JLAR 4 (2026)
10.15460/jlar.2026.4.1.0005

An Italian corpus of clinical speech data: *Corpus di Parlato Patologico Italiano in Pazienti Affetti da Schizofrenia Farmacoresistente* (CIPP-TRS)

Laura Tagliaferro[1]*

[1]University of Naples Federico II, Naples, Italy

Abstract. The *Corpus di Parlato Patologico Italiano in Pazienti Affetti da Schizofrenia Farmacoresistente* (CIPP-TRS) is a spoken language corpus consisting of semi-spontaneous speech samples collected from adults diagnosed with treatment-resistant schizophrenia, non-treatment-resistant schizophrenia, and healthy controls. The corpus includes approximately 15 minutes of recorded interviews per participant, elicited through semi-structured interviews, for a total of over seven hours of annotated audio data. All speakers are adult native speakers of Italian, with recordings reflecting a Neapolitan regional variety. The dataset is enriched with detailed sociolinguistic metadata, including age, years of schooling, and clinical status, enabling investigations into language variation across adulthood and young old age. Orthographic transcriptions are aligned with audio and annotated in XML-TEI format, allowing systematic analysis of disfluencies, pauses, peri- and paralinguistic phenomena, and other features of spoken interaction. The CIPP-TRS corpus is designed to support linguistic, phonetic, and interactional research, with particular relevance for studies on aging, clinical populations, and spoken language variability. In line with FAIR data principles (Wilkinson et al. 2016), the corpus is structured to ensure transparent metadata documentation, standardized encoding (XML-TEI), and clear licensing conditions to facilitate discoverability, accessibility, interoperability across platforms, and reuse in future research.

Keywords. spoken corpora; Italian; aging; clinical linguistics; disfluency

A peer-reviewed contribution to the *Journal of Language and Aging Research* (JLAR).
Submitted: 2025-10-10
Accepted: 2026-02-24
Published: 2026-06-29
DOI: 10.15460/jlar.2026.4.1.0005

*. Corresponding author, laura.tagliaferro@unina.it

Please cite as Tagliaferro, Laura. 2026. "An Italian corpus of clinical speech data: *Corpus di Parlato Patologico Italiano in Pazienti Affetti da Schizofrenia Farmacoresistente* (CIPP-TRS)." *Journal of Language and Aging Research* 4(1): 49–60. 10.15460/JLAR.2026.4.1.0005.

JLAR 4 (2026)
10.15460/jlar.2026.4.1.0005

1 Introduction

While linguistic change is a recognized feature of the natural aging process, clinical conditions can alter the trajectory of language across adulthood (Shafto and Tyler 2014). Deviations from age-related norms may affect fluency, lexical access, turn-taking, and the temporal organization of speech, offering valuable insights into the interaction between health status and spoken language production (Dovetto and Marra 2024). In this regard, clinical speech data represent an important resource for research on language variation and aging.

This article introduces the CIPP-TRS (*Corpus di Parlato Patologico Italiano in Pazienti Affetti da Schizofrenia Farmacoresistente*), a spoken language corpus developed to investigate linguistic variation at the intersection of adulthood, clinical pathology, and aging-related processes. The corpus was designed and constructed over the course of a three-year doctoral program (2022–2025) at the University of Naples Federico II and consists of semi-spontaneous speech samples collected from adult speakers diagnosed with treatment-resistant schizophrenia (TRS),[2] non-treatment-resistant schizophrenia (non-TRS), and healthy controls (HC).

The CIPP-TRS corpus is conceived primarily as a data resource. It provides audio recordings, aligned transcriptions, and structured annotations that enable qualitative and quantitative analyses of spoken Italian with the presence of occasional Neapolitan dialect features and inflections, with particular attention to disfluencies, pauses, lexical diversity, and peri- and paralinguistic phenomena.[3] All participants fall within a broad adult age range (18–65 years), enabling the corpus to document spoken language production across adulthood in clinical and non-clinical populations. In clinical groups, linguistic behavior is interpreted in relation to the interaction between age, clinical condition, and age of symptom onset, as earlier onset of schizophrenia has been consistently associated with more pronounced cognitive impairment, which in turn affects language production (Rajji, Ismail, and Mulsant 2009; de Boer et al. 2020; de Bartolomeis et al. 2022).

From a methodological standpoint, the corpus adopts recording conditions, transcription conventions, and annotation practices aligned with established standards in Italian spontaneous speech corpora (see Section 4.1), enabling interoperability and cross-corpus comparison.

The inclusion of detailed sociolinguistic metadata such as age, sex, and years of schooling further enhances the reuse potential and extensibility of the dataset for studies on spoken language, interaction, and aging. More broadly, the corpus design adheres to the FAIR principles (Findable, Accessible, Interoperable, Reusable; Wilkinson et al. 2016), particularly with respect to standardized encoding, rich metadata

2. Resistance to treatment is defined as the persistence of symptoms despite at least two trials of antipsychotic medications at adequate doses and durations, with documented adherence (Potkin et al. 2020). The TRS subgroup is associated with greater severity of both positive and negative symptoms, profound functional impairment, and a substantial burden on patients as well as caregivers (Kane et al. 2019).

3. Following De Mauro (2008)'s distinction, peri-linguistic phenomena encompass weakly integrated or non-integrated elements of the language system that nevertheless convey prosodic cues, such as interjections, filled pauses, and disfluencies. Paralinguistic phenomena, by contrast, accompany linguistic production but belong to other semiotic domains, including gestures, prosodic modalities (e. g., whispering, screaming), proxemics, and speech rate.

JLAR 4 (2026)
10.15460/jlar.2026.4.1.0005

documentation, and conditions that facilitate long-term discoverability and responsible reuse.

2 Methods

Semi-spontaneous speech samples were collected from 25 patients diagnosed with TRS, 25 non-TRS patients, and 25 HC. The patient groups were being treated at the Department of Neuroscience, Reproductive Sciences, and Odontostomatology, Integrated Department of Clinical care, Section on Psychiatry and Clinical Unit of Psychiatry and Psychology of the Azienda Ospedaliera Universitaria Federico II of Naples. The TRS and non-TRS groups were comparable with respect to key sociolinguistic variables, including diatopic background, age, sex, and years of schooling. Based on this distribution, a single healthy control group was constructed to match the overall sociolinguistic profile of the two clinical groups at the group level. This procedure ensured comparability across age, educational background, and regional linguistic variety while maintaining balanced sample sizes.

All participants signed an informed consent for the protection of personal data and were assessed according to the group to which they belong. The patient group was assessed with the Positive and Negative Syndrome Scale (PANSS) and the Brief Assessment Cognitive Schizophrenia Scale (BACS). The HC group was assessed using the Mini International Neuropsychiatric Interview (MINI plus) of DSM-5. To date, the annotated and revised material consists of 10 TRS patients, 10 non-TRS patients and 10 HC whose information is reported in Table 1.

The development of the CIPP-TRS involved creating selection criteria to ensure data accuracy and comparability. Participants were required to be between 18 and 65 years old and to provide valid written informed consent. This group included individuals diagnosed with schizophrenia who were either unresponsive (TRS) or responsive to treatment (non-TRS). To minimize confounding factors, we excluded patients with other psychiatric conditions such as bipolar disorder or major depression, as these could interfere with the expression of negative symptoms. Additionally, patients with organic pathologies affecting language production or articulation, intellectual disabilities, severe systemic illnesses, psychiatric conditions secondary to medical issues, or substance abuse were not included.

The HC group followed similar inclusion criteria in terms of age and consent, while also excluding individuals with any psychopathological conditions. Further exclusion criteria for controls included the presence of bipolar disorder or depression, systemic or neurological conditions affecting speech, disabilities, and the influence of psychoactive substances.

Speech data were collected through a 15-minute semi-structured interview for each participant. The interviews were recorded at the hospital using a Zoom H4nPro device (44 kHz sampling rate using 16 bit). The recordings were analyzed using PRAAT (Boersma and Weenink 2024). The dataset includes audio files (.wav), aligned TextGrid files (.txt), and orthographic transcriptions (.doc). An example of TextGrid is provided in Figure 1.

To enable meaningful comparisons, the transcription method adopted aligned with analogous Italian corpora of neurotypical spontaneous speech such as CLIPS (Corpora and Lexicon of Written and Spoken Italian; Albano Leoni, Sobrero, and Paoloni 2007) and non-neurotypical speech, i. e., CIPPS (Italian Corpus of Patients Affected with

JLAR 4 (2026)
10.15460/jlar.2026.4.1.0005

Schizophrenia; Savy 2006; Dovetto and Gemelli 2013) and FRA (Language Corpora for the Assessment of Spontaneous Speech in Alzheimer's Disease and Mild Cognitive Impairment) (Dovetto and Marra 2020).

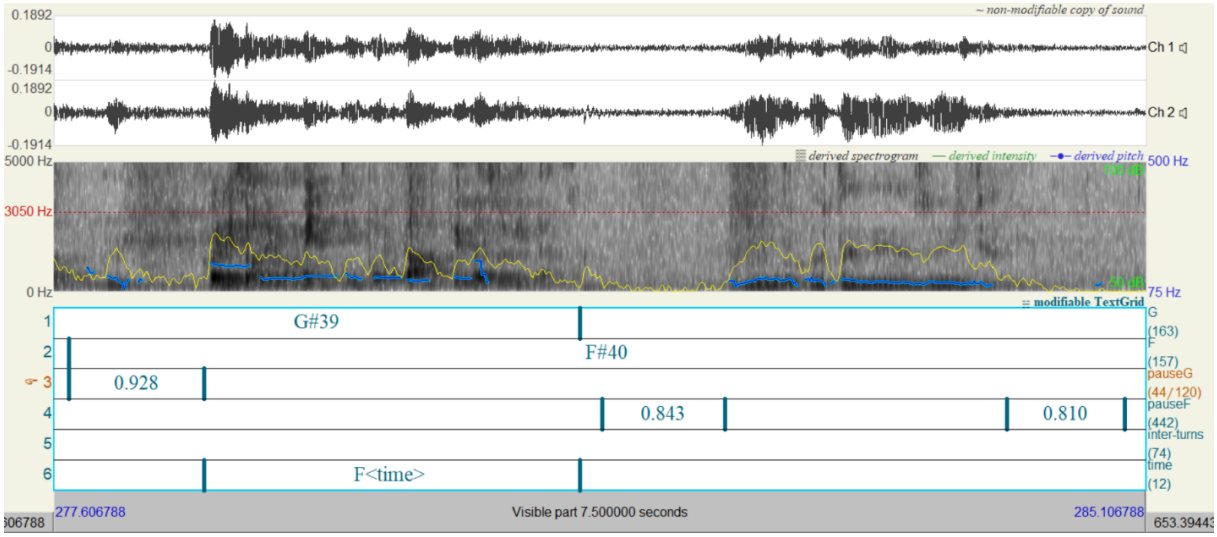

Figure 1: Illustration of the PRAAT annotation stored as a TextGrid. The first two tiers are dedicated to the Giver's (G, doctor) and Follower's (F, subject) turns. The two other subsequent tiers are dedicated to the silent pauses produced by each participant. The tier dedicated to inter-turn pauses represents the silences that were not assigned to either one of the participants. The last tier (optional) shows which one of the participants is waiting for the other to finish their turn in order to start talking again.

3 Corpus description

The majority of patients in both groups were recipients of a disability pension. At the time of the recordings, approximately 95% of the sample was unemployed, considering both part-time and full-time occupations. Nearly all patients had attempted at least one job during their lifetime, most commonly in family businesses or seasonal employment. Table 1 provides the sociolinguistic variables of the corpus with mean ages of 44 years old for TR and HC groups, and 41 years old for the non-TR group.

4 Methods

During the interviews, which lasted approximately 15 minutes, a subset of 61 questions previously employed by de Boer et al. (2020) were administered. The questions, presented in randomized order, address a range of topics intended to elicit diverse linguistic responses until the time target is reached.

JLAR 4 (2026)
10.15460/jlar.2026.4.1.0005

Table 1: Sociolinguistic Variables of the first 10 participants in each group. All participants come from Campania, Italy.

ID	Sex	Age of Onset	Age	Years of Schooling
TRSA	M	23	40	8
TRSB	M	16	36	5
TRSC	M	27	39	17
TRSD	M	27	64	13
TRSE	F	23	39	13
TRSF	M	27	51	13
TRSG	M	14	37	13
TRSH	F	16	43	13
TRSI	F	22	47	13
TRSJ	M	24	43	13
non-TRSA	M	28	40	16
non-TRSB	M	22	62	8
non-TRSC	M	20	26	13
non-TRSD	M	27	37	13
non-TRSE	M	27	62	17
non-TRSF	M	16	30	8
non-TRSG	M	19	24	13
non-TRSH	M	11	28	8
non-TRSI	M	20	45	5
non-TRSJ	F	22	55	13
HCA	M	–	62	8
HCB	M	–	48	13
HCC	F	–	44	16
HCD	F	–	38	16
HCE	M	–	20	13
HCF	M	–	46	13
HCG	M	–	53	8
HCH	M	–	45	13
HCI	M	–	60	8
HCJ	M	–	20	13

The topics covered included:

- Hypothetical or imaginative scenarios, e. g.: *What would you do if you won a million euros?*
- Experiences from the past, such as childhood and adolescence, e. g.: *Have you ever been in trouble as a child? Did your parents get angry?*
- Aspects of the present, including vacation experiences, leisure activities, family life, hobbies, and work, e. g.: *Do you go on vacation? What is your favorite vacation destination? Why?*

Generally, a single interview contained about twenty different questions with further input for detail and/or aid that the doctor provided in reformulating, explaining and adapting questions when necessary. Given the variation in interview length, the average proportion of speaking time estimates indicate that, across groups, interviewees spoke for more than three-fifths of the total interview duration, demonstrating

JLAR 4 (2026)
10.15460/jlar.2026.4.1.0005

that they actively responded to questions and maintained the communicative floor as shown in Table 2.

Table 2: Recordings' duration, token production and lexical diversity (TTR). The average percentage of time speaking was calculated considering the interviewees' dialogic turns including pauses.

	Length	Tokens	TTR	Average percentage of time speaking
TRS group	2 hours 37 minutes	13576	0.357	65.53%
non-TRS group	2 hours 42 minutes	13644	0.344	71.09%
HC group	2 hours 12 minutes	13499	0.343	66.74%

4.1 Transcription and annotation in XML

The orthographic transcription, which preceded conversion into XML format, was carried out according to the CLIPS method, whereby every attempt to speak was treated as a turn, and each turn was defined as a sequence marked by semantic–pragmatic coherence within the speaker's output (Savy 2006). Each transcription is accompanied by an informative box, which contains metadata on the collected material, the speakers, and general information about the recording. The title of each transcription and its associated audio file encodes the main identifying information about the speaker: for example, DgTRSA01N refers to the first (01) dialogical interview (Dg) of patient A with treatment-resistant schizophrenia (TRS), recorded in Naples (N). Each transcription begins with functional tables that summarize information about the speakers, the recording, and the transcription process. Examples are shown in Tables 3–6.[4]

Table 3: Text Information (modified from Savy 2006).

Code	Description	Example
MAT	Corpus type	TRS = treatment-resistant schizophrenia; non-TRS = non treatment-resistant schizophrenia etc.
MAP	Speaker identification code	A, B, C, D …
Ndl	Dialogue number	01
REG	Regional variety	N = Neapolitan

Table 4: Speaker details (modified from Savy 2006).

Code	Description	Example
INp1	First speaker (initials of name and surname, gender, age, place of birth)	V. M. S., F, 28 years old, Palermo (PA)
INp2	Second speaker (same format as INp1)	–
INp3	Occasional third speaker (same format as INp1)	–

4. The information on the first speaker in Table 4 coincides with the doctor's data, because the doctor initiates the interview.

JLAR 4 (2026)
10.15460/jlar.2026.4.1.0005

Table 5: Recording information (modified from Savy 2006).

Code	Description	Example
TYP	Recording type and format	Audio; WAV file; Zoom H4n Pro
LOC	Location of recording	–
DAT	Date of recording	–
DUR	Duration	17 min. 20 sec.
CON	Recording condition	Bad, good, excellent

Table 6: Transcription's information (modified from Savy 2006).

Code	Description	Example
DAT	Month of last revision	–
CMT	Transcriber's comments (e. g., dialectal inflections, additional details)	–
Nst	Number of turns	–

The orthographic transcriptions follow the conventions outlined by Savy (2006), as shown in Example (1).

(1) G#13: <NOISE> <inspiration> Hai mai indossato un capo di abbigliamento particolarmente strano ?
F#14: <pause dur="0,205 s"/> <tongue click> <eh> Io sì<ii> quando mi vesto<oo> <ehm> <pause dur="0,642 s"/> cioè non do importanza agli {<whispering> abiti non<nn> </whispering>} / <pause dur="0,495 s"/> <inspiration> <beh> io faccio parte dicerie artisti io sono / dice l' ariete è 'n' arte , è un po' così l' ariete {<whispering> po<oo>' </whispering>} <pause dur="0,815 s"/> {<giggle> lunatico </giggle>} <pause dur="1,399 s"/> <vocal> <exhale>
<note> The conjunction *cioè* is pronounced contracted in the form of [tʃɛ] </note>
(DgTRSI01N)

However, in order to facilitate data comparison and enable systematic extraction of linguistic information, the dataset was transformed and adapted into XML-TEI format in order to facilitate the isolation of peri- and paralinguistic phenomena. In particular, peri-linguistic phenomena include:

- Repair strategies such as false starts, generally following truncated words and word repetitions.
- Primary interjections (e. g., <ah>, <beh>).
- Filled pauses with vocalization or nasalization (e. g., <eeh>, <ehm>).
- Empty pauses (e. g., <pause dur="0,504 s"/>).
- Sound elongations (vowel/consonant lengthening).
- Underscore to indicate interruptions within lexical items (e. g., tem_po).

JLAR 4 (2026)
10.15460/jlar.2026.4.1.0005

Paralinguistic phenomena include both punctual elements (e. g., tongue clicks, lip smacking, throat clearing, laughter, coughing, breathing) and extended elements[5] annotated with opening and closing tags (e. g., <screaming>…</screaming>). The set of tags used is given in Table 7.

Table 7: Tagset for orthographic transcription and xml format.

Orthographic Transcription	XML format	Phenomenon
<name> <surname>	<name> <surname>	Proper Nouns and Surnames[6]
/ (slash)	<FS>	False starts
+ (plus sign)	<TR>	Fragmented words
_ (underscore)	_	Interruptions within the lexical element
{<repetition …</repetition>}	<RP> …</RP>	Involuntary stuttering of same words or parts.
<pause dur="0,476s"/>	<P 0,476/>	Silent pause
<eeh> <ehm>	<PE eeh> <PE ehm>	Filled pause with vocalization or nasalization.
<vv> (v = vowel) <cc> (c = consonant)	<vv> <cc>	Elongation of vowel or consonant at word initial, middle, or final position.
<inspiration> <cough> <exhale> <giggle> <laugh> <lip smacking> <tongue click> <throat clearing>	<PA inspiration> <PA cough> <PA laugh> <PA lip smacking> <PA tongue click> <PA throat clearing>	Vocal non-verbal phenomena produced by the speakers. The phenomenon is inserted in the exact collocation before any eventual punctuation.
<vocal>	<PE vocal>	Other phenomena aside from the labels mentioned.
<ah> <ah!> <eh> <eh!> <mah> <boh> <beh> <mh> <mhmh> <uh> <neh> <oh!> etc.	<PE ah> <PE ah!> <PE eh> <PE eh!> <PE mah> <PE boh> <PE mh> <PE mhmh> etc.	Primary interjections indicating surprise, awe, satisfaction etc.
<NOISE>	<NOISE>	Unidentified noises
<unclear>	<unclear>	Unintelligible words or strings of words.
<MUSIC>	<MUSIC>	Music or ringtones.
{<i.talkers…</i.talkers>}	<i.talkers> …</i.talkers>	Background unidentified voices.
<note> …<note>	<note>…<note>	Transcriber's comments.

4.2 Availability

The CIPP-TRS corpus is intended as a reusable data resource. The corpus has been designed in accordance with FAIR data principles (Wilkinson et al. 2016). The dataset will be deposited on the LISA platform by the end of January 2027, where metadata and documentation describing the structure of the corpus will be publicly available. The data will be distributed in non-proprietary formats (TextGrid, UTF-8 encoded text files)

5. Extended elements comprise noises, whispering, laughter, giggling, dialectal features, music, and background voices.
6. Proper names and surnames have been anonymized to ensure the privacy of the project participants.

JLAR 4 (2026)
10.15460/jlar.2026.4.1.0005

to ensure long-term accessibility. Annotation conventions follow established standards in spoken corpus linguistics, allowing interoperability with existing corpora and enabling additional annotation layers without altering the original files. Pseudonymized metadata and detailed documentation support reproducibility and facilitate reuse for linguistic, clinical, and computational research purposes.

4.3 Reuse potential

The CIPP-TRS corpus is a reusable resource for the analysis of spoken language in both clinical and non-clinical adult populations. The orthographic transcriptions allow researchers to perform quantitative and qualitative analyses on lexical variation, disfluency patterns, and turn organization. The annotation structure is compatible with standard corpus linguistics formats, wherein additional layers of markup, e. g., syntactic and pragmatic, can be added. The presence of detailed sociolinguistic metadata further supports studies across age groups and facilitates cross-linguistic investigations.

4.4 Connections to language and aging research

The CIPP-TRS corpus occupies a critical interface between clinical linguistics and lifespan language research. It enables the systematic investigation of disorder-specific linguistic disruption while simultaneously capturing the dynamic interplay between psychopathology and age-related language change across adulthood.

5 Ethical considerations

The CIPP-TRS corpus was developed within the framework of doctoral scholarships funded in Italy in 2022 under the National Operational Programme “Research and Innovation” (PON Ricerca e Innovazione 2014–2020), with additional support from REACT-EU funds.

All corpus data were pseudonymized by removing personally identifiable information from transcription and annotation files, and de-pseudonymization is prohibited. Access to the CIPP-TRS corpus does not include sensitive data, and researchers requesting access must agree to use the material exclusively for scientific purposes and not distribute it to third parties.

The project received ethical approval as part of the doctoral program and adheres to national and institutional guidelines for research with clinical populations. All participants provided their informed written consent to study participation and data processing, ensuring confidentiality. Data will remain pseudonymized and may only be used for scientific purposes, in compliance with DL 196/2003 on personal data protection.

6 Example

The following example analysis illustrates how the CIPP-TRS corpus can be exploited to investigate temporal and fluency-related properties of spoken language. By combining aligned audio data with detailed orthographic and XML-TEI annotations, the corpus enables the systematic extraction and quantification of phenomena such as

JLAR 4 (2026)
10.15460/jlar.2026.4.1.0005

silent pauses, filled pauses, and other disfluency markers. Measures such as pause-to-word ratio demonstrate how the dataset can support comparative analyses across speaker groups and accommodates phonetic, interactional, and discourse-oriented approaches.

In the XML excerpt (2), peri- and paralinguistic elements are marked within spoken language transcripts to allow their isolation and study in terms of numerosity and positioning.

(2)

F#42: <P 0,550/> <PA exhale> <PA inspiration> **Non lo so proprio** <P 0,293/> **come te lo dico** <PE vocal> <dialect> 'n ghe senz' </dialect> **stra**<TR> <FS><PE eeh> **non lo so** <NOISE> **che**<ee> <PA inspiration> **posso dire** <PA inspiration>
(DgTRSA01N)

Silent pauses can be investigated in relation to fluency and speech pacing through the automatic extraction of pause intervals from the TextGrid annotations. The ratio between pause duration and number of words produced serves as an illustrative metric for comparing the continuity of speech flow across groups (Table 8).

Table 8: Average pause to word ratio and standard deviation (S. D.). This measure was calculated considering the number of silent intervals within the speaker's linguistic production.

	TRS	non-TRS	HC
Pause/all words ratio	0.13	0.11	0.08
S. D.	0.060	0.054	0.037

Acknowledgements

I am grateful to my advisors, Professor Francesca Maria Dovetto and Professor Andrea de Bartolomeis, for their guidance, critical insights, and continuous support throughout this research. Their supervision was instrumental in shaping the methodological and theoretical dimensions of the study.

I also wish to acknowledge the contribution of Dr. Viviana M. Saia, Dr. Giancarlo Spennato, and Dr. Luigi Franzese, whose collaboration in the testing phase was essential for the successful collection of clinical data.

Finally, I thank my colleague and friend Sundra Sorrentino for her careful linguistic revisions and for providing valuable feedback on the transcription process.

Funding

The construction of this corpus was part of a PhD program funded jointly by the University of Naples Federico II and the European Union through FSE REACT-EU initiative IV.4.

Conflict of interest

The author has no conflicts of interest to declare.

References

Albano Leoni, Federico, Alberto A. Sobrero, and Andrea Paoloni. 2007. "Corpora e lessici di italiano parlato e scritta CLIPS." *Bolletino di Italianistica,* 121–148.

Boersma, Paul, and David Weenink. 2024. *Praat: Doing phonetics by computer.* Computer Program, Version 6.4.17. https://praat.org/.

de Bartolomeis, Andrea, Luigi Vellucci, Antonio Barone, Michele Manchia, Vincenzo De Luca, Felice Iasevoli, and Christoph U. Correll. 2022. "Clozapine's multiple cellular mechanisms: What do we know after more than fifty years? A systematic review and critical assessment of translational mechanisms relevant for innovative strategies in treatment-resistant schizophrenia." *Pharmacology & Therapeutics* 236:108236. https://doi.org/10.1016/j.pharmthera.2022.108236.

de Boer, Janna N., Alban E. Voppel, Sanne G. Brederoo, Frank N.K. Wijnen, and Iris E.C. Sommer. 2020. "Language disturbances in schizophrenia: The relation with antipsychotic medication." *NPJ Schizophrenia* 6 (1): 24. https://doi.org/10.1038/s41537-020-00114-3.

De Mauro, Tullio. 2008. *Lezioni di linguistica teorica.* Rome: Laterza.

Dovetto, Francesca M., and Monica Gemelli. 2013. *Il parlar matto: Schizofrenia tra fenomenologia e linguistica: Il corpus CIPPS.* 2nd ed. Rome: Aracne.

Dovetto, Francesca M., and Francesca Marra. 2020. *Language corpora for the assessment of spontaneous speech in Alzheimer's Disease and Mild Cognitive Impairment.* FRA Corpus.

———. 2024. "The FRA Corpus from the DISease and AGEing Project." *Journal of Language and Aging Research* 2 (2): 64–78. https://doi.org/10.15460/jlar.2024.2.2.1522.

Kane, John M., Ofer Agid, Majorie L. Baldwin, Oliver Howes, Jean-Pierre Lindenmayer, Stephen R. Marder, Mark Olfson, Steven G. Potkin, and Christoph U. Correll. 2019. "Clinical guidance on the identification and management of treatment-resistant schizophrenia." *Journal of Clinical Psychiatry* 80 (2). https://doi.org/10.4088/JCP.18com12123.

Potkin, Steven G., John M. Kane, Christoph U. Correll, Jean-Pierre Lindenmayer, Ofer Agid, Stephen R. Marder, Mark Olfson, and Oliver D. Howes. 2020. "The neurobiology of treatment-resistant schizophrenia: Paths to antipsychotic resistance and a roadmap for future research." *NPJ Schizophrenia* 6 (1): 1. https://doi.org/10.1038/s41537-020-0090-z.

Rajji, Tarek K., Zahinoor Ismail, and Benoit H. Mulsant. 2009. "Age at onset and cognition in schizophrenia: Meta-analysis." *British Journal of Psychiatry: The Journal of Mental Science* 195 (4): 286–293. https://doi.org/10.1192/bjp.bp.108.060723.

JLAR 4 (2026)
10.15460/jlar.2026.4.1.0005

Savy, Renata. 2006. "Specifiche per la trascrizione ortografica annotata dei testi raccolti." In *Italiano parlato: Analisi di un dialogo*, edited by F. Albano Leoni and R. Giordano, 1–37. Naples: Liguori.

Shafto, Meredith A., and Lorraine K. Tyler. 2014. "Language in the aging brain: The network dynamics of cognitive decline and preservation." *Science* 346 (6209): 583–587. https://doi.org/10.1126/science.1254404.

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. 2016. "The FAIR Guiding Principles for scientific data management and stewardship." *Scientific Data* 3 (1): Article 160018. https://doi.org/10.1038/sdata.2016.18.