.org/10.1016/j.comppsy.2013.10.016>
- Tandon, R., Keshavan, M.S., Nasrallah, H. a, 2008. Schizophrenia, “just the facts” what we know in 2008. 2. Epidemiology and etiology. *Schizophr. Res.* 102, 1–18. <https://doi.org/10.1016/j.schres.2008.04.011>
- Tavano, A., Sponda, S., Fabbro, F., Perlini, C., Rambaldelli, G., Ferro, A., Cerruti, S., Tansella, M., Brambilla, P., 2008. Specific linguistic and pragmatic deficits in Italian patients with schizophrenia. *Schizophr. Res.* 102, 53–62. <https://doi.org/10.1016/j.schres.2008.02.008>
- Tenback, D., van Harten, P., Slooff, C., van Os, J., 2010. Incidence and persistence of tardive dyskinesia and extrapyramidal symptoms in schizophrenia. *J. Psychopharmacol.* 24, 1031–1035. <https://doi.org/10.1177/0269881109106306>
- Thoma, P., Hennecke, M., Mandok, T., Wähner, A., Brüne, M., Juckel, G., Daum, I., 2009. Proverb comprehension impairments in schizophrenia are related to executive dysfunction. *Psychiatry Res.* 170, 132–139. <https://doi.org/10.1016/j.psychres.2009.01.026>
- Todder, D., Avissar, S., Schreiber, G., 2013. Non-Linear Dynamic Analysis of Inter-Word Time Intervals in Psychotic Speech. *IEEE J. Transl. Eng. Heal. Med.* 1, 2200107–2200107. <https://doi.org/10.1109/JTEHM.2013.2268850>
- Tolkmitt, F., Helfrich, H., Standke, R., Scherer, K.R., 1982. Vocal indicators of psychiatric treatment effects in depressives and schizophrenics. *J. Commun. Disord.* 15, 209–222. [https://doi.org/10.1016/0021-9924\(82\)90034-X](https://doi.org/10.1016/0021-9924(82)90034-X)
- Trémeau, F., Goldman, J., Antonius, D., Javitt, D.C., 2013. Inpatients with schizophrenia report impaired situational motivation but intact global and social motivation. *Psychiatry Res.* 210, 43–49. <https://doi.org/10.1016/j.psychres.2013.05.031>
- Trémeau, F., Malaspina, D., Duval, F., Corrêa, H., Hager-Budny, M., Coin-Bariou, L.,

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

- Macher, J.P., Gorman, J.M., 2005. Facial expressiveness in patients with schizophrenia compared to depressed patients and nonpatient comparison subjects. *Am. J. Psychiatry* 162, 92–101. <https://doi.org/10.1176/appi.ajp.162.1.92>
- Trigeorgis, G., Ringeval, F., Brueckner, R., Marchi, E., Nicolaou, M.A., Schuller, B., Zafeiriou, S., 2016. Adieu features? End-to-end speech emotion recognition using a deep convolutional recurrent network. 2016 IEEE Int. Conf. Acoust. Speech Signal Process. <https://doi.org/10.1109/ICASSP.2016.7472669>
- Tsanas, A., Little, M.A., McSharry, P.E., Ramig, L.O., 2011. Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson’s disease symptom severity. *J. R. Soc. Interface* 8, 842–855. <https://doi.org/10.1098/rsif.2010.0456>
- Tsanas, A., Zañartu, M., Little, M.A., Fox, C., Ramig, L.O., Clifford, G.D., 2014. Robust fundamental frequency estimation in sustained vowels: Detailed algorithmic comparisons and information fusion with adaptive Kalman filtering. *J. Acoust. Soc. Am.* 135, 2885–2901. <https://doi.org/10.1121/1.4870484>
- Van Der Ploeg, T., Austin, P.C., Steyerberg, E.W., 2014. Modern modelling techniques are data hungry: A simulation study for predicting dichotomous endpoints. *BMC Med. Res. Methodol.* 14, 1–13. <https://doi.org/10.1186/1471-2288-14-137>
- Vehtari, A., Gelman, A., Gabry, J., 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat. Comput.* 27, 1413–1432. <https://doi.org/10.1007/s11222-016-9696-4>
- Viechtbauer, W., 2010. Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* 36, 1–18.
- Walther, S., 2015. Psychomotor symptoms of schizophrenia map on the cerebral motor circuit. *Psychiatry Res. Neuroimaging* 233, 293–298. <https://doi.org/10.1016/J.PSCYCHRESNS.2015.06.010>

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

- Walther, S., Strik, W., 2012. Motor Symptoms and Schizophrenia. *Neuropsychobiology* 66, 77–92. <https://doi.org/10.1159/000339456>
- Wan, M.W., Penketh, V., Salmon, M.P., Abel, K.M., 2008. Content and style of speech from mothers with schizophrenia towards their infants. *Psychiatry Res.* 159, 109–114. <https://doi.org/10.1016/j.psychres.2007.05.012>
- Wang, K., An, N., Li, B.N., Zhang, Y., Li, L., 2015. Speech Emotion Recognition Using Fourier Parameters. *IEEE Trans. Affect. Comput.* <https://doi.org/10.1109/TAFFC.2015.2392101>
- Weed, E., Fusaroli, R., 2019. Acoustic Measures of Prosody in Right-Hemisphere Damage: A Systematic Review and Meta-Analysis. *bioRxiv* 676734. <https://doi.org/10.1101/676734>
- Williams, D.R., Rast, P., Bürkner, P.C., 2018. Bayesian meta-analysis with weakly informative prior distributions. *PsyArXiv*. <https://doi.org/10.17605/OSF.IO/7TBRM>
- Williams, M.E., Lee, S., Bone, D., Levitt, P., Black, M.P., Lee, C.-C., Narayanan, S., 2014. The Psychologist as an Interlocutor in Autism Spectrum Disorder Assessment: Insights From a Study of Spontaneous Prosody. *J. Speech, Lang. Hear. Res.* 57, 1162–1177. https://doi.org/10.1044/2014_jslhr-s-13-0062
- Yin, B., Ruiz, N., Chen, F., Khawaja, M.A., 2007. Automatic cognitive load detection from speech features. 19th Australas. Conf. Comput. Interact. Entertain. User Interfaces 249–255. <https://doi.org/10.1145/1324892.1324946>
- Zhang, J., Pan, Z., Gui, C., Zhu, J., Cui, D., 2016. Clinical investigation of speech signal features among patients with schizophrenia. *Shanghai Arch. Psychiatry* 28, 95–102. <https://doi.org/10.11919/j.issn.1002-0829.216025>

VOICE IN SCHIZOPHRENIA: REVIEW AND META-ANALYSIS

Figures Legend

Figure 1. Flow chart showing the literature search and study selection process in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines.

Figure 2. Forest plots of effect sizes (Hedges'g) for all the acoustic measures. The x-axis report effect sizes (black dot, positive values indicate that individuals with schizophrenia are higher on that acoustic measures, while negative values the opposite), posterior distribution (density plot) and original data point (white dot) for each study. The y-axis indicates the studies for which statistical estimates have been provided.

The dotted vertical line indicates the null hypothesis (no difference between the populations). The studies are grouped by the speech task used to collect voice recordings (Constr = constrained monological, Free = free monological, Social = social interaction). When adding speech task credibly improved the model, we reported below each specific task group the summary effect size for that group. Filled diamonds represent summary effect sizes. Note that multiple inputs for the same reference may refer to original studies including: 1) more than one sample; 2) more than one speech production tasks (e.g. free monologue and dialogue); 3) more than one symptoms rating scale to measure symptomatology.

Figure 3. Forest plots of effect size (Pearson's r) for the correlations between clinical symptoms and acoustic measures. The x-axis report effect sizes (black dot, positive values indicate a positive relation between acoustic measures and clinical symptoms rating, e.g. increased pause duration associated with increased rating of alogia, while negative values the opposite), posterior distribution (density plot) and original data point (white dot) for each study. The y-axis indicates the studies for which statistical estimates have been provided.

The dotted vertical line indicates the null hypothesis (no difference between the populations). The studies are grouped as indicated in Figure 2