EDICIONES COMPLUTENSE

Revista de Investigación en Logopedia
e-ISSN: 2174-5218

ESTUDIOS

Exploring language in schizophrenia: The use of low frequency words as a linguistic marker schizophrenia symptoms

Alfonso Martínez Cano
Faculty of Health Sciences, Department of Medical Sciences. University of Castilla-La Mancha, Talavera de la Reina, Spain

José Luis Santos
Hospital of Virgen de la Luz, Department of Mental Health, Cuenca, Castilla-La Mancha, Spain.

Manuela Martínez-Lorca (corresponding author)
Faculty of Health Sciences, Department of Psychology. University of Castilla-La Mancha, Talavera de la Reina, Spain

Begoña Polonio-López
Faculty of Health Sciences, Technological Innovation Applied to Health Research Group. University of Castilla-La Mancha, Talavera de la Reina, Spain

https://dx.doi.org/10.5209/rlog.103749 Received June 2nd, 2025; First revision September 9th 2025; Accepted 7th January 2026

Abstract: Our study aims to assess lexical uniqueness in Spanish-speaking individuals with schizophrenia, as well as analyze the peculiarity in word usage concerning symptom predominance. Furthermore, we aim to investigate whether word associations could elucidate the usage of uncommon terms by this population. The sample consists of 50 people diagnosed with schizophrenia (56% with a predominance of positive symptoms and 44% with negative symptoms). The semantic and phonological verbal fluency test was administered. Through the ESPAL database, the frequencies of the generated words were analyzed. Just as the associations of the words were studied through an analysis of blinded judges. People with schizophrenia use a greater number of low-frequency words, especially in positive symptoms in both phonological verbal fluency and semantic verbal fluency, the differences being statistically significant. We observed a greater association of words depending on the sound in the phonological fluency tests while in semantic fluency a greater semantic association was observed both with statistical significance. Our results support the validity of the method for assessing word distinctiveness by people with schizophrenia, especially when they present positive symptoms. The use of rare words could be used as an objective measure for the distinction of symptomatology in schizophrenia.
Keywords: Language; Low-frequency words; Negative symptomatology; Positive symptomology; Schizophrenia.

ES Explorando el lenguaje en la esquizofrenia: el uso de palabras de baja frecuencia como marcador lingüístico de los síntomas de la esquizofrenia

Resumen: Nuestro estudio busca evaluar la singularidad léxica en personas españolas con esquizofrenia, así como analizar la peculiaridad en el uso de palabras en relación con el predominio de síntomas. Además, buscamos investigar si las asociaciones de palabras podrían dilucidar el uso de términos poco comunes en esta población. La muestra está compuesta por 50 personas con diagnóstico de esquizofrenia (56% con predominio de síntomas positivos y 44% con síntomas negativos). Se administró la prueba de fluidez verbal semántica y fonológica. A través de la base de datos ESPAL, se analizaron las frecuencias de las palabras generadas. Asimismo, se estudiaron las asociaciones de las palabras mediante un análisis interjueces. Las personas con esquizofrenia utilizan un mayor número de palabras de baja frecuencia, especialmente en los síntomas positivos, tanto en la fluidez verbal fonológica como en la fluidez verbal semántica, siendo las diferencias estadísticamente significativas. Se observó una mayor asociación de palabras en función del sonido en las pruebas de fluidez fonológica, mientras que en la fluidez verbal semántica se observó una mayor asociación semántica, ambas con significación estadística. Nuestros resultados respaldan la validez del método para evaluar la distinción de palabras en personas con esquizofrenia, especialmente cuando

presentan síntomas positivos. El uso de palabras poco frecuentes podría utilizarse como una medida objetiva para la distinción de la sintomatología en la esquizofrenia.
Palabras clave: Esquizofrenia; Lenguaje; Palabras de baja frecuencia; Sintomatología negativa; Sintomatología positiva.

Cómo citar: Martínez-Cano, A., Santos, J.L., Martínez-Lorca, M., & Polonio-López, B. (2026). Exploring language in schizophrenia: The use of low frequency words as a linguistic marker schizophrenia symptoms. *Revista de Investigación en Logopedia* 16(1), e103749, https://dx.doi.org/10.5209/rlog.103749

Introduction

Schizophrenia has long been characterized by alterations of language, which constitute one of its primary clinical and diagnostic features (American Psychiatric Association, 2013). Previous studies evidence syntactic deficits (e.g. inappropriate and disordered sentences) (Jo et al., 2023), pragmatic deficits (e.g. ignoring the interlocutor or misinterpreting interactions) (Chakrabarty et al., 2023; Martínez et al., 2018), voice deficits (e.g. abnormal pauses, intonation, and fundamental frequency) (Martínez et al., 2018; Martinez-Cano et al., 2024; Oomen et al., 2022; Parola et al., 2020), as well as problems in verbal fluency, especially in terms of lexical access (Agurto et al., 2023; Allen et al., 1993; Gourovitch et al., 1996; Spence et al., 2000). Recent studies conducted in Spanish-speaking populations have reported difficulties in these areas and have even proposed their use as potential biomarkers or as tools to monitor disease progression (Ayuso et al., 2023; Figueroa et al., 2019; 2022; Martínez-Cano, et al., 2025a, 2025b).

The unusual or atypical use of vocabulary by individuals with schizophrenia is a language anomaly that has been extensively reported in the literature(Andreasen, 1986; Coleman et al., 1996), and various scales have been implemented as a diagnostic measure of the disease (Liddle et al., 2002). Authors have even posited its use as a predictor of the disorder (Baskak et al., 2008). Despite its being a common clinical observation in schizophrenia and there existing theories that explain its role in thought disorder (Crow, 1980; Liddle et al., 2002; Salavera et al., 2013), studies using quantitative methods to assess this anomaly are limited, arguably as a result of the inherent subjectivity of categorizing words as uncommon.

Linguist George Zipf (1945) established Zipf's Law (Alvarado & Arango, 2011), which involves analyzing the frequency of occurrence of words in an extensive corpus, ordering them according to their frequency of appearance. According to Erar (2001) and other researchers, Zipf's Law can be interpreted as the probability that a specific word appears with a certain frequency in a text. Extending this concept to speech, we can consider it as the probability of using one word instead of another. Extensive databases, such as ESPAL (Duchon et al., 2013), have compiled data on the frequency of use of words, ranking them from highest to lowest usage.

There have been extensive studies that have taken data on word frequencies into account. For example, in research on Alzheimer's, better recall of frequent words compared to infrequent words has been observed (Balota et al., 2002; Wilson et al., 1983).

Likewise, it has been documented that children tend to respond more quickly in contexts where frequent words are used (Pearson & Studt, 1975), and that these words are acquired and learned earlier (Morrison et al., 1992). Additionally, people with aphasia and language development disorder have been found to process sight words more effectively (Bose et al., 2007; Huck et al., 2017; Perea et al., 2005).

Although substantial work has been conducted in Spanish-speaking contexts on verbal fluency in first-episode psychosis (Figueroa et al.,2019) and on receptive vocabulary in chronic schizophrenia (Ayuso et al., 2023), few studies have examined the frequency of word use among Spanish populations with schizophrenia. In English-speaking contexts, however, some studies report that individuals with schizophrenia tend to use low-frequency words (Baskak et al., 2008; Liddle et al., 2002; Salavera et al., 2013).

Studies conducted in English, such as that by Baskak (2008), have analyzed the uniqueness of words used by people with schizophrenia. These authors ranked the peculiarity of each word uttered according to its frequency of use as evaluated by a group of blinded raters, finding that individuals with schizophrenia use fewer common words compared to those produced by the control group and the patients' healthy siblings. Although this method is reliable, it is still costly and has an inherently subjective component.

These studies (Baskak, 2008) have evaluated verbal fluency (semantic and phonological), and word production should be analyzed according to time intervals. Previous studies with both children and adults have evidenced a change in task performance and words produced as function of the time elapsed (Crowe, 1998; Hurks et al., 2006). However, this parameter has not always been considered in the study of schizophrenia (Barattieri di San Pietro et al., 2023). Besides these authors have posited that schizophrenia patients may generate a larger number of low-frequency words due to an excessive use of associations based on the sounds of words rather than their meaning (Baskak et al., 2008; Salavera et al., 2013).

Therefore, the objectives of this work are: (**1**) Establish an objective, reliable and rapid method for assessing word peculiarity in Spanish-speaking persons with schizophrenia (**2**) Determine whether the occurrence of unusual or atypical word use differs depending on positive or negative symptomatology of schizophrenia. (**3**) Analyzed the associations between the words produced by individuals with schizophrenia.

Method

Participants

This was a cross-sectional and observational with a sample of 50 adult participants diagnosed with schizophrenia, divided into two groups, one with positive symptoms (n=28) and the other with negative symptoms (n=22). Table 1 shows the sociodemographic and clinical variables of the sample.

The participants' diagnoses were confirmed by a team of psychiatrists, who determined symptomatology using the clinical interview for the Positive and Negative Schizophrenia Syndrome Scale (PANSS) (Kay et al., 1987; Peralta & Cuesta, 1994). To distinguish between symptoms, the test's restrictive system was used, based on the number of items scored above 3 on the positive and negative subscales. A positive profile is considered positive when there are more than two items above 3 on the positive subscale and fewer than three on the negative subscale; a negative profile is considered negative when the opposite is true, and a mixed profile is considered when both criteria are met simultaneously.

The collaborating psychiatrist was unknow with the users' words participants; he simply performed the psychopathological assessments and provided us with their medication information. There was great variability in the subjects' medications, including medications with sedative, anticholinergic, and β-adrenergic blocking effects. Furthermore, different methods of administration were observed, as some had depot (injectable) medications and others took them orally. In summary, it is worth noting that atypical antipsychotics, such as Leponex (over 28%) and risperidone (18%), were prevalent. It should be noted that more than half of the users were taking multiple medications. All participants were in two stages of the disease: multiple episodes currently in the acute phase, and multiple episodes currently in partial remission (APA, 2013; Caballo et al., 2014).

The inclusion criteria were as follows: a) diagnosis of schizophrenia according to DSM-5; b) an interval of at least one and a half years since diagnosis to avoid false positives, and c) positive/negative symptomatology at the time of recruitment. Our exclusion criteria were the following: a) comorbid diagnosis; A minimum of eighteen months since diagnosis was required for all participants, in order to reduce the likelihood of false positives and ensure the exclusion of comorbid conditions. Individuals with a primary diagnosis other than schizophrenia were excluded from the study, as confirmed by the collaborating psychiatrists b) cognitive deficits that might affect the reliability of test results, and c) diagnosis of residual schizophrenia.

Table 1. Demographic and clinical characteristics of the sample analyzed.

Variables	Total participants (n=50)	Positive symptoms (n=28)	Negative symptoms (n=22)
Age			
Mean ± SD	52.18±11.68	53.07±11.58	51.04 ± 11.98
Range	25-72	26-66	25-72
Gender			
Male	34 (68 %)	17 (60.7%)	17 (77.3%)
Female	16 (32 %)	11 (39.3%)	5 (22.7%)
Education			
Primary	20 (40 %)	13 (46.6%)	7 (31.8%)
Secondary	19 (38%)	10 (35.7%)	9 (40.9%)
High	11 (22%)	5 (17.9%)	6 (27.3%)
Age at 1º psychotic episode			
Mean ± SD	20.24±4.10	20.89±4.20	19.40 ±3.91
Range	13-28	15-28	13-27
N. of psychotic episodes			
Mean ± SD	5.16 ± 1.41	5.25 ± 1.37	5.04 ± 1.49
Range	2-8	3-7	2-8

Note: Values are presented as mean ± standard deviation or n/N. (number).

Instruments

Phonological verbal fluency test. The Controlled Oral Word Association Test (Benton et al., 1983) (COWAT) has traditionally been used to evaluate this aspect of language. Participants are asked to generate words

beginning with h the phonemes /f/, /a/, and /s/, the most used in the English language. However, when this test is administered in the Spanish population, the phonemes /p/, /m/, and /r/ are used, with evaluation being based on scales proposed in other studies (Casals-Coll et al., 2013; Peña-Casanova et al., 2009).

The following instruction was given: "*I'm going to say a letter and I want you to tell me as many words as you can beginning with that letter. Let's start. Tell me as many words as you can beginning with P*" They were also told that certain types of words, such as proper nouns, numbers and words stemming from the same root would not be counted. They subsequently continued with the other phonemes, /M/ and /R/.

Participants were given a minute for each letter once they had said the first word. The words generated by each patient in the corresponding time were recorded for each phoneme. The score was the number of different words generated in the period of a minute divided into different intervals: the first 15"; between 15-30"; between 30-45" and between 45-60", for each of the phonemes in question. At no point was help provided, although participants were encouraged to say as many words as possible for each of the phonemes.

Semantic verbal fluency test. The semantic lexical evocation test measures the capacity to generate nouns belonging to the same paradigm. This test, widely used by other authors (Benito-Cuadrado et al., 2002; Ramírez et al., 2005), evaluates the number of words participants can say from the same category (in this case, "animals") in one minute. The instruction given was: *"I want you to tell me all the names of animals you can think of. They can be sea, air or land animals, animals, domestic or wild, etc. Say as many as you can!"* Participants were told that extinct animals, diminutives, or names derived from the same species would not be scored. As in phonological fluency, the words were scored in different time intervals; no help was provided at any point, although they were encouraged to continue saying words.

As well as the total number of words generated both phonologically and semantically in the four-time intervals: 1st (from 0 to 15 seconds), 2nd (from 15 to 30 seconds), 3rd (from 30 to 45 seconds) and 4th (from 45 to 60 seconds), it was decided to analyze the frequency of the words generated, taking the EsPAL (Duchon et al., 2013) database as a reference. This database is extensively used in linguistic research to identify the characteristics of words. To assess word frequency and avoid bias, the lexical frequency per million (Log10) of words in the EsPAL database (Duchon et al., 2013) was used. Similarly, to determine whether words were frequent or infrequent, following other linguistic studies (Pérez et al., 2003; Soares et al., 2018; Van Heuven et al., 2014), word evaluations on the Zipf scale were used. Thus, all words with a score equal to or less than 3.5 were low use.

Furthermore, following studies that report that patients with negative symptoms emit fewer words than patients with positive symptoms (Tsakanikos & Claridge, 2005), which could generate discrepancies in the results, the percentage of frequent and infrequent words emitted by each of the participants is analyzed.

Finally, following previous studies (Baskak et al., 2008), we analyzed the associations between words, with the aim of determining the type of linguistic connections patients with schizophrenia make between words. To this end, using interrater agreement, all the word pairs were evaluated, comparing the words generated by each participant with each other, i.e., the first and the second, the second and the third, the third and the fourth, and so on, until the list of words was completed. To identify all the associations, each word emitted was evaluated for the presence or absence of a semantic, phonological, or semantic and phonological relationship with the word previously generated.

Following these studies we was carried out by two professionals: a clinical psychologist and a speech-language therapist. Their training consisted of a detailed explanation of the classification criteria: if two words shared similar sounds (e.g., palo and pato), the relationship was considered phonological; if the words were semantically related (e.g., rinoceronte and elefante), the relationship was classified as semantic; and if they shared both sounds and meanings (e.g., leopardo and guepardo), it was categorized as mixed.

To ensure consistency, a linguist prepared five examples for each criterion so that both raters clearly understood how to apply them. The evaluation was conducted in pairs for each participant, meaning that the ratings of one evaluator were compared with those of the other for each subject.

After the analysis conducted by each evaluator, a simple arithmetic mean was calculated by summing their scores and dividing the result by two. This average value was then used for the statistical analysis of the types of connections between the words produced by the participants.

Procedure

The participants were recruited from the following four mental health associations in Madrid and Castilla-La Mancha (Spain): the Madrid Association of Families and Friends of Persons with Schizophrenia; the VIVIR Association in Cuenca; the Center for Psychosocial Rehabilitation in Cuenca (province of Ciudad Real); and the Talavera Association of Friends, Families and Persons with Mental Ilnness in Talavera de la Reina (province of Toledo). Additionally, the Mental Health Unit at the Virgen de la Luz Hospital in Cuenca collaborated in the study. Before recruiting the participants, the project was approved by the Drug Research Ethics Committee of the corresponding healthcare, with code 11/2017. We collected data from May 2018 to May 2019.

Data analysis

The data were analyzed using the Statistical Package for Social Science 29.0 (SPSS®, IBM® Corp., Armonk, NY, USA). For the descriptive and inferential statistical analysis, different tests were implemented according

to the scale of the variable. Distribution normality was assessed using the Kolmogorov-Smirnov test. Most of the data had a normative distribution so parametric tests were used.

In instances of normally distributed data, independent samples t- tests for quantitative variables were used, and Chi-squared tests were conducted for categorical variables to reveal group-wise differences, when appropriate.

When the design included repeated measures, the assumption of sphericity was verified. For these analyses, 2x2 repeated measures ANOVA were performed. The Bonferroni test was used for post-hoc comparisons. Finally, interrater agreement was evaluated using the Kappa index.

The results were revaluated with a reliability of 95%, and statistical significance was set at $p < 0.05$.

Results

Results for phonological and semantic verbal fluency

Table 2 shows the results for phonological and semantic verbal fluency, as well as the total number of frequent and infrequent words. The total number of words generated is low, with the phoneme /R/ being the lowest, while the number of infrequent words is high.

Table 2. Results on the phonological and semantic fluency tests.

Variable	/P/	/M/	/R/	Animals
Total words				
Mean ± SD	13.32 ± 5.08	12.56 ± 5.19	11.88 ± 5.43	16.44 ± 5.76
Range	4-23	3-23	2-22	5-25
Word frequency				
Mean ± SD	8,26 ± 8,88	8,59±10,55	5.42±10,45	8,13 ± 4,22
Range	-4,51-54,27	-4,57-68,06	-4,50-69,78	0,87-16,98
Zipf scores				
Mean ± SD	41,19 ± 17,13	39,49 ± 13,64	34,54 ± 13,85	47,89 ± 11,79
Range	2,21-71,05	8,87-71,05	7,03-70,12	18,24 - 73,76
Low-frequency words				
Mean ± SD	6,72 ± 5,98	6,10 ± 5,54	6,12 ± 5,32	6,36 ± 5,27
Range	0-17	0-13	0-17	0-14
Percentage low-frequency words				
Percentage** ± SD	42.24 ± 34,61	40,72 ± 34,46	45,10 ± 34,61	40,35 ± 31,45
Range*	0-100	0-93,75	0-100	0-93,33
High-frequency words				
Mean* ± SD	5,50 ± 2,46	5,48 ± 1.94	4,50 ± 1,96	7,10 ± 2,70
Range*	0-12	1-10	1-9	2-13
Percentage high-frequency words				
Percentage ** ± SD	57,76 ± 35,52	59,28 ± 33,45	54,90 ± 32.61	59,70 ± 31,24
Range*	0-100	6,25-100	0-100	6,67-93,33

Note: Values are presented as mean ± standard deviation or n (percentages) *The mean of the sum of the frequencies of the low-frequency words was calculated. **The percentages will be calculated by establishing a rule of three, if the total number of words produced was 100, the total number of infrequent words would be X.

Significant relationships between variables

To determine the possible effect of certain sociodemographic and clinical variables on the performance of the groups (positive and negative symptomatology) in the verbal fluency tests and to ensure equivalence between them and buffer the possible effect of some of these variables, we performed an analysis of the variables in both groups finding: sex ($\chi 2(1) = 1.552$; $p = .213$), age ($Z = -.871$, $p = .284$), educational level ($\chi 2(2) = 1.241$; $p = .538$), age at first episode ($Z = -1.168$, $p = .243$) and the number of psychotic episodes ($Z = .579$, $p = .563$). No effect was revealed for any of the variables analyzed, indicating that the groups were homogeneous.

Table 3 shows the statistically significant differences between positive and negative symptomatology and the linguistic characteristics of the words generated according to the different time intervals analyzed. The participants with negative symptomatology generated fewer words, but these were of higher frequency. In addition, individuals with positive symptoms produced a greater number of low-frequency words in all the verbal fluency tests, both phonological and semantic. With respect to the time intervals, these differences were always notable in the last interval and, on occasions, in the first and third intervals.

Table 3. Relationships between the linguistic characteristics of words according to type of symptomatology.

Measures	Positive symptomatology	Negative symptomatology	*T*	*gl*	*DoM*	*P*
P						
Frequency I1	1,74±2,18	4,43±1,64	-4,754	48	-2,68	<,001
Frequency I2	1,69±2,33	2,82±1,71	-1,882	48	-1,13	,066
Frequency I3	0,41±1,85	2,12±1,63	-3,386	48	-1,71	<,001
Frequency I4	0,08±1,14	1,46±1,08	-4,305	48	-1,37	<,001
Total Frequency	5,84±10,57	11,60±4,6	1,796	48	-5,76	,022
Total High-Freq. Word	4,79±2,67	6,47±1,77	1,324	48	-1,68	,016
Perc. of High-Freq	29,33±15,08	96,18±7,72	8,235	48	-66,83	<,001
Total Low-Freq. Word	11,31±3,20	0,38±0,74	36,897	48	10,92	<,001
Perc. Low frecuency Word	70,65±15,66	5,01±9,13	7,126	48	64,17	<,001
M						
Frequency I1	1,45±1,88	4,84±1,44	-6,910	48	-3,38	<,001
Frequency I2	1,05±1,77	3,09±1,43	-4,336	48	-2,03	<,001
Frequency I3	0,74±1,56	2,90±1,34	-5,089	48	-2,15	<,001
Frequency I4	0,17±0,93	1,59±1,09	-4,764	48	-1,41	<,001
Total Frequency	5,39±12,52	13,01±4,22	-2,633	48	-7,61	,010
Total High-Freq. Word	5,13±2,1	5,95±1,59	-1,483	48	-0,81	,145
Perc. of High-Freq	33,75±13,88	96,66±9,29	-18,08	48	-62,90	<,001
Total Low-Freq. Word	10,34±3	0,23±0,70	15,087	48	10,10	<,001
Perc. Low frecuency Word	66,25±14,45	3,33±9,29	17,903	48	64,46	<,001
R						
Frequency I1	0,88±1,98	3,08±1,1	-4,565	48	-2,19	<,001
Frequency I2	1,08±1,35	2,06±1,59	-2,371	48	-0,99	,022
Frequency I3	-0,10±1,54	1,41±1,39	-3,597	48	-1,52	<,001
Frequency I4	-0,32±1,60	1,35±0,96	-4,269	48	-1,68	<,001
Total Frequency	3,69±13,23	7,81±3,49	-1,386	48	-4,11	,172
Total High-Freq. Word	4±1,92	5,14±1,95	-2,056	48	-1,14	,045
Perc. of High-Freq	28,89±14,45	92,32±11,51	-16,636	48	-63,42	<,001
Total Low-Freq. Word	10,24±2,73	0,42±0,67	16,053	48	9,81	<,001
Perc. Low frecuency Word	71,11±12,21	7,68±11,18	19,343	48	65,36	<,001
Animals						
Frequency I1	2,97±2,10	5,15±1,41	-4,121	48	-2,17	<,001
Frequency I2	1,52±1,57	2,91±1,02	-3,550	48	-1,39	<,001
Frequency I3	0,85±1,23	1,91±1,08	-3,139	48	-1,05	,003
Frequency I4	0,47±1,24	1,45±0,57	-3,345	48	-0,97	,002
Total Frequency	5,65±3,54	11,54±2,25	-6,683	48	-5,89	<,001
Total High-Freq. Word	5,96±2,42	8,66±2,28	-3,977	48	-2,70	<,001
Perc. of High-Freq	35,19±11,26	94,05±8,36	-20,228	48	-58,86	<,001
Total Low-Freq. Word	10,55±2,19	0,57±0,81	19,815	48	9,98	<,001
Perc. Low frecuency Word	64,81±12,27	5,94±8,36	19,131	48	59,31	<,001

Note: T: T-Students; DF: Degrees of freedom; DoM: Difference of means; Perc: Percentage; Freq: Frequency; P: significance; the low-frequency (<=3.50 Scale Zpif) and high-frequency (>3.50 Scale Zpif) words according to the ESPAL were counted.

A mixed-design ANOVA (2×2 design) was conducted to examine differences in the number of high- and low-frequency words produced by participants, with *word frequency* (high vs. low) as the within-subjects factor and *symptomatology* (positive vs. negative symptoms) as the between-subjects factor. Sex, age, and educational level were included as covariates.

The analysis revealed no significant interaction effects between these covariates and word frequency: sex ($F(1,46) = 1.814$, $p = .18$), age ($F(1,46) = 0.753$, $p = .39$), and education ($F(1,46) = 0.363$, $p = .55$).

In the phonological verbal fluency task, significant main effects of *word frequency* ($F(1,46) = 4.502$, $p = .039$) and *group* ($F(1,46) = 5.610$, $p = .022$) were observed, as well as a *frequency × group interaction* ($F(1,46) = 26.955$, $p < .001$). Post-hoc comparisons indicated that participants with negative symptoms produced

significantly more high-frequency words and significantly fewer low-frequency words compared to participants with positive symptoms (all $p < 0.05$).

In the semantic verbal fluency task, significant effects of group ($F(1,46) = 119.131$, $p < .001$) and a frequency × group interaction ($F(1,46) = 160.917$, $p < .001$) emerged. Participants with negative symptoms produced significantly more high-frequency words and significantly fewer low-frequency words than participants with positive symptoms (all $p < .001$).

Complete results are presented in Tables 4 and 5 and illustrated in Figure 1.

Table 4. Results of the analysis of variance to compare word frequency in the two study groups in phonological and semantic verbal fluency

Analysis of phonological verbal fluency	***F***	***DF***	***P***
Effect word frequency	4,502	1-46	.039
Group effect	5,610	1-46	.022
Interaction frequency*Group	26,955	1-46	<.001
Analysis of semantic verbal fluency			
Effect word frequency	0,194	1-46	.662
Group effect	119,131	1-46	<.001
Interaction frequency*Group	160,917	1-46	<.001

Note: F: ANOVA 2X2; DF: Degrees of freedom; P: Significance.

Table 5. Pairwise analysis between the frequency of word use and symptoms in the phonological and semantic verbal fluency test.

	Symptomatology	**Frecuencia**	$\bar{x}$	**SD**	**95% Confidence Interval**		***P***
					Lower	**Upper**	
Phonological verbal fluency	Positive symptoms	High frequency word	7,46	1,803	3,834	11,091	.014
		Low frequency word	11,323	0,476	10,366	12,281	.014
	Negative symptoms	High frequency word	11,538	2,123	7,264	15,812	.014
		Low frequency word	0,363	0,560	-0,765	1,491	.014
Semantic verbal fluency	Positive symptoms	High frequency word	5,947	0,448	5,046	6,848	<.001
		Low frequency word	10,606	0,325	9,952	11,260	<.001
	Negative symptoms	High frequency word	8,692	,527	7,631	9,754	<.001
		Low frequency word	0,496	,383	-,275	1,266	<.001

Note: $\bar{x}$: Means; SD: Standard deviation; P: Signification.

Regarding the post-hoc analysis, significant differences were observed between the groups with predominant positive and negative symptoms in phonemic verbal fluency. Participants with negative symptoms produced a smaller total number of words; however, most of these were high-frequency words, with very few low-frequency items. In contrast, participants with positive symptoms generated a greater number of words overall, but a higher proportion of these were low-frequency.

In the semantic verbal fluency task, a similar pattern was observed. Individuals with positive symptoms produced a larger number of animal names, although many of them were low-frequency. Conversely, participants with negative symptoms generated fewer responses overall, but these tended to be more frequent words (see Figure 1).

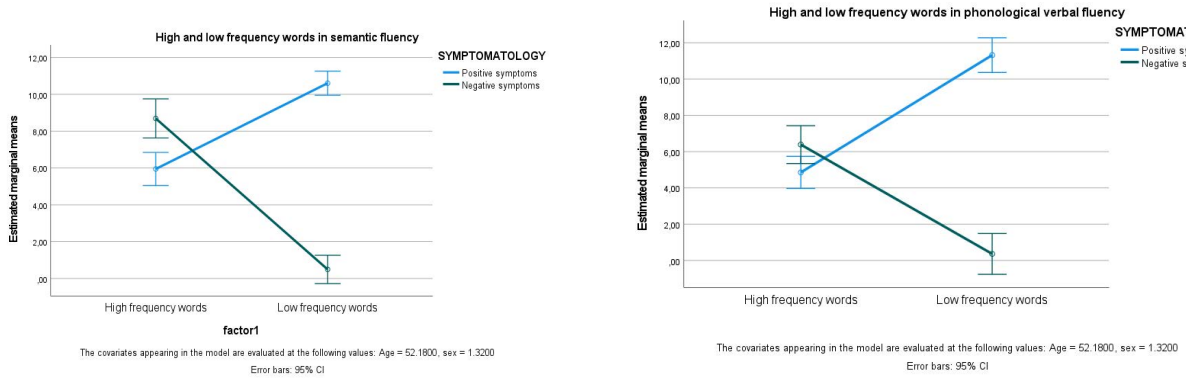

Figure 1. High- and low-frequency words according to study group and the phonological verbal fluency test and semantic verbal fluency test

Results of interrater agreement and linguistic comparisons

An interrater analysis was conducted, using the Kappa statistics to test the level of agreement on the relationships between words. We have indicated above, the evaluation was conducted in pairs for each participant, meaning that the ratings of one evaluator were compared with those of the other for each subject.

The data reveal agreement for both phonological and semantic verbal fluency. In the phonological relationships, the level of agreement is 0.544, in the semantic relationships, the level is 0.824, and between the scores on both, it is 0.804. As indicated in the instruments section, each evaluator analyzed the multiple word pairs produced by each participant. After obtaining the scores from both evaluators, the average score for each word pair was calculated. Subsequently, to maintain the independence of observations, the total mean score for phonological pairs and the total mean score for semantic pairs were obtained for each participant. These means per participant were used as the unit of analysis in a Student t test for paired samples, to test for significant differences between the two types of relationships (see Table 6).

Table 6. Results of the linguistic comparisons after the interrater analysis.

Measures	x̄	SD	Total mean	Total SD	DF	Cohen's D	T	*P*
P								
Phonological relations	2.5	1.682	2.29	1.70	49	1.70	7.629	<.001
Semantic relations	0.20	0.476						
M								
Phonological relations	1.71	1,669	1.50	1.80	49	1.72	5.888	<.001
Semantic relations	0.21	0,441						
R								
Phonological relations	1.69	1,494	1.18	2.25	49	2.25	3.699	<.001
Semantic relations	0.51	1,390						
Animals								
Phonological relations	0.21	0,589	-4,76	3.21	49	3.24	-10.468	<.001
Semantic relations	4.97	2,983						

Note: x̄: Mean; SD: Standard deviation; DF: Degrees of freedom; T: Student's t-test of paired samples; P: Significance

Discussion

With respect to our first aim, the results confirm that analyzing the words generated in verbal fluency tests using the ESPAL database (Duchon et al., 2013) is an objective, reliable and effective method for assessing word frequency in Spanish speakers with schizophrenia. Furthermore, this approach mitigates the subjectivity inherent in relying on rater judgments (Robles Garrote et al., 2015), which could help in studying and diagnosing the disease. This finding brings something new to the field because infrequent words have been a feature associated with an increase in formal thought disorder in schizophrenia (Crow, 1980; Liddle et al., 2002; Salavera et al., 2013), so if an objective test is available to assess them, it would improve the evaluations and monitoring of patients' symptoms.

Regarding the total number of words generated, it is observed that people with negative symptoms produce a significantly lower amount than those with less marked symptoms, a pattern previously described in the literature (Tsakanikos & Claridge, 2005). An analysis of the scores obtained by the 50 participants diagnosed with schizophrenia shows that their performance falls between the 29th and 40th percentiles in phonological verbal fluency, and between the 19th and 28th percentiles in semantic verbal fluency, in accordance with the scales established by Peña-Casanova et al. (2009). However, according to the scales proposed by Benito-Cuadrado et al. (2002), semantic verbal fluency would be within the average range. Results comparable to lower percentiles are obtained when using more recent scales, such as those of the Portellano-Pérez and Martínez-Arias (2020) verbal fluency test. The lower score observed in semantic verbal fluency reveals a more pronounced deficit in this modality, which is consistent with previous evidence on the semantic alterations characteristic of schizophrenia (Berberian et al., 2016; Bozikas et al., 2005).

It is important not only to analyze the total generated but also the time intervals., our study confirms that the most significant differences occur in the last interval. Studies with different populations using phonological and semantic verbal fluency tests report that it is in the later time intervals that individuals begin to generate lower-frequency words, as the lexicon is exhausted and the search for new words requires greater effort and becomes less productive and automatic (Crowe, 1998; Hurks et al., 2006). This might be the case with our participants, although it is true that low- frequency words also appear in the first-time interval, which suggests that the low- frequency lexicon employed by people with schizophrenia is that habitually used.

Our findings suggest that not only should time be considered, but also the characteristics of the language, since we observed that, depending on the letter used, lexical access is better or worse, with the letter P yielding the best results, possibly because it is the first letter of a greater number of words in Spanish (Duchon et al., 2013).

Furthermore, in contrast to various studies in English, our participants scored higher on semantic verbal fluency than phonological verbal fluency (Bozikas et al., 2005), although the meta-analysis conducted by Doughty & Done (2009) found no significant differences between semantic and phonological verbal fluency. Therefore, we consider it necessary to continue studying the possible influences of language and its characteristics on the linguistic symptomatology of schizophrenia.

Secondly, our results confirm that the type of schizophrenia symptomatology has a different effect on patients' linguistic deficits (Tsakanikos & Claridge, 2005). Individuals with positive symptoms generate a larger number of words than persons with negative symptoms, as has been studied and evidence in various works (Allen et al., 1993; Bora et al., 2019; Martínez Cano et al., 2019; Tsakanikos & Claridge, 2005).

However, a detailed analysis of the words generated reveals that the participants with positive symptoms generate a greater number of low frequency words compared to those with negative symptoms. An illustrative example of this distinction can be found when analyzing a selection of words generated by persons with schizophrenia, which includes terms such as, words that begin with m in Spanish generated by people with schizophrenia and a predominance of positive symptoms: Maleficient, Wickedness, Malevolent, Cursed, Jujube, Scoundrel. However, the words with m generated by a person with schizophrenia but with a predominance of negative symptoms are: "mesa, mama, mar, mundo, música".

Our study confirms the use of infrequent words in Spanish-speaking people with schizophrenia, as some English studies have shown (Baskak et al., 2008; Bora et al., 2019; Liddle et al., 2002). Nonetheless, our study delved into the analysis of the linguistic differences between individuals with schizophrenia according to their symptomatology, in terms of the frequency of words produced.

We observed that individuals presenting positive symptoms of the disorder tended to produce not only a greater number of words overall but also a higher proportion of low-frequency terms. Conversely, participants with negative symptoms generated fewer words, and these were generally more frequent.

Previous studies had already reported differences in the quantity of verbal production depending on symptom type (Tsakanikos & Claridge, 2005), but not in the frequency of the words employed. Our findings regarding word frequency may reflect hyperactivation of semantic nodes in individuals with positive symptoms and hypoactivation of those nodes in individuals with negative symptoms (Almeida & Radanovic, 2021; Bansal & Chatterjee, 2021; Doughty & Done, 2009; Kim et al., 2005; Leeson et al.,2005; Nestor et al., 2001).

The use of low frequency words by individuals with positive symptoms is thus confirmed. Several authors have reported deficits in verbal fluency in other languages, such as Italian (Tavano et al., 2008), Korean (Jo et al., 2023) and English (Baskak et al., 2008), suggesting that the linguistic deficit in this area may be universal and characteristic of the disease. The fact that our results evidence that low- frequency or uncommon words are present to a greater extent in the positive symptomatology of schizophrenia leads us to consider whether the formal thought disorder characteristic of schizophrenia might be due to a primary language deficit (Crow, 2000; Zegers, 2010), together with a cognitive deficit. This could explain the role of delirium as a deficient and inadequate activation of both internal and external language(Hoffman, 1986; Spence et al., 2000).

Previous studies have suggested a relationship between language deficits and formal thought disorders, indicating that those with poorer language (especially in the areas of semantics and pragmatics) present more severe formal thought disorder (Jo et al., 2023; Salavera et al., 2013). Future research should more clearly elucidate the role of low-frequency lexicon in thought disorder.

Furthermore, we found that low-frequency words appear in greater number in semantic verbal fluency, which could confirm that one of the language deficits in schizophrenia is due to failures in lexical storage and in the semantic network, an aspect that has been extensively researched (Henry & Crawford, 2004; Martínez-Cano et al., 2019; Tandon et al., 2024).

Our final aim was to determine whether the use of low frequency words by persons with schizophrenia might be due to their associating words more with their sound than their meaning, giving rise to what other authors have termed "word salad"(Andreasen, 1989).

Previous studies, such as that conducted by Baskak et al.,(2008) reported a greater association between words based on their phonetic characteristics rather than their meaning, especially compared to a group of persons without schizophrenia matched for age and sex. However, our results show this tendency is only present in the phonological verbal fluency tests, and not in the semantic verbal fluency tests. This might lead us to believe that patients with schizophrenia use the same strategies as healthy controls to perform verbal fluency tests, but use them less effectively, as they tap into phonological associations for letters and semantic associations for categories, as revealed by studies in other languages, such as Greek (Bozikas et al., 2005). This observation suggests a compelling line of research requiring in-depth attention in future studies, especially as regards possible discrepancies between languages and symptoms.

It is worth noting that these relationships were evaluated through agreement between blinded raters. Despite having achieved a high degree of consensus, this process is not exempt from the subjectivity inherent to evaluation by judges. In addition, it should be underlined that the lowest kappa index was observed in phonological relationships, an aspect that is a subject of ongoing debate.

Finally, different studies suggested that the delay the first psychotic episode is diagnosed, is related to a lower and slower recovery, therefore, is necessary use precise tools such as the measurement of low- frequency words, that could advance early diagnosis (Arango et al., 2019; Ayesa-Arriola et al., 2023).

Conclusions

Our study confirms firstly that the words generated in verbal fluency tests analyzed through the ESPAL database constitute an objective, reliable, and rapid method for assessing word peculiarity in Spanish-speaking people with schizophrenia.

Secondly, we confirmed that symptomatology affects the type of words generated. People with positive symptoms, in addition to producing a greater number of words, are less frequent; due to their size, people with negative symptoms generate fewer words overall but are more frequent. Therefore, if new studies confirm that the frequency of use of the words is related to the symptoms, we could have an objective criterion for monitoring the disease.

Finally, word associations are linked to the method of verbal fluency search. In phonological verbal fluency, phonological associations predominate, while in semantic verbal fluency, those corresponding to meaning predominate. This suggests that people with schizophrenia perform correct lexical search processes.

In the future, it is important to continue analyzing and studying linguistic aspects associated with schizophrenia as possible risk markers for this condition, integrating neurocognitive and neurolinguistic methods.

Limitations

The present study is not without limitations, one of which is that the medication administered to the patients was not considered, which might have affected our results. Additionally, the study did not include a healthy control group, which might have been of interest to compare results. This decision was taken because individuals without this disorder are generally assumed not to use uncommon or peculiar words in their everyday language.

We believe that future studies should control medication, for example by changing all antipsychotics to a single measure such as Chlorpromazine, following the recommendation of other studies (Andreasen et al., 2010; Leucht et al., 2014, 2014).

Authorship declaration

Alfonso Martínez-Cano: Conceptualization; Project administration; Software; Writing—original draft.

Jose Luis Santos: Supervision; Formal analysis.

Manuela Martínez-Lorca: Conceptualization; Supervision; Methodology; Writing—review; Editing.

Begoña Polonio-López: Supervision; Writing Original draft; Visualization; Formal analysis; Software; Editing.

Conflict of interest

There is no conflict of interest to declare.

References

Agurto, C., Norel, R., Wen, B., Wei, Y., Zhang, D., Bilgrami, Z., Hsi, X., Zhang, T., Pasternak, O., Li, H., Keshavan, M., Seidman, L. J., Whitfield-Gabrieli, S., Shenton, M. E., Niznikiewicz, M. A., Wang, J., Cecchi, G., Corcoran, C., & Stone, W. S. (2023). Are language features associated with psychosis risk universal? A study in Mandarin-speaking youths at clinical high risk for psychosis. *World Psychiatry*, *22*(1), 157-158. https://doi.org/10.1002/wps.21045.

Allen, H. A., Liddle, P. F., & Frith, C. D. (1993). Negative features, retrieval processes and verbal fluency in schizophrenia. *British Journal of Psychiatry*, *163*(DEC.), 769-775. https://doi.org/10.1192/bjp.163.6.769.

Almeida, V. N., & Radanovic, M. (2021). Semantic priming and neurobiology in schizophrenia: A theoretical review. *Neuropsychologia*, *163*. https://doi.org/10.1016/j.neuropsychologia.2021.108058.

Alvarado, R. U., & Arango, C. R. (2011). La ley de Zipf y el punto de transición de Goffman en la indización automática. *Investigación Bibliotecológica: Archivonomía, bibliotecología e información*, *25*(54), Article 54. https://doi.org/10.22201/iibi.0187358xp.2011.54.27482.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-V)* (Washington D.C. American Psychiatric Association, Ed.).

Andreasen, N. C. (1986). *Scale for the Assessment of Thought, Language, and Communication (TLC)* (Vol. 12, Número 3).

Andreasen, N. C. (1989). The scale for the assessment of negative symptoms (SANS): Conceptual and theoretical foundations. *British Journal of Psychiatry*, *155*(NOV. SUPPL. 7), 49-52. https://doi.org/10.1192/s0007125000291496.

Andreasen, N. C., Pressler, M., Nopoulos, P., Miller, D., & Ho, B.-C. (2010). Antipsychotic dose equivalents and dose-years: A standardized method for comparing exposure to different drugs. *Biological Psychiatry*, *67*(3), 255–262. https://doi.org/10.1016/j.biopsych.2009.08.040.

Arango, C., Baeza, I., Bernardo, M., Cañas, F., de Dios, C., Díaz-Marsá, M., García-Portilla, M. P., Gutiérrez-Rojas, L., Olivares, J. M., Rico-Villademoros, F., Rodríguez-Jiménez, R., Sánchez-Morla, E. M., Segarra, R., & Crespo-Facorro, B. (2019). Long-acting injectable antipsychotics for the treatment of schizophrenia in Spain. *Revista de Psiquiatria y Salud Mental*, *12*(2), 92-105. https://doi.org/10.1016/j.rpsm.2018.03.006.

Ayesa-Arriola, R., Miguel-Corredera, M., de la Foz, V. O.-G., Neergaard, K. D., Correa-Ghisays, P., Setién-Suero, E., & Crespo-Facorro, B. (2023). Education and long-term outcomes in first episode psychosis: 10-year follow-up study of the PAFIP cohort. *Psychological Medicine*, *53*(1), 66-77. https://doi.org/10.1017/S0033291721001112.

Balota, D. A., Burgess, G. C., Cortese, M. J., & Adams, D. R. (2002). The word-frequency mirror effect in young, old, and early-stage Alzheimer's disease: Evidence for two processes in episodic recognition performance. *Journal of Memory and Language*, *46*(1), 199-226. https://doi.org/10.1006/jmla.2001.2803.

Barattieri di San Pietro, C., Luzzatti, C., Ferrari, E., de Girolamo, G., & Marelli, M. (2023). Automated clustering and switching algorithms applied to semantic verbal fluency data in schizophrenia spectrum disorders. *Language, Cognition and Neuroscience*. https://doi.org/10.1080/23273798.2023.2178662.

Baskak, B., Ozel, E. T., Atbasoglu, E. C., & Baskak, S. C. (2008). Peculiar word use as a possible trait marker in schizophrenia. *Schizophrenia Research*, *103*(1-3), 311-317. https://doi.org/10.1016/j.schres.2008.04.025.

Benito-Cuadrado, M. M., Esteba-Castillo, S., Böhm, P., Cejudo-Bolívar, J., & Peña-Casanova, J. (2002). Semantic verbal fluency of animals: A normative and predictive study in a Spanish population. *Journal of Clinical and Experimental Neuropsychology*, *24*(8), 1117-1122. https://doi.org/10.1076/jcen.24.8.1117.8376.

Benton, A. L., Hamsher, K., & Sivan, A. B. (1983). Controlled oral word association test (COWAT). En *Multilingual aphasia examination* (3rd ed., p. 234). AJA Associates.

Berberian, A. A., Moraes, G. V., Gadelha, A., Brietzke, E., Fonseca, A. O., Scarpato, B. S., Vicente, M. O., Seabra, A. G., Bressan, R. A., & Lacerda, A. L. (2016). Is semantic verbal fluency impairment explained by executive function deficits in schizophrenia? *Brazilian Journal of Psychiatry*, *38*(2), 121–126. Scopus. https://doi.org/10.1590/1516-4446-2015-1663.

Bora, E., Yalincetin, B., Akdede, B. B., & Alptekin, K. (2019). Neurocognitive and linguistic correlates of positive and negative formal thought disorder: A meta-analysis. *Schizophrenia Research*, *209*, 2-11. https://doi.org/10.1016/j.schres.2019.05.025.

Bose, A., van Lieshout, P., & Square, P. A. (2007). Word frequency and bigram frequency effects on linguistic processing and speech motor performance in individuals with aphasia and normal speakers. *Journal of Neurolinguistics*, *20*(1), 65-88. https://doi.org/10.1016/j.jneuroling.2006.05.001.

Bozikas, V. P., Kosmidis, M. H., & Karavatos, A. (2005). Disproportionate impairment in semantic verbal fluency in schizophrenia: Differential deficit in clustering. *Schizophrenia Research*, *74*(1), 51-59. https://doi.org/10.1016/j.schres.2004.05.001.

Caballo, V. E., Salazar, I. C., & Carrobles, J. A. (2014). *Manual de psicopatología y trastornos psicológicos* (2.ª ed.). Pirámide.

Casals-Coll, M., Sánchez-Benavides, G., Quintana, M., Manero, R. M., Rognoni, T., Calvo, L., Palomo, R., Aranciva, F., Tamayo, F., & Peña-Casanova, J. (2013). Estudios normativos españoles en población adulta joven (proyecto NEURONORMA jóvenes): Normas para los test de fluencia verbal. *Neurologia*, *28*(1), 33-40. https://doi.org/10.1016/j.nrl.2012.02.010.

Chakrabarty, M., Bhattacharya, K., Chatterjee, G., Biswas, A., & Ghosal, M. (2023). Pragmatic deficits in patients with schizophrenia and right hemisphere damage: A pilot study. *International Journal of Language and Communication Disorders*, *58*(1), 169-188. https://doi.org/10.1111/1460-6984.12778.

Coleman, M. J., Levy, D. L., Lenzenweger, M. F., & Holzman, P. S. (1996). Thought disorder, perceptual aberrations, and schizotypy. *Journal of abnormal psychology*, *105*(3), 469–473. https://doi.org/10.1037//0021-843x.105.3.469

Crow, T. J. (1980). Positive and Negative Schizophrenic Symptoms and the Role of Dopamine. *The British Journal of Psychiatry*, *137*(4), 383-386. https://doi.org/10.1192/S0007125000071919.

Crow, T. J. (2000). Schizophrenia as the price that Homo sapiens pays for language: A resolution of the central paradox in the origin of the species 1. *Brain Research Reviews*, *31*(2-3), 118-129. https://doi.org/10.1016/S0165-0173(99)00029-6.

Crowe, S. F. (1998). Decrease in performance on the verbal fluency test as a function of time: Evaluation in a young healthy sample. *Journal of Clinical and Experimental Neuropsychology*, *20*(3), 391-401. https://doi.org/10.1076/jcen.20.3.391.810

Doughty, O. J., & Done, D. J. (2009). Is semantic memory impaired in schizophrenia? A systematic review and meta-analysis of 91 studies. *Cognitive Neuropsychiatry*, *14*(6), 473-509. https://doi.org/10.1080/13546800903073291.

Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior Research Methods*, *45*(4), 1246-1258. https://doi.org/10.3758/s13428-013-0326-1.

Erar, A. (2001). Bibliometrics or informetrics: Displaying regularity in scientific patterns by using statistical distributions. *Hacettepe Journal of Mathematics and Statistics, 31*, 113–125.

Gourovitch, M. L., Goldberg, T. E., & Weinberger, D. R. (1996). Verbal fluency deficits in patients with schizophrenia: Semantic fluency is differentially impaired as compared with phonologic fluency. *Neuropsychology, 10*(4), 573–577. https://doi.org/10.1037/0894-4105.10.4.573

Henry, J. D., & Crawford, J. R. (2004). A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology*, *18*(2), 284-295. https://doi.org/10.1037/0894-4105.18.2.284.

Hoffman, R. E. (1986). Verbal hallucinations and language production processes in schizophrenia. *The behavioral and brain sciences,* 9, 503-548. 10.1017/s0140525x00046781

Huck, A., Thompson, R. L., Cruice, M., & Marshall, J. (2017). Effects of word frequency and contextual predictability on sentence reading in aphasia: An eye movement analysis. *Aphasiology*, *31*(11), 1307-1332. https://doi.org/10.1080/02687038.2017.1278741.

Hurks, P. P. M., Vles, J. S. H., Hendriksen, J. G. M., Kalff, A. C., Feron, F. J. M., Kroes, M., Van Zeben, T. M. C. B., Steyaert, J., & Jolles, J. (2006). Semantic category fluency versus initial letter fluency over 60 seconds as a measure of automatic and controlled processing in healthy school-aged children. *Journal of Clinical and Experimental Neuropsychology*, *28*(5), 684-695. https://doi.org/10.1080/13803390590954191.

Jo, Y. T., Park, S. Y., Park, J., Lee, J., & Joo, Y. H. (2023). Linguistic anomalies in the language of patients with schizophrenia. *Schizophrenia Research: Cognition*, *31*. https://doi.org/10.1016/j.scog.2022.100273.

Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia bulletin*, *13*(2), 261–276. https://doi.org/10.1093/schbul/13.2.261

Kim, J. J., Ho Seok, J., Park, H. J., Soo Lee, D., Chul Lee, M., & Kwon, J. S. (2005). Functional disconnection of the semantic networks in schizophrenia. *Neuroreport*, *16*(4), 355–359. https://doi.org/10.1097/00001756-200503150-00010

Leeson, V. C., Simpson, A., McKenna, P. J., & Laws, K. R. (2005). Executive inhibition and semantic association in schizophrenia. *Schizophrenia Research*, *74*(1), 61-67. https://doi.org/10.1016/j.schres.2004.07.011.

Leucht, S., Samara, M., Heres, S., Patel, M. X., Woods, S. W., & Davis, J. M. (2014). Dose equivalents for second-generation antipsychotics: The minimum effective dose method. *Schizophrenia Bulletin*, *40*(2), 314–326. https://doi.org/10.1093/schbul/sbu001.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). Oxford University Press.

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

Martínez, A., Martínez-Lorca, M., Santos, J. L., & Martínez-Lorca, A. (2018). Protocolo de evaluación de la prosodia emocional y la pragmática en personas con esquizofrenia. *Revista de Investigación en Logopedia*, *8*(2), 129-146. https://doi.org/10.5209/RLOG.59892.

Martínez-Cano, A., Martínez-Lorca, M., Santos Gómez, J. L., & Martínez-Lorca, A. (2019). Evaluación de la dimensión semántica y fonológica en pacientes con esquizofrenia. *Revista de la Asociación Española de Neuropsiquiatría*, *39*(135), 133-155. https://doi.org/10.4321/s0211-57352019000100008.

Martínez-Cano, A., de la Sacristana, R. F.-B. G., Martín-Conty, J. L., Mordillo-Mateos, L., Bernal-Jimenéz, J. J., Polonio-López, B., & Martínez-Lorca, M. (2024). Fundamental Frequency of the Voice in Schizophrenia and Its Value as a Biomarker of the Disease. *Journal of Voice: Official Journal of the Voice Foundation*, S0892-1997(24)00394-1. https://doi.org/10.1016/j.jvoice.2024.11.005.

Martínez-Cano, A., Martínez-Lorca, A., Criado, J. J., & Martínez-Lorca, M. (2025a). Macrotextual, microtextual and writing analysis of texts written by people with schizophrenia differentiated by their symptoms. *Journal of Writing Research*, *17*(2), 287–308. https://doi.org/10.17239/jowr-2025.17.02.04.

Martínez-Cano, A., Polonio-López, B., Bernal-Jiménez, J. J., Martín-Conty, J. L., Mordillo-Mateos, L., & Martinez-Lorca, M. (2025b). Semantic processing deficits and their use as early biomarkers in schizophrenia. *Healthcare*, *13*(16), 1958. https://doi.org/10.3390/healthcare13161958.

Morrison, C. M., Ellis, A. W., & Quinlan, P. T. (1992). Age of acquisition, not word frequency, affects object naming, not object recognition. *Memory & Cognition*, *20*(6), 705-714. https://doi.org/10.3758/BF03202720.

Nestor, P. G., Han, S. D., Niznikiewicz, M., Salisbury, D., Spencer, K., Shenton, M. E., & McCarley, R. W. (2001). Semantic disturbance in schizophrenia and its relationship to the cognitive neuroscience of attention. *Biological Psychology*, *57*(1), 23-46. https://doi.org/10.1016/S0301-0511(01)00088-6.

Oomen, P. P., de Boer, J. N., Brederoo, S. G., Voppel, A. E., Brand, B. A., Wijnen, F. N. K., & Sommer, I. E. C. (2022). Characterizing speech heterogeneity in schizophrenia-spectrum disorders. Journal of Psychopathology and Clinical Science, 131(2), 172-181. https://doi.org/10.1037/ABN0000736.

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

Pearson, P. D., & Studt, A. (1975). Effects of word frequency and contextual richness on children's word identification abilities. *Journal of Educational Psychology*, *67*(1), 89-95. https://doi.org/10.1037/h0078675.

Peña-Casanova, J., Quiñones-Úbeda, S., Gramunt-Fombuena, N., Quintana-Aparicio, M., Aguilar, M., Badenes, D., Cerulla, N., Molinuevo, J. L., Ruiz, E., Robles, A., Barquero, M. S., Antúnez, C., Martínez-Parra, C., Frank-García, A., Fernández, M., Alfonso, V., Sol, J. M., & Blesa, R. (2009). Spanish multicenter normative studies (NEURONORMA project): Norms for verbal fluency tests. *Archives of Clinical Neuropsychology*, *24*(4), 395-411. https://doi.org/10.1093/arclin/acp042.

Peralta Martín, V., & Cuesta Zorita, M. J. (1994). Validacion de la escala de los síndromes positivo y negativo (PANSS) en una muestra de esquizofrenicos españoles [Validation of positive and negative symptom scale (PANSS) in a sample of Spanish schizophrenic patients]. *Actas Luso-Espanolas de Neurologia, psiquiatria y Ciencias Afines*, *22*(4), 171–177.

Perea, M., Rosa, E., & Gómez, C. (2005). The frequency effect for pseudowords in the lexical decision task. *Perception & Psychophysics*, *67*(2), 301-314. https://doi.org/10.3758/BF03206493

Pérez, M. Á., Alameda, J. R., & Vega Cueto, F. (2003). Frecuencia, longitud y vecindad ortográfica de las palabras de 3 a 16 letras del Diccionario de la Lengua Española (RAE, 1992). *Revista Electrónica de Metodología Aplicada*, *8*(2), 1–10. https://doi.org/10.17811/rema.8.2.2003.1-10

Portellano, J., & Martínez-Arias, R. (2020). *TFV. Test de Fluidez Verbal*. TEA Ediciones.

Ramírez, M., Ostrosky-Solís, F., Fernández, A., & Ardila-Ardila, A. (2005). Fluidez verbal semántica en hispanohablantes: Un análisis comparativo. *Revista de neurología*, *41*(3), 463-468. https://doi.org/10.33588/rn.4108.2004597

Robles Garrote, P., Del, M., Rojas, C., Garrote, R., & Rojas, P. Y. (2015). La validación por juicio de expertos: Dos investigaciones cualitativas en lingüística aplicada validation by expert judgements: Two cases of qualitative research in applied linguistics. *Nebrija Universidad*, *18*, 124-139. https://doi.org/10.26378/rnlael918259.

Salavera, C., Puyuelo, M., Antoñanzas, J. L., & Teruel, P. (2013). Semantics, pragmatics, and formal thought disorders in people with schizophrenia. *Neuropsychiatric Disease and Treatment*, *9*, 177-183. https://doi.org/10.2147/NDT.S38676.

Spence, S. A., Liddle, P. F., Stefan, M. D., Hellewell, J. S. E., Sharma, T., Friston, K. J., Hirsch, S. R., Frith, C. D., Murray, R. M., Deakin, J. F. W., & Grasby, P. M. (2000). Functional anatomy of verbal fluency in people with schizophrenia and those at genetic risk. Focal dysfunction and distributed disconnectivity reappraised. *British Journal of Psychiatry*, *176*(JAN.), 52-60. https://doi.org/10.1192/bjp.176.1.52.

Soares, A. P., Iriarte, Á., de Almeida, J. J., Simões, A., Costa, A., Machado, J., França, P., Comesaña, M., Rauber, A., Rato, A., & Perea, M. (2018). Procura-PALavras (P-PAL): A Web-based interface for a new European Portuguese lexical database. *Behavior Research Methods, 50*(4), 1461–1481. https://doi.org/10.3758/s13428-018-1058-z

Tandon, R., Nasrallah, H., Akbarian, S., Carpenter, W. T., DeLisi, L. E., Gaebel, W., Green, M. F., Gur, R. E., Heckers, S., Kane, J. M., Malaspina, D., Meyer-Lindenberg, A., Murray, R., Owen, M., Smoller, J. W., Yassine, W., & Keshavan, M. (2024). The schizophrenia syndrome, circa 2024: What we know and how that informs its nature. *Schizophrenia Research*, *264*, 1-28. https://doi.org/10.1016/j.schres.2023.11.015.

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. *Schizophrenia Research*, *102*(1-3), 53-62. https://doi.org/10.1016/j.schres.2008.02.008.

Tsakanikos, E., & Claridge, G. (2005). More words, less words: Verbal fluency as a function of «positive» and «negative» schizotypy. *Personality and Individual Differences*, *39*(4), 705-713. https://doi.org/10.1016/j.paid.2005.02.019.

Van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521.

Wilson, R. S., Bacon, L. D., Fox, J. H., Kramer, R. L., & Kaszniak, A. W. (1983). Word frequency effect and recognition memory in dementia of the alzheimer type. *Journal of Clinical Neuropsychology*, *5*(2), 97-104. https://doi.org/10.1080/01688638308401157.

Zegers, O. D. (2010). Schizophrenia, language and evolution (or the schizophrenias as logopathies). *Actas Espanolas De Psiquiatria*, *38*(1), 1-7. https://actaspsiquiatria.es/index.php/actas/article/view/692.

Zipf, G. K. (1945). *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology.* Cambridge. https://doi.org/10.2307/2572028.