

Journal of Psychopharmacology
2025, Vol. 39(9) 940–949
© The Author(s) 2025



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/02698811251319455
journals.sagepub.com/home/jop



Quantitative natural language processing markers of psychoactive drug effects: A pre-registered systematic review

Sachin Ahuja¹, Farida Zaher^{2,3}
and Lena Palaniyappan^{2,3,4} 

Abstract

Psychoactive substances used for recreational purposes have mind-altering effects, but systematic evaluation of these effects is largely limited to self-reports. Automated analysis of expressed language (speech and written text) using natural language processing (NLP) tools can provide objective readouts of mental states. In this pre-registered systematic review, we investigate findings from applying the emerging field of computational linguistics to substance use with specific focus on identifying short-term effects of psychoactive drugs. From the literature identified to date, we note that all the studied drugs – stimulants, 3,4-methylenedioxyamphetamine (MDMA), cannabis, ketamine and psychedelics – affect language production. Based on two or more studies per substance, we note some emerging patterns: stimulants increase verbosity; lysergic acid diethylamide reduces the lexicon; MDMA increases semantic proximity to emotional words; psilocybin increases positive sentiment and cannabis affects speech stream acoustics. Ketamine and other drugs are understudied regarding NLP features (one or no studies). One study provided externally validated support for NLP and machine learning-based identification of MDMA intoxication. We could not undertake a meta-analysis due to the high degree of heterogeneity among outcome measures and the lack of sufficient number of studies. We identify a need for harmonised speech tasks to improve replicability and comparability, standardisation of methods for curating and analysing speech and text data, theory-driven inquiries and the need for developing a shared ‘substance use language corpus’ for data mining. The growing field of computational linguistics can be utilized to advance human behavioral pharmacology of psychoactive substances. Achieving this will require concerted efforts towards consistency in research methods.

Keywords

Acoustics, semantics, computational linguistics, intoxication, speech markers

Background

Psychoactive drugs (PADs) including psychedelics/hallucinogens, entactogens, psychostimulants and cannabinoids alter our mental representations and cognition. PADs in these classes are commonly used recreationally and can cause effects that mimic other altered mental states such as psychosis (Brady et al., 1991; Johns, 2001; Krystal et al., 1994). Language, a complex and commonly used cognitive process, is often altered by the acute effects of PADs (Salomé et al., 2000). Nevertheless, linguistic markers are not routinely evaluated when assessing PAD intoxication in clinical, legal or other service settings.

Testing levels of PAD in blood or urine is often used to quantify the amount of consumption (Margolis et al., 2016). For several substances, their systemic levels may not reflect their neurocognitive effects. In other words, a positive test may not translate to observed behaviours (Durback et al., 1998), for various reasons including the individual variations in PAD metabolism (Vizeli et al., 2021). Developing objective markers of the effects of PADs on current mental state would be beneficial for psychiatric assessment in urgent and acute settings, to determine whether a clinical presentation (of, say, psychotic symptoms) is substance-induced.

In research settings, being able to objectively assess the psychoactive effects of a PAD on the mental state would help reduce the overreliance on subjective self-reports, especially as expectancy effects and unblinding are common in this setting (Earleywine

et al., 2023; Muthukumaraswamy et al., 2021). Wilful manipulation of one’s speech/writing style to mimic an intoxicated state is possible, but many PADs affect the effortful control of psychomotor functions including language. This makes intentional change of linguistic markers less likely when one is truly under the acute effects of PADs, prompting some to call speech as a ‘window to the intoxicated mind’ (Bedi et al., 2014). Computerised analysis of verbal descriptions obviates the need for introspective self-awareness that is often altered by PADs and provides an alternative with a more empirical quantitative approach that may complement reliability of clinical ratings (Blaabjerg et al., 2020; Espejo et al., 2024; Ross and Leichner, 1988).

By computerised analysis of verbal production, we refer to the automated analysis of expressed language (speech and written text) using natural language processing (NLP) tools. NLP is a

¹Schulich School of Medicine and Dentistry, University of Western Ontario, London, ON, Canada

²Douglas Mental Health University Institute, Montreal, QC, Canada

³Department of Psychiatry, McGill University, Montreal, QC, Canada

⁴Robarts Research Institute, University of Western Ontario, London, ON, Canada

Corresponding author:

Lena Palaniyappan, Douglas Mental Health University Institute, T-Pavilion, 6875 Boulevard Lasalle, Verdun, QC H4H 1R3, Canada.
Email: lena.palaniyappan@mcgill.ca

Table 1. Eligibility criteria.

Inclusion criteria	Exclusion criteria
Participants consumed a drug (or had a diagnosed substance-use disorder) from one of the following PAD categories ^a at any timepoint before an audio recording or written transcript was collected: psychedelic, hallucinogen, entactogen, stimulant, cannabis or cannabinoid	The study only included patients with a diagnosed psychiatric disorder, or patients with smoking-induced laryngeal changes ^b
There was an audio recording or text transcript of participants' speech or writing collected in a controlled clinical or research setting	Transcripts were collected from publicly available drug experience repositories, social media or blog posts
The audio recordings or text transcripts were processed using a computer-based analysis	The study only used manual language analyses ^c
	Review studies, conference abstracts and manuscripts in pre-print servers
	The study does not report findings relevant to the effect of PADs on NLP markers ^d

^aAs classified by a public database on PAD effects (PsychonautWiki Contributors, 2024).

^bLee et al. (1999), Mueller and Wilcox (1980).

^cFor example manual cloze analyses (Amarel and Cheek, 1965; Honigfeld, 1965; Weil and Zinberg, 1969).

^dAlexander-Emery et al. (2005), Cox et al. (2021).

subdiscipline of computer science focused on allowing computers to understand and interpret human (i.e. non-programming, 'natural') language. NLP methods convert raw natural language data – text or audio of an individual's language production – into numerical or categorical features and provide various readouts of *how much* is being said (rate and amount of speech, verb density, idea density), *what* is being said (meaning, sentiment and topics conveyed and diversity and similarity among words), and *how* it is being said (the structure of the word sequences, prosody and acoustic aspects of speech) (Le Glaz et al., 2021). From studies that employ human raters, PADs are well established as agents that affect the above aspects of language. For example, ketamine causes disorganised speech resembling psychosis and is termed 'ketamine-induced thought disorder' (Adler et al., 1998; Krystal et al., 1994). Furthermore, methamphetamine improves whereas 3,4-methylenedioxymethamphetamine (MDMA) negatively affects verbal fluency, reflecting these PADs' psychoactive effects on other cognitive processes like attention and memory (Marrone et al., 2010). Finally, cannabis causes speech with greater discontinuity in thought, hypothesised to be caused by cognitive effects such as intruding associations (Roth et al., 1975). A recent review of psychedelics also highlights the value of objective analysis of language production to generate reliable markers of their psychoactive effects (Tagliazucchi, 2022). The quantitative features generated by the NLP analyses can enter machine learning (ML) classifiers to accurately identify if a given speech sample comes from one group (e.g. after cannabis use) or the other (no cannabis use) (Le Glaz et al., 2021).

Here, we review the current literature on how NLP markers are affected by PADs and the accuracy to which classifiers can predict drug status to support the synchronisation of research efforts towards these future applications. Our research questions are:

1. How do PADs including cannabis/cannabinoids, psychostimulants, psychedelics/hallucinogens and entactogens affect language production as measured by NLP markers?

2. How much accuracy has been achieved in predicting a person's recent use of PADs using computerised language analyses?

To this end, we undertook a pre-registered systematic literature search with a view to meta-analytically combine accuracy metrics provided sufficient number of eligible studies are identified for such statistical pooling. This systematic review followed the Preferred Reporting Items for Systematic Review and Meta-analysis (PRISMA) 2020 guidelines. Details of the protocol for this systematic review were registered on PROSPERO (International prospective register of systematic reviews, Ahuja et al., 2023).

Methods

Search strategy and eligibility criteria

For this systematic review, a MEDLINE and Embase search was conducted on 12 June 2023 and 5 March 2024, for articles with terms according to the following two concepts: first, a specific PAD in the psychedelic, hallucinogen, entactogen, stimulant, cannabis or cannabinoid classes, or a related substance-use disorder; and second, a term related to either language or NLP. The specific search terms we utilised are reported in the Supplemental Materials and Methods. Overall, our search found 1051 unique abstracts (discarding duplicates): 961 from the initial search in June 2023, and 90 from the updated search in March 2024. Our eligibility criteria are outlined in Table 1.

Study selection and data collection process

The Covidence software (Covidence, 2023) was used to manage study selection (Figure 1) and data collection. For articles from the first search in June 2023, study screening and subsequent data collection was conducted by two authors independently (SA and FZ), and conflicts were resolved via discussion between the two authors. One author reviewed studies from the updated search in

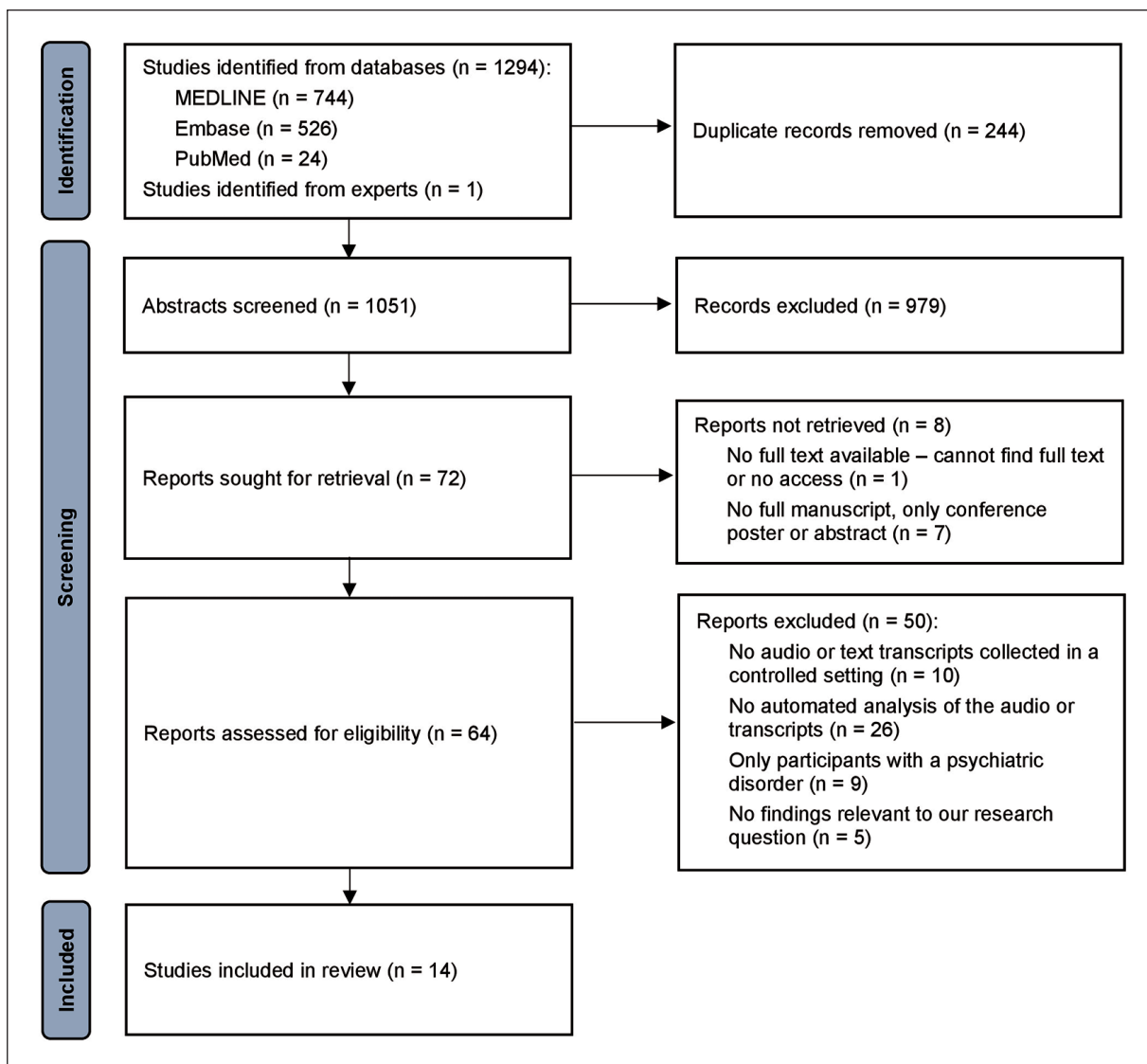


Figure 1. PRISMA flow diagram.

March 2024 (no studies were eligible). The senior author (LP) reviewed the final list independently.

Data items

The variables extracted from each study included the study aims, recruitment and eligibility criteria, participant demographic variables, study design, drug name and dosing regimen, language task and timing, NLP and classification analyses employed and study results including the effect of study drug on NLP markers and the classification accuracy between study conditions. For variables missing from included publications, the corresponding authors were contacted.

We could not undertake a meta-analysis due to the high degree of heterogeneity among outcome measures and the lack of sufficient number of studies. Therefore, data were synthesised qualitatively for all variables except demographics, which were synthesised quantitatively using the Cochrane Review Manager

(Review Manager Web, 2020) and reported as pooled data (Supplemental Table 1). Further details on statistical methods are described in the Supplemental Materials and Methods.

Risk of bias assessments

Different tools were used to assess risk of bias based on each study's design: the Risk of Bias in Non-randomized Studies – of Exposures (ROBINS-E) tool was used for cross-sectional studies; the Risk of Bias in Non-randomized Studies – of Interventions (ROBINS-I) tool was used for non-randomised trials; and the Risk of Bias 2 (ROB 2) tool was used for randomised crossover trials. The risk of bias assessments can be found in Supplemental Figure 2.

Results

Figure 2 provides a summary of the fourteen included studies categorised by PAD. This figure summarises the types of

Study Overview					Computerized Language Analyses			
Study	Location	Design	Language Data	Bias Risk	NLP Analyses	Results of NLP Analyses	ML Classifiers	
MDMA & other Psychostimulants	Wardle 2012 [30]	Chicago, USA	Speech about significant other after oral d-amphetamine and placebo (English)	+	B	(B) Amphetamine increased talkativeness.		
	Bedi 2014 [10]	Chicago, USA	Speech about significant other after oral MDMA, methamphetamine, and placebo (English)	+	A, B, H, S, T	(B) Methamphetamine increased word count. (H) MDMA increased semantic proximity of speech to "friend", "support", and "empathy". (T) Methamphetamine decreased the frequency of repeated words.	W, Y	
	Wardle 2014 [31]	Chicago, USA	Speech about significant other after oral MDMA and placebo (English)	+	F	(F) MDMA increased use of positive emotion words.		
	Baggott 2015 [32]	Chicago, USA	Speech about significant other after oral MDMA and placebo (English)	-	A, F, S	(F) Across phrases which described the person of interest, MDMA increased words in the "cognitive mechanisms" category ¹ .	V	
	Agurto 2020 [33]	Chicago, USA	Speech about significant other and free monologue after oral MDMA, oxytocin, and placebo (English)	+	A-C, G, H, L-P, R, S	Significant other speech task: MDMA 1.5 mg/kg increased (B) total word count, (C) lexical richness, and (G) number of ideas; (H) MDMA 0.75 mg/kg decreased semantic proximity to "sad". Free monologue speech task: (H) MDMA 0.75 mg/kg increased semantic similarity to "think", "talk", and "feeling". (L-P, R) MDMA did not significantly affect acoustic markers.	V, W, X	
Ket.	Covington 2007 [34]	Cambridge, UK; Vancouver, Canada	Speech describing pictures from TAT after IV ketamine, in schizophrenia patients, and in controls (English)	+	A, C, G, S	Ketamine increased (C) repetitiveness and decreased (S) verb density compared to placebo within-subjects but not between-groups. (G) Ketamine increased the variance in idea density between participants.		
LSD	Sanz 2021 [35]	London, UK; Natal, Brazil	Verbal experience report after IV LSD or placebo (English), and verbal dream report by schizophrenia patients, bipolar patients, and healthy controls (Portuguese)	+	A, B, C, E, S, T	(B) LSD increased word count at time 1 and 2. (C) LSD increased speech disorganization (Shannon's information entropy) at time 1 and 2. (E) LSD decreased semantic similarity to "mood" at time 1 and 2. (T) LSD reduced the lexicon and increased recurrence metrics (word loops and parallel edges). (S) LSD increased semantic variability and reduced redundancy of speech.	V	
	Wießner 2023 [36]	Campinas, Brazil	Verbal and written experience reports and storytelling at multiple timepoints after oral LSD and placebo (Portuguese)	+	B, E, F, T	(B) LSD decreased verbosity. (E) LSD decreased 3 measures of average semantic distance: distance of each word to all predecessors, between neighbouring words, and between all words. (F) LSD decreased time- and space-related words, and increased words in multiple categories (dictionary, verbs, punctuation, they, present, and biology-related words). (T) LSD decreased markers of the lexicon, and global and local connectivity.		
Psilocybin	McCulloch 2022 [37]	Copenhagen, Denmark	Sober written report post-psilocybin experience (Danish)	+	A, D	(D) The words with the highest significance to reports of psilocybin-induced Complete Mystical Experiences, included "Universe", "dad", "Magnetic Resonance", "beautiful", "simultaneous", "infinite", etc., and these terms were found across many reports.		
	Sanz 2022 [38]	Buenos Aires, Argentina	Verbal experience report after oral psilocybin and placebo (Spanish)	+	A, B, E, K	(B) Psilocybin increased word count for 3 of 6 interview questions. (E) Psilocybin did not affect semantic variability. (K) Psilocybin increased positive sentiment in response to the "perception" question.	V	
Cannabis & THC	Zeidenberg 1973 [39]	New York, USA	Free monologue speech at onset of effects of oral THC and placebo (English)	-	L	(L) THC increased mean pause duration, decreased mean phrase duration, and increased mean vowel duration ¹ .		
	Shamei 2019 [40] & 2021 [41]	Victoria, Canada	Reading standardized passage and word lists before and after cannabis administration (English)	+	L, M, O, R	(L) Cannabis increased the standard variation of voice onset time for voiced phonemes. (O) Cannabis reduced the total pitch range and the intrasyllabic trajectory, without affecting intersyllabic trajectory. Decrease in phonation trajectory was not significant (p=.06). (R) Cannabis reduced shimmer but not jitter.	W, Z	
	Vogel 2021 [42]	Adelaide, Australia	Acoustic-focused speech tasks in cannabis naive vs. experienced adults (English)	+	L, N-Q	(N) Cannabis-experienced speech had a lower alpha ratio on the reading task, and this effect correlated with the duration of abstinence from cannabis. (L, O-Q) No other markers differed between groups after controlling for confounders.		
	Murray 2022 [43]	Chicago, USA	Written self-report after oral THC and placebo (English)	+	B, D, E, I, J, U	THC did not affect (B) word count, (E) semantic coherence, or (U) sentence perplexity. (E) THC dose-dependently increased color entropy. (J) Two topics emerged, related to placebo and THC writing.	W	

Icon Legend	Design	Bias Risk	NLP Analyses	ML Classifiers
	Cross-sectional study Non-randomized trial Single-blind, within-subjects RCT Double-blind, within-subjects RCT	Low risk of bias Some concerns of bias High risk of bias		

Alphabetical Index	NLP Analyses				ML Classifiers
	Lexical Analyses A. Lemmatization B. Word count C. Complexity metrics ^{III} D. Term frequency-inverse document frequency ^V	Semantic Analyses E. Word embedding F. Linguistic Inquiry and Word Count G. Idea density metrics ^V H. Latent semantic analysis I. Document embedding J. Latent Dirichlet allocation K. Sentiment analysis	Acoustic Analyses L. Temporal features ^{VI} M. Mel scale frequency markers N. Noise features ^{VII} O. Pitch P. Vowel space features ^{VIII} Q. Intensity (decibel) measures R. Stability features ^{IX}	Other S. Parts of speech T. Speech graphs U. Sentence perplexity ^X	V. Random forest W. Support vector machine X. k-Nearest Neighbours Y. Linear discriminant analysis Z. Neural network

Figure 2. Natural language data source, NLP analyses and findings and ML classifiers used for included studies categorised by PAD. See bottom of figure for icon legend and computer analysis alphabetical index – this categorisation scheme was adapted from Le Glaz et al. (2021). Bold letters are NLP analyses whose results were used as inputs for the study’s classifiers.
¹After we applied a post hoc Bonferroni adjusted *p*-value cutoff of 0.0012 for 43 hypothesis tests – one for each Linguistic Inquiry and Word Count (LIWC) word category – only this LIWC category was significant.
^{II}Significance testing was not performed in the original study, but based on our post hoc paired *t*-tests, all associations were significant (Supplemental Materials and Methods).
^{III}Includes metrics related to the total lexicon, such as Honore’s statistic, Brunet and Shannon’s index and Covington and Morris’s perseveration metric.
^{IV}Does not include use of LIWC program solely for total word count.
^VBased on the number of prepositions in text estimated using a parts-of-speech (POS) tagging-based method.
^{VI}Temporal features such as the distribution of pauses and utterances, or the voice onset time.
^{VII}For example, harmonics-to-noise ratio, noise-to-harmonics ratio, mean autocorrelation and alpha ratio of spectral tilt.
^{VIII}For example, total area, ‘a-i-u’ area and formants 1,2,3 distribution.
^{IX}For example, jitter, shimmer and voice breaks.
^XAssessed using context-dependent transformer-based approaches (BERT, GPT-2).

computerised analysis methods used (NLP and ML classifiers), as well as the directional effects of each PAD on NLP markers employed.

Nature of participants across studies

Across the included studies, there were a total of 406 participants (range 4–71 per study). No participants had a history of psychiatric or substance use disorders, except for comparator subgroups that is, schizophrenia (Covington et al., 2007; Sanz et al., 2021) and bipolar disorder (Sanz et al., 2021) groups. However, participants' history of prior PAD use varied: previous use of the studied PAD (or related drugs) was required for inclusion in nine studies (Agurto et al., 2020; Baggott et al., 2015; Bedi et al., 2014; Murray and Srinivasa-Desikan, 2022; Shamei and Bird, 2019; Vogel et al., 2021; Wardle et al., 2012; Wardle and De Wit, 2014; Wießner et al., 2023), while individuals with prior psychedelic drug use history were excluded in one study (Sanz et al., 2021).

The mean age of the included sample was 26.8 years ($N=375$), with 59% of the participants being males. The most represented ethnic group in the sample was White/Caucasian (self-reported) and the language that formed the medium of assessment was mostly English (ten studies) followed by Portuguese (two studies).

Study designs across drugs

We identified double-blind randomised controlled trial (RCTs; $n=8$) as the major study type, while single-blind RCTs, non-randomised trials and cross-sectional studies were also reported. While most studies had a within-subject contrast, one study (Vogel et al., 2021) reported only between-subjects data. There were numerous approaches used for eliciting verbal production – two studies relied on written materials (McCulloch et al., 2022; Murray and Srinivasa-Desikan, 2022), while 12 had speech data. When speech was elicited, instructions were highly variable with five studies collecting speech about an important person in the participant's life (significant other) (Agurto et al., 2020; Baggott et al., 2015; Bedi et al., 2014; Wardle et al., 2012; Wardle and De Wit, 2014), three studies collecting free monologues (Agurto et al., 2020; Vogel et al., 2021; Zeidenberg et al., 1973), three studies collecting experience reports (Sanz et al., 2021, 2022; Wießner et al., 2023), two studies conducting reading tasks (Shamei and Bird, 2019; Vogel et al., 2021) and two studies providing picture prompts (Covington et al., 2007; Wießner et al., 2023). We also noted a wide range of time intervals for language tasks after the cessation of drug intake – 30 min (Covington et al., 2007; Shamei and Bird, 2019) to 24 h (Wießner et al., 2023); while some of this was dictated by the pharmacokinetics of the specific PAD, even within a single class of agent, both the dose (50–75 µg lysergic acid diethylamide [LSD]) and the time interval after administration varied (e.g. 75–180 min for MDMA, 120–225 mins for LSD and 30–100 min for cannabis/THC). Furthermore, the length of time provided for speaking or writing varies between a 5-s sustained vowel production (Vogel et al., 2021) to 10-min speech samples (Agurto et al., 2020; Bedi et al., 2014; Covington et al., 2007; Murray and Srinivasa-Desikan, 2022).

Regarding bias in study designs (Supplemental Figure 2), we noted two studies had higher risk than others based on bias in

measurement of outcomes and selection of the reported result (Baggott et al., 2015; Zeidenberg et al., 1973). Consistent with the initial epochs in a field of research, we note an exploratory approach in statistical testing, analysing multiple variables without corrections for inflated type 1 error; for example, 20 or more themes for semantic proximity analysed (in Agurto et al., 2020; Bedi et al., 2014), and 43 variables from Linguistic Inquiry and Word Count (LIWC) analysed (in full transcript and selected parts, Baggott et al., 2015).

Methods used for speech/text analysis

The most common lexical analyses used were lemmatisation (a pre-processing method used by seven studies) and word count (six studies). A range of semantic analyses were employed that can be broadly grouped under dictionary/corpus-based (LIWC and parts-of-speech (POS) tagging-based idea density), distributional semantics (latent semantic analysis, word embedding and document-embedding models) and topic modelling approaches (Latent Dirichlet allocation and sentiment analysis). Word embedding, the most common semantic analysis (four studies), was done using the Word2Vec (Murray and Srinivasa-Desikan, 2022; Sanz et al., 2021) and FastText (Sanz et al., 2022; Wießner et al., 2023) models. A popular semantic approach used in three studies (Agurto et al., 2020; Bedi et al., 2014; Sanz et al., 2021) was using word embeddings or LSA to test the similarity of the verbal output against a predetermined group of words representing distinct themes of mental states to provide values of *semantic proximity* for each selected theme. Even with this approach, the same set of words were not used for similarity assessment across studies: Bedi et al (2014) used terms for mood states affected by MDMA intoxication, Agurto et al. (2020) added an additional item to this set and Sanz et al. (2021) used terms relevant to LSD-intoxicated speech. Temporal analysis was the most common acoustic approach (three studies). Overall, we found no pair of studies using the same approach for speech or text analysis among our search results.

In the section below, we summarise the drug-specific results reported in more than one study with a view of identifying potential NLP variables for further research. As only one study on ketamine was included (Covington et al., 2007), the effects of ketamine reported in that single study (higher repetitiveness of words with a reduction in verb density) are not discussed any further.

Replicated drug-specific results

Psychostimulants (amphetamine and methamphetamine). Both Wardle et al. (2012, 20-mg oral *d*-amphetamine, 180-min post-ingestion, 5 mins talking about others) and Bedi et al. (2014, 20-mg oral methamphetamine, 130-min post-ingestion, 10 mins talking about others) reported an increase in talkativeness (word count) with stimulants compared to placebo. This effect was not statistically significant at 10 mg dose of *d*-amphetamine, though no significant differences were reported between the 10 and 20 mg doses. Bedi et al. (2014) also reported a methamphetamine-related reduction in the sequential repetition of the same word (e.g. 'I . . . I think we are late'), and of the 20 different themes tested for semantic proximity, noted a reduction

in words related to ‘compassion’ with methamphetamine compared to placebo. These effects are yet to be replicated. Taken together with studies reporting a significant increase of manually counted syllables 90 mins after oral methamphetamine (Marrone et al., 2010), increased verbosity has been a replicated feature of psychostimulants.

3,4-Methylenedioxymethamphetamine. Bedi et al. (2014) reported an increase in semantic proximity to the positive emotional words ‘friend’ and ‘support’ after 1.5 mg/kg and ‘empathy’ after 0.75 mg/kg of oral MDMA (130-min post-ingestion, 10 mins talking about others), with no changes in verbosity and word repetition patterns in Mota’s graphs, compared to placebo. Agurto et al. (2020) were unable to replicate semantic proximity changes for the same themes despite using a comparable paradigm (0.75 and 1.5 mg/kg oral MDMA, 5 mins talking about others, 75 to 105-min post-ingestion); instead, they reported decreased proximity to ‘sad’ at 0.75 mg/kg dose and increased word count, lexical richness and number of ideas at 1.5 mg/kg dose compared to placebo. Wardle et al. (2014) also used a comparable paradigm in their article (0.75 and 1.5 mg/kg oral MDMA, 5 mins talking about others, 140-min post-ingestion) but did not study semantic proximity. Using dictionary-based word counts, they noted that both doses increased the proportion of positive emotional words, without affecting percentage of negative words used to describe another person. Baggott et al. (2015) used the same task and counting procedure as Wardle et al. but did not see any changes in positive or negative emotion words or overall verbosity 90 mins after MDMA ingestion (1.5 mg/kg). Instead, they noted a decrease in relative terms and motion related words, an increase in words with sexual and social content, words reflecting discrepancies, future and death, and when focusing on phrases that specifically referred to the person being discussed, an increase in references to cognitive processes and insight. Thus, there are no directly replicated results regarding the specific themes affected following MDMA ingestion, though at a broader level, an increase in social and positive emotion content recurs across studies.

LSD. Sanz et al. (2021) administered intravenous LSD 75 µg or placebo over two sessions and recorded speech (on the experience of undergoing neuroimaging without time limit) at two time points 120–150 min and 225 min post-infusion. LSD affected the word trajectory structure by increasing verbosity and recurrence along with reduced global connectivity (i.e. long and disconnected with reduced lexicon), affected semantic content by increasing proximity to the word ‘mood’, increasing entropy and semantic variability (reduced coherence). Using a similar experience report approach but with 50 µg oral dose of LSD, Wießner et al. (2023) failed to see any alteration in structure and semantics. But in the same sample, the use of picture description (‘story telling’) revealed *reduced* verbosity and global connectivity with higher recurrence (i.e. short and disconnected with reduced lexicon), affected semantic content with less ‘time’ and ‘space’ related, but more biology-related words, and *decreased* semantic distance between words (more coherence). While there are no directly replicated results on the semantic content and coherence after LSD, reduced lexicon with a disconnected word trajectory has been demonstrated as a post-acute effect in two different tasks, at two different doses.

Psilocybin. In Sanz et al.’s (2022) microdosing study, volunteers received placebo or a 0.5 g dose of oral-dried *Psilocybe cubensis* mushrooms (with 0.32 mg psilocybin, 0.48 mg psilocin, 0.025 mg baeocystin and 0.0063 mg norbaeocystin on average) and interviewed 150-min post-ingestion on six topics including feelings, expectation and creativity with no time constraints. Active dose induced higher verbosity especially when asked about mood, alertness and perception, and more positive sentiment when asked about perception, with no changes in semantic variability. While there were no other methodologically similar studies, McCulloch et al. (2022) generated unconstrained, freely written qualitative reports 6–8 h after a medium-high dose (12–30 mg) to compare themes present among those who experienced a Complete Mystical Experience (CME) versus those who did not (groups determined using a 30-item self-report questionnaire). Those who experienced CMEs used words with positive sentiment, relationships and time/space dimensions (e.g. ‘universe’, ‘dad’, ‘beautiful’, ‘simultaneous’, ‘infinite’, ‘purple’, ‘in relation to’, ‘ray’, ‘happy’ and ‘brother’), while those who did not experience the active dose effect used words with negative sentiment more often (e.g. ‘gloomy’, ‘cycle’, ‘evil’, ‘cold’ and ‘need’). Thus, while there are no directly replicated results on the structure and content of verbal expression after psilocybin use, a preferential expression of positive sentiment has been reported in two studies at least in a subgroup of volunteers.

Cannabis and THC. Speech studies on cannabis have generally focussed on acoustic effects but have several methodological shortcomings. The unblinded test–retest designs used by Zeidenberg et al. (1973) included only four participants (male psychiatric resident physicians; 15 mg oral THC, re-tested at 90 mins; 5-min monologue), while Shamei et al. (2021; Shamei and Bird, 2019) included eight (50% females; variable doses, re-tested at 30 mins; reading passages and word lists without time constraint) medicinal cannabis users. These two studies did not test the same speech variables; Zeidenberg et al. (1973) reported increased pause length and vowel duration with reduced phrase duration while Shamei and Bird (2019) reported altered speech stream acoustics (reduced shimmer, pitch range, reduced trajectory of pitch within syllables with increased variability of voice onset time). Vogel et al. (2021) also observed increased pause length with decreased voice onset time in recreational cannabis users without any experimental cannabis administration, indicating a sustained, longer-term effect of cannabis use on speech acoustics.

Murray et al. (2022, oral THC 7.5 mg, 15 mg or placebo administration, 10 mins of written reports 100 mins later) calculated semantic perplexity (unlikelihood of a sentence in a transcript), semantic ‘entropy’ (complement of semantic distance among words in a sentence determined on the basis of 300 dimensional vectors from Word2Vec) and colour ‘entropy’ (semantic distance limited to 8-colour dimensions based on 100 images per word, that is, banana is more colour-distant from apple than from lemon). While number of words generated and semantic distance did not change with THC, a dose-dependent increase in ‘colour entropy’ was noted along with a trend level increase in semantic perplexity. Certain words identified using Latent Dirichlet Allocation (LDA)-based topic modelling revealed distinct topics with words ‘time’, ‘moving’, ‘weird’, ‘drug’ and ‘embarrassed’ representing THC-related effects. These methods are yet to be repeated in other samples.

In summary, amidst the methodologically weak studies on the verbal production after cannabis use, there are no directly replicated results; nonetheless, pause length and voice onset times (speech stream markers) appear to be interesting candidates.

Classification results

In total, 7 of 14 identified studies reported the results of a ML-based classification that aimed to discriminate the speech pattern induced by active drug from placebo. Attempts for external validation has been made only for the MDMA versus placebo contrast, with better-than-chance classification accuracies alongside several limitations in feature selection and sample size, warranting a cautious interpretation.

Bedi et al. (2014) built a Support Vector Machine (SVM) classification model using semantic proximity to ‘rapport’, ‘love’ and ‘support’, as well as word count, and reported an accuracy of 69% in classifying between placebo and methamphetamine-influenced speech, 88% accuracy between MDMA 1.5 mg/kg and placebo and 84% accuracy between MDMA 1.5 mg/kg and methamphetamine (leave one out cross-validations). While speech after MDMA 0.75 mg/kg dose was not clearly separable from placebo, a linear discriminant analysis conducted to classify between all four study conditions and had an accuracy of 59%, with chance corresponding to 25%. Baggott et al. (2015) built a random forest classifier using a bag-of-words approach from the transcripts (using frequency of all words as predictors) and reported an out-of-bag accuracy of 72% for the final model classifying speech after 1.5 mg/kg MDMA from placebo. Agurto et al. (2020) assessed diverse NLP markers (88 acoustic features, 21 semantic proximity concepts (as in Bedi et al., 2014, with the addition of ‘disdain’), and lexical/syntactic features including total word count, number of ideas, propositional density, 6 parts of speech and 2 measures of lexical richness) with SVM and observed accuracies ranging from 71% to 87% (nested leave-one-participant-out cross-validation) when separating speech after MDMA from placebo in the original dataset, with major contributions from semantic proximity (‘anxiety’) and acoustic markers (mel-frequency cepstral coefficients, pitch). When using the same speech NLP features to validate the classifiers on other datasets (Bedi et al., 2014; Wardle and De Wit, 2014), the highest classification accuracies of 66%–92% were found in classifying between MDMA 0.75 mg/kg and placebo.

Sanz et al. (2021) used two random forest classifiers to separate LSD and placebo conditions: one using semantic similarity with ten predefined terms had an area under the curve (AUC) of 0.757, while the one using structural metrics had an AUC of 0.749. For psilocybin versus placebo, Sanz et al. (2022) utilised word count and sentiment scores from six interview questions and built a random forest classifier with stratified cross-validation, reporting an AUC of 0.79. Murray and Srinivasa-Desikan (2022) built an SVM classifier using document embeddings with Term Frequency Inverse Document Frequency (TF-IDF) format and reported an accuracy of 87.5% (AUC of 0.94) in classifying writing under the influence of THC from placebo. Shamei et al. (2021) built SVM, and convolutional neural network (CNN) classifiers using Mel spectrograms from a sustained vowel production task and reported accuracies of 67.2% (female) and 68.2% (male) in

classifying cannabis intoxicated speech (gender-specific CNN models with *n*-fold cross-validation; pre-post design).

Discussion

We reviewed the current literature on how PADs affect NLP markers of verbal expressions and whether intoxication/use status can be accurately predicted with speech and language markers. The result of our systematic search indicates that PADs do affect language production measured by computerised analyses, though consistent patterns required to achieve sufficiently accurate predictions of a person’s recent substance use are not yet available. We discuss the interpretations from the most consistent results, gaps identified in the literature and recommendations for future studies in the sections below.

We found insufficient data and prominent methodological heterogeneity that precluded a pooled synthesis of reported effects across all psychoactive substances reported here. Nonetheless, in more than one study, stimulants (*d*-amphetamine and methamphetamine) increased word count (Bedi et al., 2014; Wardle et al., 2012), in line with the findings from manual analyses of speech (Jaffe et al., 1973; Marrone et al., 2010; Stitzer et al., 1978). Similarly, MDMA seemed to increase social and positive emotion content in speech, consistent with its well-established prosocial effects on subjective mood states (Kamilar-Britt and Bedi, 2015). While these NLP effects do not seem to be dose-dependent (since semantic effects occurred at MDMA 0.75 mg/kg more than 1.5 mg/kg, Agurto et al., 2020), the emerging pattern of MDMA’s effects on sentiment suggests the value of including semantic proximity markers in future studies. Under the influence of LSD (peak effects at 1.5 to 4 h, Dolder et al., 2017), speech becomes more recurrent with a reduction in the use of unique words (Sanz et al., 2021), leading to less connected structure (Wießner et al., 2023). The degree of this disorganisation is limited when compared to the speech of patients with schizophrenia (Mota et al., 2017). Finally, psilocybin appears to increase positive sentiment with ‘microdosing’ (Sanz et al., 2022) and with larger doses in those who report mystical-type psychoactive effects (Garcia-Romeu et al., 2015; Griffiths et al., 2016). Taken together, these observations provide a useful starting point for pursuing automated NLP markers of psychoactive effects that may track their physiological effects in social and clinical settings.

We notice several gaps in the research that aims to apply computational linguistics/NLP in the study of mental state changes induced by PADs. Applying the framework suggested by Miles (2017; Nyanchoka et al., 2019), we tabulate the identified gaps in Table 2.

Several directions for future studies emerge from this work. Some of these are tabulated in Table 2. Classification accuracies of drug status are often above chance and have been validated using external cohorts in the case of Agurto et al. (2020). Nevertheless, classifiers have large variation in accuracy across studies, with no replication studies of model accuracy. There is an urgent need for the use of a shared protocol for speech data collection, deriving consensus on experimental parameters (e.g. post-ingestion timing of acquisition, recording equipment standards, pre-processing routines) and collaborative sharing of

Table 2. Some of the identified gaps in the NLP literature on psychoactive drugs.

Gap domain and contributing observations	Key questions/needs
<p>Knowledge gap: Classifier performance of NLP markers not evaluated against human judgement of intoxication; languages other than English not studied, classification not attempted yet for certain drugs</p> <p>Evidence gap: Results from the reviewed studies are contradictory, needing more evidence and multi-method approaches for clarity</p>	<ul style="list-style-type: none"> • When assessing intoxication, can objective NLP markers perform better than or provide incremental value to clinical judgement? • Do the NLP markers tracking PAD effects generalise across languages? • Can NLP-based classifiers accurately identify psychostimulant use on an individual basis?
<p>Empirical gap: Replications are lacking for several observations, while other propositions are yet to be evaluated empirically</p>	<ul style="list-style-type: none"> • Does verbosity increase in dose-dependent manner with stimulants? Does MDMA affect verbosity in the same way as stimulants? • What is the effect of the duration of interview on semantic and acoustic markers sensitive to PADs? • For each PAD, what speech tasks are the most sensitive to minimum detectable changes? • Does ketamine affect repetitiousness and idea density as reported? • Do stimulants reduce sequential word repeats and reduce 'compassion' in speech? • Does LSD reduce verbosity? Is 'storytelling' superior to 'experience reporting' to elicit LSD-related speech changes? • Does cannabis affect semantic coherence/perplexity? Is cannabis-induced alteration in speech stream a long-term effect that persists beyond its peak physiological effects? • Does microdosing with psilocybin affect speech output for selected themes only?
<p>Population gap: Clinical populations for whom end results are applicable are under-researched</p>	<ul style="list-style-type: none"> • Do NLP markers change in persons with SUD to the same extent as in healthy volunteers with experimental exposure? • Are the effects of PADs sufficiently distinct from the effect of disorders such as schizophrenia, depression and mania? • Some speech markers are more sensitive than others to the social background (race/ethnicity, cultural background, socioeconomic status). How does this affect the drug effect?
<p>Theoretical gap: Pharmacological and cognitive theories not being applied in variable selection</p>	<ul style="list-style-type: none"> • Cannabis may induce paranoia by increasing self-relevance of stimuli. Does semantic content reflect this effect? • As pharmacological response shows inter-individual variations, can we identify clusters/subgroups of subjects with varying speech patterns using unsupervised classifiers?
<p>Methodological gap: Lack of control arm and blinding is common, dialogical/social interactions with two or more people remain unstudied</p>	<ul style="list-style-type: none"> • Are speech changes induced by cannabis observable in double-blind RCT design? • What is the influence of interviewer's facilitations in sentiment and content analysis reported in MDMA, LSD and psilocybin studies?
<p>Action gap: Observed research practices deviate from perspectives gained from other fields</p>	<ul style="list-style-type: none"> • Need to attempt replication of prior studies when new experiments are designed • Need for overlapping protocols of language data acquisition affecting comparisons across substances • Need for accessible/shared corpora to enable external validation of classifiers • Lack of structured reporting of data capture, curation, processing pipelines, model selection, parameter tuning and algorithmic decisions • Missed opportunities of collecting language samples during therapeutic trials (e.g. MDMA in PTSD, psychedelics/ketamine in depression)

SUD: substance use disorders; RCT: randomised controlled trial; PTSD: post-traumatic stress disorder; PAD: psychoactive drug; NLP: natural language processing; MDMA: 3,4-methylenedioxymethamphetamine.

recorded speech samples to enable systematic studies of the effects of various NLP analytical pipelines. Tagliazucchi (2022) recommended that researchers conducting therapeutic trials of PADs routinely collect natural language samples and analyse them using NLP to validate the psychometric properties of the NLP markers. Additionally, we recommend that ongoing studies where speech is being collected as a by-product of planned assessments analyse this speech data using NLP. Future studies should attempt to replicate promising NLP findings such as those identified in this review to accelerate clinical translation. Future work on PAD-induced speech effects should also explore other PADs, study designs, dosing regimens, languages, speech tasks and NLP markers (such as markers known to be affected in psychosis (Corcoran et al., 2020) and depression (Koops et al.,

2021)) while keeping the most likely use cases for a speech-based ML classifier in mind. For speech recordings to be used in the identification of an individual's substance use status in clinical settings, designing studies that compare one PAD against another, as well as against other altered mental states such as psychosis or mood episodes, are essential.

A limitation of the current pre-registered review is that it does not include several psychoactive substances that may affect speech, but are less likely to present with a particular challenge in identifying states of intoxication, and have limited NLP research to date (e.g. depressants, neuroleptics). Also, a numerical synthesis of effect size differences or classification accuracies was not attempted due to the high degree of methodological heterogeneity. Given these issues, we urge caution when interpreting the

findings on drug-specific NLP changes reported in our review. Our observations are best considered as a starting point for further experimental work in this field.

In summary, language production appears to be consistently affected by PADs. Objective, computerised analyses appear sensitive to detect these changes. With continued efforts towards investigating promising NLP markers and open questions in the field, speech - a highly accessible and non-invasive (Palaniyappan et al., 2022) source of biosocial markers (Palaniyappan, 2021) - can be leveraged to expand human behavioural pharmacology.

Acknowledgements

The authors acknowledge the support of Ms. Alanna Marson, professional healthcare librarian, in helping design the search strategy.

Author contributions

SA and LP designed the study. SA conducted the literature searches. SA and FZ performed study screening and data extraction. SA conducted the risk of bias assessments and wrote the first draft of the manuscript. LP and SA contributed to manuscript revision, and all authors read and approved the final submission. SA had full access to the data and assumes responsibility for its integrity and is available for data-related queries at sahuja2026@meds.uwo.ca.

Availability of data, code and other materials

The online version contains a supplementary datasheet of the methodological variables extracted from the primary studies included in this review.

Declaration of conflicting interests

The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: LP reports personal fees for serving as chief editor from the Canadian Medical Association Journals, speaker/consultant fee from Janssen Canada and Otsuka Canada, SPM Course Limited, UK, Canadian Psychiatric Association; book royalties from Oxford University Press; investigator-initiated educational grants from Janssen Canada, Sunovion and Otsuka Canada outside the submitted work. SA and FZ have no competing interests to disclose.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: L. Palaniyappan acknowledges research support from the Canada First Research Excellence Fund, awarded to the Healthy Brains, Healthy Lives initiative at McGill University (through New Investigator Supplement to LP); Monique H. Bourgeois Chair in Developmental Disorders and Graham Boeckh Foundation (Douglas Research Centre, McGill University) and salary award from the Le Fonds de recherche du Québec-Santé (FRQS). F. Zaher's research was undertaken thanks in part to a studentship from the Canada First Research Excellence Fund awarded to the Healthy Brains, Healthy Lives initiative at McGill University.

Registration and protocol

This systematic review followed the Preferred Reporting Items for Systematic Review and Meta-analysis (PRISMA) 2020 guidelines. Details of the protocol for this systematic review were registered on PROSPERO (International prospective register of systematic reviews, Ahuja et al., 2023). Two amendments were made from information provided in the protocol. Firstly, the following exclusion criterion was added: 'The study does not report findings relevant to the effect of PADs

on NLP markers'. Secondly, while in the protocol, we stated that we would not conduct a formal risk of bias assessment for the included studies, and we subsequently decided to go ahead with conducting formal risk of bias assessments in order to improve the methodological rigor.

PRISMA systematic review checklist

Attached with manuscript submission.

ORCID iD

Lena Palaniyappan  <https://orcid.org/0000-0003-1640-7182>

Supplemental material

Supplemental material for this article is available online.

References

- Adler CM, Goldberg TE, Malhotra AK, et al. (1998) Effects of ketamine on thought disorder, working memory, and semantic memory in healthy volunteers. *Soc Biol Psychiatry* 43: 81–816.
- Agurto C, Cecchi GA, Norel R, et al. (2020) Detection of acute 3,4-methylenedioxymethamphetamine (MDMA) effects across protocols using automated natural language processing. *Neuropsychopharmacology* 45: 823–832.
- Ahuja S, Zaher F and Palaniyappan L (2023) Effect of psychoactive drugs on natural language processing markers of speech. *PROSPERO*. Epub ahead of print 2023. https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42023448625
- Alexander-Emery S, Cohen LM and Prenskey EH (2005) Linguistic analysis of college aged smokers and never smokers. *J Psychopathol Behav Assess* 27: 11–16.
- Amarel M and Cheek FE (1965) Some effects of LSD-25 on verbal communication. *J Abnormal Psychol* 70: 453–456.
- Baggott MJ, Kirkpatrick MG, Bedi G, et al. (2015) Intimate insight: MDMA changes how people talk about significant others. *J Psychopharmacol* 29: 669–677.
- Bedi G, Cecchi GA, Slezak DF, et al. (2014) A window into the intoxicated mind? Speech as an index of psychoactive drug effects. *Neuropsychopharmacology* 39: 2340–2348.
- Blaabjerg ES, Hemmingsen RPA, Høegh E, et al. (2020) Variability between psychiatrists on domains of the mental status examination. *Nordic J Psychiatry* 74: 287–292.
- Brady KT, Lydiard RB, Malcolm R, et al. (1991) Cocaine-induced psychosis. *J Clin Psychiatry* 12: 509–512.
- Corcoran CM, Mittal VA, Bearden CE, et al. (2020) Language as a biomarker for psychosis: A natural language processing approach. *Schizophr Res* 226: 158–166.
- Covidence (2023) Covidence systematic review software. www.covidence.org. Melbourne, Australia: Veritas Health Innovation. Available at: www.covidence.org (accessed 16 May 2024).
- Covington MA, Riedel WJ, Brown C, et al. (2007) Does ketamine mimic aspects of schizophrenic speech? *J Psychopharmacol* 21: 338–346.
- Cox DJ, Garcia-Romeu A and Johnson MW (2021) Predicting changes in substance use following psychedelic experiences: natural language processing of psychedelic session narratives. *Am J Drug Alcohol Abuse* 47: 444–454.
- Dolder PC, Schmid Y, Steuer AE, et al. (2017) Pharmacokinetics and pharmacodynamics of lysergic acid diethylamide in healthy subjects. *Clin Pharmacokinet* 56: 1219–1230.
- Durback LF, Scharman EJ and Brown BS (1998) Emergency physicians perceptions of drug screens at their own hospitals. *Vet Hum Toxicol* 40: 234–237.
- Earleywine M, Mian MN and De Leo JA (2023) Expectancies for subjective and antidepressant effects in psilocybin users. *J Human Psychol* 1–21.

- Espejo T, Riedel HB, Messingschlager S, et al. (2024) Predictive value and interrater reliability of mental status and mobility assessment in the emergency department. *Clin Med (Lond)* 24: 100027.
- Garcia-Romeu A, Griffiths RR and Johnson MW (2015) Psilocybin-occasioned mystical experiences in the treatment of tobacco addiction. *Curr Drug Abuse Rev* 7: 157–164.
- Griffiths RR, Johnson MW, Carducci MA, et al. (2016) Psilocybin produces substantial and sustained decreases in depression and anxiety in patients with life-threatening cancer: A randomised double-blind trial. *J Psychopharmacol* 30: 1181–1197.
- Honigfeld G (1965) Temporal effects of LSD-25 and epinephrine on verbal behavior. *J Abnormal Psychol* 70: 303–306.
- Jaffe J, Dahlberg CC, Luriai J, et al. (1973) Effects of LSD-25 and dextroamphetamine on speech rhythms in psychotherapy dialogues. *Biological Psychiatry* 6: 93–96.
- Johns A (2001) Psychiatric effects of cannabis. *BJPsych* 178: 116–122.
- Kamilar-Britt P and Bedi G (2015) The prosocial effects of 3,4-methylenedioxyamphetamine (MDMA): Controlled studies in humans and laboratory animals. *Neurosci Biobehav Rev* 57: 433–446.
- Koops S, Brederoo SG, de Boer JN, et al. (2021) Speech as a biomarker for depression. *CNS Neurol Disord Drug Targets* 22: 152–160.
- Krystal JH, Karper LP, Seibyl JP, et al. (1994) Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans. Psychotomimetic, perceptual, cognitive, and neuroendocrine responses. *Arch Gen Psychiatry* 51: 199–214.
- Le Glaz A, Haralambous Y, Kim-Dufor DH, et al. (2021) Machine learning and natural language processing in mental health: Systematic review. *J Med Internet Res* 23: e15708.
- Lee L, Stemple JC, Geiger D, et al. (1999) Effects of environmental tobacco smoke on objective measures of voice production. *Laryngoscope* 109: 1531–1534.
- Margolis A, Rosca P, Kurs R, et al. (2016) Routine drug screening for patients in the emergency department of a state psychiatric hospital: A naturalistic cohort study. *J Dual Diagn* 12: 218–226.
- Marrone GF, Pardo JS, Krauss RM, et al. (2010) Amphetamine analogs methamphetamine and 3,4-methylenedioxyamphetamine (MDMA) differentially affect speech. *Psychopharmacology* 208: 169–177.
- McCulloch DE-W, Grzywacz MZ, Madsen MK, et al. (2022) Psilocybin-induced mystical-type experiences are related to persisting positive effects: A quantitative and qualitative report. *Front Pharmacol* 13: 841648.
- Miles DA (2017) *A taxonomy of research gaps: Identifying and defining the seven research gaps*. In: *Doctoral Student Workshop: Finding Research Gaps-Research Methods and Strategies*, Dallas, Texas, pp. 1–15.
- Mota NB, Copelli M and Ribeiro S (2017) Thought disorder measured as random speech structure classifies negative symptoms and schizophrenia diagnosis 6 months in advance. *NPJ Schizophr* 3: 18.
- Mueller PB and Wilcox JC (1980) Effects of marijuana smoking on vocal pitch and quality. *Ear Nose Throat J* 59:506–509.
- Murray CH and Srinivasa-Desikan B (2022) The altered state of consciousness induced by $\Delta 9$ -THC. *Conscious Cogn* 102: 103357.
- Muthukumaraswamy SD, Forsyth A and Lumley T (2021) Blinding and expectancy confounds in psychedelic randomized controlled trials. *Expert Rev Clin Pharmacol* 14: 1133–1152.
- Nyanchocka L, Tudur-Smith C, Thu VN, et al. (2019) A scoping review describes methods used to identify, prioritize and display gaps in health research. *J Clin Epidemiol* 109: 99–110.
- Palaniyappan L (2021) More than a biomarker: Could language be a bio-social marker of psychosis? *NPJ Schizophr* 7: 42.
- Palaniyappan L, Alonso-Sanchez MF and MacWhinney B (2022) Is collaborative open science possible with speech data in psychiatric disorders? *Schizophr Bull* 48: 963–966.
- PsychonautWiki Contributors (2024) Main Page. Available at: https://psychonautwiki.org/wiki/Psychoactive_substance_index (accessed 10 February 2025).
- Review Manager Web (2020) Review Manager Web (RevMan Web). 1.22.0. The Cochrane Collaboration.
- Ross CA and Leichner P (1988) Residents performance on the mental status examination. *Canad J Psychiatry* 33: 108–111.
- Roth WT, Rosenbloom MJ, Darley CF, et al. (1975) Marijuana effects on TAT form and content. *Psychopharmacologia (Berl.)* 43.: 261–266.
- Salomé F, Boyer P and Fayol M (2000) The effects of psychoactive drugs and neuroleptics on language in normal subjects and schizophrenic patients: A review. *Eur Psychiatry* 15: 461–469.
- Sanz C, Pallavicini C, Carrillo F, et al. (2021) The entropic tongue: Disorganization of natural language under LSD. *Conscious Cogn* 87: 103070.
- Sanz C, Cavanna F, Muller S, et al. (2022) Natural language signatures of psilocybin microdosing. *Psychopharmacology* 239: 2841–2852.
- Shamei A and Bird S (2019) An acoustic analysis of cannabis-intoxicated speech. *Canad Acoust* 47: 108–109.
- Shamei A, Sullivan PR, Liu Y, et al. (2021) Automated detection of cannabis-intoxication from speech. *Canad Acoust* 49.
- Stitzer ML, Griffiths RR and Liebson I (1978) Effects of d-Amphetamine on speaking in isolated humans. *Pharmacol Biochem Behav* 9: 57–63.
- Tagliazucchi E (2022) Language as a window into the altered state of consciousness elicited by psychedelic drugs. *Front Pharmacol* 13: 812227.
- Vizeli P, Straumann I, Holze F, et al. (2021) Genetic influence of CYP2D6 on pharmacokinetics and acute subjective effects of LSD in a pooled analysis. *Scient Rep* 11: 10851.
- Vogel AP, Pearson-Dennett V, Magee M, et al. (2021) Adults with a history of recreational cannabis use have altered speech production. *Drug Alcohol Depend* 227: 108963.
- Wardle MC and De Wit H (2014) MDMA alters emotional processing and facilitates positive social interaction. *Psychopharmacology* 231: 4219–4229.
- Wardle MC, Garner MJ, Munafò MR, et al. (2012) Amphetamine as a social drug: Effects of d-amphetamine on social processing and behavior. *Psychopharmacology* 223: 199–210.
- Weil AT and Zinberg NE (1969) Acute effects of marijuana on speech. *Nature* 222: 434–437.
- Wießner I, Falchi M, Daldegan-Bueno D, et al. (2023) LSD and language: Decreased structural connectivity, increased semantic similarity, changed vocabulary in healthy individuals. *Eur Neuropsychopharmacol* 68: 89–104.
- Zeidenberg P, Clark WC, Jaffe J, et al. (1973) Effect of oral administration of $\Delta 9$ tetrahydrocannabinol on memory, speech, and perception of thermal stimulation: Results with four normal human volunteer subjects. Preliminary report. *Comprehens Psychiatry* 14: 549–556.