



Differentiating PNES from epileptic seizures using conversational analysis on French patients: A prospective blinded study

Julien Biberon, Astrid de Liège, Bertrand de Toffol, Nadège Limousin, Aline-Marie Florence, Coline Duwicquet, Wissam El-Hage

► To cite this version:

Julien Biberon, Astrid de Liège, Bertrand de Toffol, Nadège Limousin, Aline-Marie Florence, et al.. Differentiating PNES from epileptic seizures using conversational analysis on French patients: A prospective blinded study. *Epilepsy & Behavior*, 2020, 111, pp.107239 -. 10.1016/j.yebeh.2020.107239 . hal-03490936

HAL Id: hal-03490936

<https://hal.science/hal-03490936>

Submitted on 22 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Differentiating PNES from epileptic seizures using conversational analysis on French patients: a prospective blinded study

Julien Biberon^{a,*}, Astrid de Liège ^a, Bertrand de Toffol^{a, b}, Nadège Limousin^a, Wissam El-Hage^{b, c}, Aline-Marie Florence^{d, e}, Coline Duwicquet^a

^a Department of Neurology, University hospital of Tours, France

^b UMR 1253, iBrain, Université de Tours, Inserm, Tours, France

^c Centre de Psychotraumatologie CVL, Pôle de Psychiatrie, CHRU de Tours, Tours, France

^d Laboratoire de santé publique, CHRU de Tours, Tours, France

^e Equipe “éducation, éthique, santé” EA 7505, Université de Tours, Tours, France

* Corresponding author

Declarations of interest: none

1. INTRODUCTION

Psychogenic non-epileptic seizures (PNES) resemble epileptic seizures but are not caused by the occurrence of excessive cortical neuronal discharges. PNES are related to psychological processes and classified as conversion disorder with seizures or convulsions in the ICD-10 (F44.5) [1]. They are classified as functional neurological symptom disorders or as dissociative identity disorders in the DSM-5 [2, 3]. The diagnosis of PNES is an important therapeutic issue in establishing appropriate psychotherapy and in avoiding unnecessary consumption of potentially iatrogenic care (anti-epileptic drugs, visits to emergencies, intensive care, etc.) [4]. There are no pathognomonic clinical characteristics of PNES and their diagnosis remain challenging [5, 6]. The video-EEG recording of psychogenic seizures is currently the "gold-standard" for the positive diagnosis of PNES. However, video-EEG recording is not available everywhere, especially in non-industrialized countries. Long access duration for this examination can delay positive diagnosis and appropriate care. Other diagnostic tools are therefore needed.

Work carried out by a Bielefeld research team in Germany was aimed at analyzing the linguistic content of the seizure descriptions by patients. Authors showed that patients used a different communicative style to talk about their seizures between epileptic and PNES. This communicative style could predict diagnosis of PNES and epilepsy [7-9]. Their work was replicated on English and on Italian speaking patients [9-12]. Based on these previous studies, a Sheffield research group developed a linguistic rating score, the diagnostic scoring aid (DSA), based on 17 interactive and linguistic characteristics. Assisted by the DSA, linguists correctly predicted the diagnosis of PNES and epilepsy in 80% of cases [12].

The aim of our study was to evaluate the contribution of linguistic analysis in the differential diagnosis between epileptic and PNES, in a French patient population. Our primary objective was to assess the sensitivity and specificity of the linguistic diagnosis on this population. Our secondary objectives were: to evaluate the sensitivity and specificity of the linguistic diagnosis assisted by the DSA and to identify the most relevant elements of the DSA, to assess inter-observer reproducibility for linguistic diagnosis, assisted or not assisted by the score, to analyze metaphorical concepts and diagnostic labels used by patients.

2. METHODS

2.1. Patients

Patients were included by the department of neurology (University Hospital of Tours, France) between November 2016 and June 2018. Were eligible all patients whose diagnosis of PNES or epilepsy had not been previously confirmed by video-EEG recording and who were hospitalized in the department for an extended video-EEG monitoring (between 48 hours and 5 days). Patients were included only if they had at least one habitual seizure during monitoring and if the ictal electro-clinical features allowed an unequivocal diagnosis of epilepsy or PNES. Included patients had to be native French speakers, over 18 years of age. Exclusion criteria comprised previously established diagnosis of PNES or epilepsy, intellectual disability or aphasia.

2.2.Interview

A neurologist conducted the interview at the beginning of the monitoring (two neurologists conducted all interviews, ADL and JB). Physicians were not aware of the initial diagnostic suspicion or seizure descriptions reported in previous interviews. They were purposely unaware of patients' medical records. The interview of each patient followed an identical structure, as developed by the German EpiLing project and Reuber et al. [12-13]. It consisted of a semi-structured interview containing the 5 following questions: *What do you expect from this hospitalization? Can you describe your very first seizure? Can you describe your latest seizure? Can you describe the most impressive or serious seizure? How do you spend your free time?*

Throughout the interview, neurologists had to remain neutral, let patients express themselves, and make sure not to interrupt them. Neurologists had to adopt an empathic and respectful attitude and tolerate prolonged silences but could ask open-ended questions or pick up points from the patients' replies in order to help them focus on the seizures' descriptions with the aim of defining their linguistic profiles.

2.3.Linguistic Analysis

The original audio recording and transcribed interview were then analyzed and linguistic analysis was carried out off line independently by two neurologists from the neurology department. *Raters* were purposely unaware of the patient's medical record and of the video-EEG diagnosis. First, based on qualitative analysis, *raters* had to determine the epileptic or psychogenic origin of described seizures. Linguistic aspects previously described to distinguish PNES and epileptic seizures were scrutinized: formulation work (pause, reformulation of subjective symptoms), focus on seizure description, spontaneous reference to

attempted seizure suppression, seizure description by negation, and description of periods of reduced consciousness or self-control [8, 13].

We adapted the Diagnostic Scoring Aid (DSA) from the previous work of Sheffield into French [12]. This score included 17 items based on previously described linguistic aspects, in order to discriminate epileptic seizures from PNES. Each of the 17 items was rated "1" if item was more in keeping with epilepsy, "-1" if item was more in keeping with PNES and "0" if rater did not know or if item was impossible to rate. A total DSA score for each patient was derived by adding the 17 sub-scores. We established the diagnostic value and inter-rater reliability of individual sub-scores. Then we analyzed the diagnostic value of the DSA, weighting the five most reliable and accurate items by a factor of two. We also analyzed the diagnostic value of a simplified DSA including only these 5 items.

Terms used by patients to describe their seizures were scrutinized. The terms "crise" (seizure), "trou noir" (memory lapse), "malaise" (blackout) were counted in a patient's interview [8, 14]. Raters also quantified metaphors used by patients to describe their seizures and classified them according to three previously established categories: "the seizure is an AGENT/FORCE", "the seizure is an EVENT/SITUATION" and "the seizure is a PLACE/SPACE" [8, 15-17] (Appendix 2).

2.4.Ethics

This study was validated by the clinical research support ethics group (ERERC, Espace de Réflexion Ethique Région Centre-Val de Loire, France). All patients interviewed received an oral explanation and information notice about the study. They were free to accept or refuse to participate in this study. Their consent was obtained after a period of reflection. Their final decision had no impact on their care.

2.5.Statistical Analysis

The data were analyzed using the SPSS statistical software (version 21, IBM incorporation, USA). Data from the clinical sample were presented according to their mean and standard deviation from the mean. Means were compared with the non-parametric Mann-Whitney-Wilcoxon test. Quantitative data were analyzed using the Spearman test.

We performed a ratio comparison with the non-parametric chi square test for diagnostic label analysis and metaphorical analysis and we calculated odd ratios. Two-sided P values <0.05 were considered significant.

For each rater, a nonparametric receiver operating characteristic (ROC) curve was based on DSA scores in order to assess the sensitivity and specificity of the diagnostic procedure, assisted by DSA. ROC analysis was made for each of the 3 DSA (original DSA, weighted DSA and simplified DSA). Inter-rater reproducibility of linguistic diagnosis and analysis assisted by DSA was assessed by the concordance measurement and the coefficient κ . A very strong agreement corresponds to a value of κ between 0.8-1.00, a strong agreement between 0.6-0.79, a moderate agreement between 0.4 and 0.59, a weak agreement between 0.2 and 0.39, a very weak agreement between 0-0.19 [18]. Concordance measurements were also evaluated for each sub-score of the DSA. Diagnosis accuracy for each of these sub-scores was assessed with a non-parametric Chi square test. For this analysis, the Bonferroni correction was applied and P values <0.0029 were considered significant.

3. RESULTS

Among 53 eligible patients, 32 were included: 13 patients (40.6%) had one or more PNES (13 women) and 19 patients had one or more epileptic seizures (9 women). Twenty-one patients were excluded: 12 had no seizure during monitoring, interviews of 5 patients were not reliable (4 due to intellectual disability, 1 due to aphasia) 4 patients had attacks of other origins.

Patients' demographic and clinical characteristics are presented in Table 1. Patients with PNES were younger than patients with ES at time of diagnosis ($p=0.048$). The average age at disease onset in patients with PNES was 24.5 years and 37 years in patients with ES ($p=0.049$). Patients with ES (12.8 ± 0.6) had more average years of schooling than patients with PNES (10.4 ± 0.7) ($p=0.03$). Patients with PNES took antidepressants more frequently than patients with epilepsy ($p=0.005$). Diagnosis delay seemed longer for patients with PNES but differences did not reach significance ($p=0.079$).

There were no significant differences between patients with PNES and patients with ES concerning duration of monitoring, duration of interview and treatment with AED ($p>0.07$). Psychiatric comorbidities were known for 4 patients with PNES: Post-traumatic stress disorder with comorbid depression ($n=2$), PTSD ($n=1$), depression and anxiety ($n=1$); and only one epileptic patient suffered from anxiety. Epileptic patients had temporal lobe seizures (13), bi-temporal seizures (2), and frontal lobe seizures (3). One patient had a clonic tonico-clonic seizure.

Diagnosis based on qualitative analysis

Based on qualitative analysis, *Rater 1* predicted the correct diagnosis in 27 out of 32 patients (84%) and *Rater 2* in 28 out of 32 patients (88%). Interrater reliability of qualitative analysis was satisfactory ($k=0.68$, interrater agreement=84.4%). Patients with epilepsy exhibited an ongoing effort to describe subjective seizure symptoms, the beginning and end of the phase of reduced self-control or possible strategies for interrupting or preventing seizures. They performed extensive formulation work including repairs, reformulations, hesitations, changes in grammatical construction and restarts. They tried to reconstruct what happened by referring to witnesses' reports and to what they still remembered before and after the "gap".

Patients with PNES provided little information about subjective symptoms and often focused more on seizure circumstances or consequences. They may have resisted the interviewer's attempts to focus on one particular seizure. They often considered the phase of reduced self-control as the entire seizure and often highlighted what they did not experience or restricted seizure description to absolute negations. They often denied any knowledge or memory of their seizures when the interviewer asked them to contour more precisely this gap. They could provide detailed seizure description only after repeated prompting but rarely spontaneously. Patients with PNES could sometimes refer to witnesses' descriptions but often failed to describe the circumstances or consequences of the seizure.

3.1.Diagnosis based on the DSA

Mean total DSA scores were significantly higher for epileptic patients than for patients with PNES: *Rater 1*: 9.2 (range -6 to 16) versus -4.5 (range -15 to 7); *Rater 2*: 7.7 (range 6 to 16) versus -6.2 (range -14 to 14). The differences were significant for both raters ($p=2.10^{-4}$). Inter-rater reliability and predictive value of each item for each rater are presented in Table 2. Five items were selected to construct the weighted DSA and the simplified DSA. These items concerned: "description of 'gaps'" (phases of reduced self-control or recollection), "Response to challenges of statements about 'gaps'", "Subjective seizure symptoms", "'Formulation efforts' associated with the description of subjective seizure symptoms ('formulation efforts' include restarts, reformulations, neologisms)" and "consistence of Metaphoric seizure conceptualization".

For each DSA (original, weighted, and simplified) and each rater (1 and 2) ROC are summarized in Figure 1. Using the original DSA score and the optimal diagnostic cut-off score determined by ROC (5), *Rater 1* would have correctly diagnosed 84% of the

cases (27/32). The Area Under the Curve (AUC) was 0.893 (SE 0.06, asymptotic significance 2.10^{-4}). For *Rater 1*, using the weighted DSA score AUC improved to 0.909 (SE 0.05, asymptotic significance 1.10^{-4}) and, with the optimal diagnostic cut-off score determined by ROC (6), *Rater 1* would have correctly diagnosed 81% of the cases (26/32 patients). Using the simplified DSA score AUC improved to 0.927 (SE 0.045, asymptotic significance 5.10^{-5}) and, with the optimal diagnostic cut-off score determined by ROC (1.5), *Rater 1* would have correctly diagnosed 88% of the cases (28/32 patients).

Using the original DSA score and the optimal diagnostic cut-off score determined by ROC (0.5), *Rater 2* would have correctly diagnosed 84% of the cases (27/32). The Area Under the Curve (AUC) was 0.899 (SE 0.06, asymptotic significance 2.10^{-4}). Using the weighted DSA score AUC improved to 0.907 (SE 0.061, asymptotic significance 1.10^{-4}) and, with the optimal diagnostic cut-off score determined by ROC (-0.5), *Rater 2* would have correctly diagnosed 88% of the cases (28/32 patients). Using the simplified DSA score AUC improved to 0.903 (SE 0.066, asymptotic significance 1.10^{-4}) and, with the optimal diagnostic cut-off score determined by ROC, *Rater 2* would have correctly diagnosed 91 % of the cases (29/32).

3.2. Inter-rater reliability

Inter-rater reliability of analysis based on DSA score was good ($k=0.68$, inter-rater agreement=81.3%). Rating of patients on individual items is presented on Table 3. Neurologists were in perfect agreement on 280/544 items (51.5%). They frankly disagreed on 25/544 items (4.6%).

3.3. The Use of Diagnostic Labels by Patients

Patients with PNES used less diagnostic labels than epileptic patients (28 vs. 96). On average, patients with PNES used only 2 diagnostic labels during interviews whereas patients with epileptic seizures used 5 diagnostic labels, but this difference was not significant ($p=0.9$). The French word "crise" (seizure) was the most frequently used (124 uses) among 32 interviews. Four patients with PNES used this label 10 times whereas 16 epileptic patients used this term 70 times. Three patients with PNES used the diagnostic label "malaise" (blackout) 5 times and 4 patients with epileptic seizures used this label 21 times. Five patients with PNES used the term "trou noir" (memory lapse) 13 times whereas 2 epileptic patients used this label 5 times. Six epileptic patients and

one patient with PNES used other diagnostic labels such "perte de conscience" (loss of consciousness) (4 times), "blanc" (blank) (3 times), "flou total" (complete blur), "vide" (void), "épisode" (episode), "rupture" (breach). These differences in using diagnostic labels were not significant.

3.4. Seizure metaphors

A total of 81 metaphors were used during interviews. Patients with epileptic seizures preferentially used metaphors from the AGENT/FORCE category (7 out of 19 patients, 25 metaphors) compared to patients with PNES (2 out of 13 patients, 2 metaphors). This difference was significant ($p=0.03$).

There was a trend for patients with epilepsy to use metaphors from the EVENT/SITUATION category (8 out of 19 patients, 14 metaphors) more frequently than patients with PNES (1 out of 13 patients, 2 metaphors) but this difference was not significant ($p=0.08$). Patients with PNES tended to use metaphors from the SPACE/PLACE category more frequently (11 out of 13 patients, 18 metaphors) than patients with epilepsy (6 out of 19 patients, 11 metaphors) but this difference did not reach significance ($p=0.051$). There was no difference in terms when using metaphors from other categories between the 2 groups. Three patients with epilepsy used 6 metaphors from other categories and 3 patients with PNES used 3 metaphors from other categories ($p=0.70$).

4. DISCUSSION

Our study showed that the diagnostic value of conversational analysis originally described on German patients and reproduced on English and Italian patients could be replicated on French-speaking patients with good inter-rater reliability. We also showed that neurologists could employ conversational analysis with a good diagnostic yield.

In our work, qualitative evaluation performed by neurologists, based on 17 previously described linguistic features, could correctly predict diagnosis with good sensibility and specificity [7, 9-12, 19]. Surprisingly, when these items were scored using DSA, accuracy seemed to be less important than qualitative evaluation: rater 2 correctly diagnosed 88% of the cases (28/32 patients) with qualitative evaluation but only 84% of the cases (27/32 patients) using DSA. In Italian studies, Papagno et al correctly identified 49 cases out of 61 (80%) using DSA and Cornaggia et al correctly identified 9

cases out of 10 (90%) using DSA [10-11]. In the study of Reuber et al, based on their qualitative evaluation, linguists predicted the correct diagnosis in 85% of the cases (17/20 patients). In this work, using the DSA and the optimal diagnostic cut-off score determined by the ROC curve, *rater 1* and *rater 2* would have correctly categorized 80% and 75% of the cases respectively [12]. All items did not provide the same diagnostic accuracy or inter-rater reliability. Unlike qualitative analysis, scoring did not allow a rater to place emphasis on one or more items throughout the interview in each particular case. This difference could explain the inferior accuracy of DSA compared to qualitative analysis, although the same 17 linguistic features were used to score DSA. In our work, simplified DSA obtained the best accuracy. Using simplified DSA and optimal cut-off, each rater would have correctly classified 1 more patient compared to qualitative analysis. Rater 1 correctly classified 1 more patient compared to the original DSA and Rater 2 classified 2 more patients compared to the original DSA. In their study, Jenkins et al. demonstrated that using a simpler questionnaire containing 12 items, neurologists could predict a diagnosis with good accuracy (85% of sensitivity and 77% of specificity). They correctly classified 82% of the cases (27/33 patients) [20]. They found six items significantly associated with the final diagnosis: “the patient readily volunteers descriptions of seizure symptoms”, “In response to enquiries, the patient readily provides more detailed seizure descriptions”, “The patient provides detailed seizure descriptions”, “The patient focuses more on the symptoms of the seizures rather than the consequences of seizures or the situations in which they occurred”, “The patient’s seizure descriptions are characterized by formulation efforts” and “the interview was challenging for the neurologist”. Four of these 6 items are similar to 4 of the 5 items we selected: "description of 'gaps'", "Response to challenge of statements about 'gaps'", "Subjective seizure symptoms", "‘Formulation effort’ associated with the description of subjective seizure symptoms”. Surprisingly, only one item in our study corresponded to those found significantly associated to the correct diagnosis by Reuber et al: “consistence of metaphoric seizure conceptualization” [12]. These differences could be due to the clinical experience and the initial training of the raters: in Jenkins’ and our studies, raters were neurologists whereas in Reuber’ study raters were linguists. This can influence the way of conducting the interview or analyzing the conversation. Neurologists are not professionally trained in linguistic analysis. They could have incorrectly taken into account some linguistic clues or subtleties or not sufficiently let patients follow their own agenda of interview.

Furthermore, given their usual method of interviewing and collecting descriptions of seizures, neurologists could have been influenced by factual features and specific symptoms contained in patients' descriptions. However, the two raters knew medical literature on this subject, and had read previous work published by German, English and Italian teams. They carefully followed German and English guidelines concerning the method of interviewing patients and of analyzing the conversations following the 17 linguistic features previously described by these teams. Jenkins et al. already described the neurologists' capacity to change their method of interviewing patients and their capacity to provide a pertinent conversational analysis [21]. They argued for practical interest and feasibility of CA during routine seizure consultations [20]. Training sessions as described by Jenkins could improve diagnosis of PNES based on linguistic analysis.

In our work and in previous studies, DSA and qualitative analysis were based on interviews in their entirety. Plug et al showed that time point at which patients provide descriptions, vary according to the final diagnosis [22]. Patients with epilepsy rapidly volunteered detailed subjective symptoms, whereas patients with PNES provided detailed description after repeated prompting later during the interview. This feature could have been more taken into account in qualitative analysis than with DSA and could explain the difference in diagnostic accuracy. Systematic conversational analysis at each step (open question, first seizure, worst seizure...) could provide supplementary clues for the diagnosis and improve the diagnostic accuracy of DSA.

Contrary to former studies, we did not find any differences between groups in terms of the use of diagnosis labels or metaphors from the EVENT/SITUATION and the SPACE/PLACE categories. Only the use of metaphors from the AGENT/FORCE category was significantly more important in patients with epilepsy than in patients with PNES. These findings were concordant with those of the Sheffield group [15, 16]. It is interesting to note that patients from our study used fewer metaphors than patients in the English study. In addition to our low sample size, these fewer numbers of metaphors could be due to the shorter length of the interview, a different method of conducting the interview or to differences between English-speaking and French-speaking patients.

It is particularly of importance to note that all these studies were conducted in European countries. In a multicenter international study, Asadi-Pooya et al recently highlighted that the semiological presentation of PNES could vary between continents, whereas the

demographic characteristics and predisposing factors were similar. Cultural, religious and ethnic differences probably explain this observation [23]. Similarly, the communicative style of patients is probably influenced by these differences and linguistic analysis could be not as relevant in continents other than Europe. Additional studies are therefore needed to determine the diagnostic yield of linguistic analysis in these continents.

Limits

The population of our study was small in number and this factor decreases the statistical power of the study. The small sample size was due to the methodology used, which implied diagnosis confirmation by video-EEG recordings of seizures as a gold standard test. Moreover, our population stemmed from a tertiary center, which may affect the external validity of our results. Patients were mostly excluded because no seizure was recorded during the monitoring. We suppose that seizures were less frequent for these patients but this question was not investigated in our study. Differences between groups concerning conversational features may have been exacerbated by long disease duration, more frequent psychological comorbidities, drug therapies, diagnostic wanderings, multiple medical interviews and tests experienced by patients. However, other work from the Sheffield teams demonstrated the diagnosis pertinence of CA during routine consultations in a seizure clinic [20]. These results, although diagnosis was not confirmed by video-EEG recordings, suggested that CA could also be useful for patients that presented few events or with shorter disease duration.

Patients with PNES are known to present trauma history, anxiety, mood disorders or personality disorders more frequently than patients with epilepsy [24]. Attachment or coping styles could also be different between these two groups. How the linguistic and interactional capacities would be influenced by these differences was not investigated in our work and our study was not designed to explore this potential interaction. Some studies suggested that linguistic and interactional capacities could provide information concerning coping behavior [15, 19]. For instance, authors suggested that "focusing resistance" could reflect an avoidant coping style or dissociative process, which is thought to cause most PNES. This interaction could therefore be useful in clinical practice in order to understand potential resources of patients and to guide patient care and psychological therapy.

5. CONCLUSIONS

As in the German, English and Italian studies, our work confirms the diagnostic value of conversational analysis on French-speaking patients in order to differentiate patients presenting seizures. Inter-rater and inter-study reliability support the fact that conversational analysis is not only a hunch but can be standardized and be useful for PNES diagnosis. Conversational analysis should probably be placed in the clusters of arguments to improve the confidence in PNES diagnosis. Our study also confirms that neurologists can perform relevant and reliable conversational analyses. Conversational analyses, after training sessions, could be useful in countries or in neurology departments where prolonged video-EEG monitoring is not easily available. The diagnostic value of conversational analysis in "first seizure clinic", for patients presenting both PNES and ES or in other continents needs further investigations and the exact place of CA within algorithms of clinical decision-making has to be specified.

REFERENCES

- [1] World Health Organization. International statistical classification of diseases and related health problems. World Health Organization; 2015.
- [2] American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 5th Edition: DSM-5. 5 edition. Washington, D.C: American Psychiatric Publishing; 2013.
- [3] M. Beghi, F. Peroni, C.M. Cornaggia. Reply to: We need a functioning name for PNES: Considering dissociative seizures. *Epilepsy & Behavior* 2020; [Available online ahead of print at. <https://doi.org/10.1016/j.yebeh.2020.107084>]
- [4] Reuber M, Baker GA, Gill R, Smith DF, Chadwick DW. Failure to recognize psychogenic nonepileptic seizures may cause death. *Neurology* 2004;62:834–5.
- [5] Devinsky O, Gazzola D, LaFrance WC. Differentiating between nonepileptic and epileptic seizures. *Nat Rev Neurol* 2011;7:210–20.
- [6] Duwicquet C, de Toffol B, Corcia P, Bonnin M, El-Hage W, Biberon J. Are the clinical classifications for psychogenic nonepileptic seizures reliable? *Epilepsy Behav* 2017;77:53-57.
- [7] Gülich E, Schöndienst M. “Das ist unheimlich schwer zu beschreiben.” : Formulierungsmuster in Krankheitsbeschreibungen anfallskranker Patienten: differentialdiagnostische und therapeutische Aspekte. *Psychotherapie & Sozialwissenschaft : Zeitschrift für qualitative Forschung und klinische Praxis* 1999;1.
- [8] Reuber M, Micoulaud-Franchi J-A, Gülich E, Bartolomei F, McGonigal A. Comment ce que disent les patients peut nous renseigner sur leurs crises non épileptiques psychogènes. *Neurophysiologie Clinique/Clinical Neurophysiology* n.d.
- [9] Schwabe M, Howell SJ, Reuber M. Differential diagnosis of seizure disorders: a conversation analytic approach. *Soc Sci Med* 2007;65:712–24.
- [10] Cornaggia CM, Gugliotta SC, Magaüdda A, Alfa R, Beghi M, Polita M. Conversation analysis in the differential diagnosis of Italian patients with epileptic or psychogenic non-epileptic seizures: a blind prospective study. *Epilepsy Behav* 2012;25:598–604.
- [11] Papagno C, Montali L, Turner K, Frigerio A, Sirtori M, Zambrelli E, et al. Differentiating PNES from epileptic seizures using conversational analysis. *Epilepsy*

- Behav 2017;76:46–50.
- [12] Reuber M, Monzoni C, Sharrack B, Plug L. Using interactional and linguistic analysis to distinguish between epileptic and psychogenic nonepileptic seizures: a prospective, blinded multirater study. *Epilepsy Behav* 2009;16:139–44.
 - [13] Schwabe M, Reuber M, Schöndienst M, Gülich E. Listening to people with seizures: how can linguistic analysis help in the differential diagnosis of seizure disorders? *Commun Med* 2008;5:59–72.
 - [14] Plug L, Sharrack B, Reuber M. Seizure, Fit or Attack? The Use of Diagnostic Labels by Patients with Epileptic or Non-epileptic Seizures. *Appl Linguist* 2010;31:94–114.
 - [15] Plug L, Sharrack B, Reuber M. Seizure metaphors differ in patients' accounts of epileptic and psychogenic nonepileptic seizures. *Epilepsia* 2009a;50:994–1000.
 - [16] Plug L, Sharrack B, Reuber M. Metaphors in the description of seizure experiences: Common expressions and differential diagnosis. *Language and Cognition* 2011;3:209–33.
 - [17] Surmann V. Anfallsbilder: metaphorische Konzepte im Sprechen anfallskranker Menschen. thesis. 2005.
 - [18] Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 1977;33:159.
 - [19] Monzoni C, Reuber M. Conversational displays of coping resources in clinical encounters between patients with epilepsy and neurologists: a pilot study. *Epilepsy Behav* 2009;16:652–9.
 - [20] Jenkins L, Cosgrove J, Chappell P, Kheder A, Sokhi D, Reuber M. Neurologists can identify diagnostic linguistic features during routine seizure clinic interactions: results of a one-day teaching intervention. *Epilepsy Behav* 2016;64:257–61.
 - [21] Jenkins L, Reuber M. A Conversation Analytic Intervention to Help Neurologists Identify Diagnostically Relevant Linguistic Features in Seizure Patients' Talk. *Research on Language and Social Interaction* 2014;47:266–79.
 - [22] Plug L, Sharrack B, Reuber M. Conversation analysis can help to distinguish between epilepsy and non-epileptic seizure disorders: a case comparison. *Seizure* 2009b;18:43–50.

- [23] Asadi-Pooya AA, Valente K, Restrepo AD, D' Alessio L, Homayoun M, Bahrami Z, Alessi R, Paytan AA, Kochen S, Myers L, Sawchuk T, Buchhalter J, Taha F, Lazar LM, Pick S, Nicholson T. Adult-onset psychogenic nonepileptic seizures: A multicenter international study. *Epilepsy Behav* 2019;98:36-39.
- [24] Hingray C, Biberon J, El-Hage W, de Toffol B. Psychogenic non-epileptic seizures (PNES). *Rev Neurol (Paris)* 2016;172:263–9.

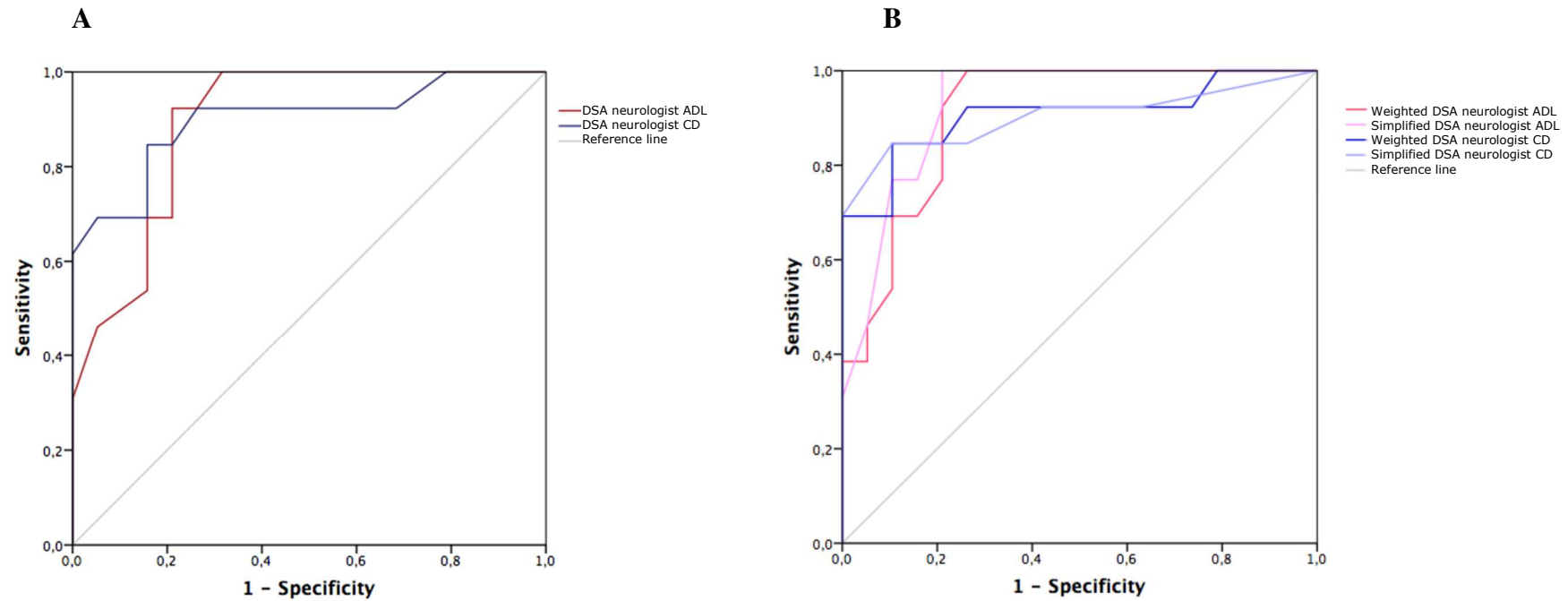


Figure 1: ROC curves for original (A) diagnostic score aid (DSA), weighted and simplified DSA (B). Red curves concern rater 1 (ADL) and blue curves concern rater 2 (CD).

Tables

- **Table 1:** Demographic and conversational characteristics
- **Table 2:** Diagnostic value and interrater reliability of individual items
- **Table 3:** Rating of patients on individual items (rater 1/rater 2)

	Patients with PNES (n=13)	Patients with epilepsy (n =19)	p
Sex ratio, M/F	0/13	10/9	0.0016
Age, years	32.7 ± 3.5	43.2 ± 3.6	0.048
Duration, years	8.2 ± 1.8	4.5 ± 1	0.079
Number of years of school completed, years	10.4 ± 0.7	12.8 ± 0.6	0.03
Current Psychotropic drugs use	62%	5%	0.001
Current Antiepileptic drugs use	62%	79%	0.29
Duration of monitoring, days	4.5 ± 0.2	4.7 ± 0.9	0.20
Duration of interview, minutes	17.2 ± 1.8	14.7 ± 0.8	0.17
DSA Rater 1 (ADL)	-4.5 ± 1.7	9.2 ± 1.8	2.10⁻⁴
DSA Rater 2 (CD)	-6.2 ± 2.2	7.7 ± 1.7	2.10⁻⁴

Table 1: Demographic and conversational characteristics. DSA: Diagnostic Score Aid.

Item	Significance of correct classification, p (chi-squared test)		Kappa (raters 1 and 2)
	Rater 1	Rater 2	
1	n.s.	0.00078	0.0455
2	n.s.	n.s.	0.2668
3	n.s.	n.s.	0.2615
4	0.0019	0.0014	0.4028
5	0.00057	0.0012	0.3905
6	n.s.	0.0021	0.274
7	0.0011	0.00058	0.3672
8	n.s.	n.s.	0.2358
9	n.s.	n.s.	0.2558
10	0.0013	n.s.	0.3918
11	0.000046	n.s.	0.422
12	0.00073	0.00094	0.5669
13	n.s.	n.s.	0.3618
14	n.s.	n.s.	0.3855
15	0.000063	0.00046	0.4173
16	n.s.	n.s.	0.2448
17	0.0014	n.s.	0.0533

Table 2: Diagnostic value and interrater reliability of individual items. n.s.: non-significant.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P28	P29	P30	P31	P32
I1	O/N	O/E	O/N	E/E	N/N	E/N	E/N	O/N	E/E	O/E	E/E	E/O	E/O	E/E	E/E	N/N	O/N	E/N	O/N	E/E	E/O	E/O	E/E	E/N	E/N	E/N	O/N	E/E	E/E	E/O	E/N	E/O
I2	N/O	O/E	N/O	E/E	O/O	E/O	E/E	O/N	E/E	E/E	E/E	E/E	E/O	E/E	E/E	N/O	O/O	E/O	O/O	E/E	E/E	O/O	E/O	O/N	O/O	E/N	O/O	E/E	E/O	E/E	E/O	E/O
I3	N/O	O/O	O/O	E/O	N/O	N/N	O/O	E/N	E/O	E/E	O/O	E/E	N/O	O/O	E/E	N/O	E/N	O/O	N/O	E/E	O/O	N/O	O/O	O/N	O/N	O/N	N/O	E/E	O/O	O/O	E/E	N/O
I4	N/N	O/E	N/N	E/E	O/N	N/N	O/N	N/N	E/E	E/E	E/E	O/N	E/E	E/E	E/E	N/N	N/O	E/E	N/O	E/E	O/E	N/E	E/E	O/N	E/N	O/O	E/E	E/E	E/E	E/E	E/O	E/O
I5	N/O	O/E	O/O	E/E	O/O	O/N	O/O	N/N	E/O	O/E	E/E	O/O	O/E	E/E	E/E	N/O	O/N	E/E	O/O	E/E	O/E	O/O	E/E	O/O	O/N	O/N	O/E	E/E	E/E	E/E	E/E	O/O
I6	N/N	E/E	E/O	E/E	E/N	O/O	O/O	N/O	E/O	E/E	E/E	E/O	E/O	E/E	E/E	O/O	E/O	E/E	E/E	E/E	E/E	E/O	E/E	E/N	E/N	E/N	E/O	E/E	E/E	E/E	E/O	E/E
I7	N/O	N/O	N/O	E/E	N/O	N/O	N/O	N/N	E/E	O/E	E/E	O/O	O/O	O/E	E/E	N/N	N/N	O/O	N/N	E/E	O/E	N/O	E/O	N/N	N/N	N/N	N/O	E/E	O/E	E/E	E/O	O/O
I8	N/N	O/O	N/O	E/E	N/N	N/O	E/E	N/N	E/E	O/E	E/E	O/O	O/N	O/E	E/O	N/N	O/N	O/N	O/O	E/E	E/E	O/O	E/O	O/N	O/N	O/N	O/O	E/O	O/E	E/O	E/O	O/O
I9	N/O	N/O	N/N	E/E	N/N	N/N	N/O	N/N	E/E	O/E	E/E	E/E	E/O	E/E	E/E	N/N	O/N	E/E	E/O	O/E	E/E	O/E	E/E	O/E	N/N	O/N	O/N	E/O	E/O	E/E	E/N	E/E
I10	N/N	O/E	O/N	E/E	O/N	N/N	N/N	N/N	E/O	E/E	E/E	N/N	E/O	E/E	E/E	O/N	O/N	E/O	O/O	E/E	O/E	O/O	E/E	N/N	O/O	N/N	E/O	E/O	E/E	E/E	O/N	O/E
I11	N/O	E/E	O/O	E/O	O/O	O/N	N/N	O/N	E/O	E/E	E/E	O/N	E/O	E/E	E/E	O/N	O/O	E/E	O/O	E/E	O/E	O/E	E/E	O/O	O/O	O/O	O/O	E/E	E/E	E/E	O/O	E/O
I12	O/N	O/E	O/O	E/E	N/O	N/O	O/O	N/N	E/E	O/E	E/E	O/O	O/O	E/E	E/E	N/N	N/N	O/O	O/O	E/E	E/E	O/O	O/O	N/N	N/N	O/N	O/E	E/E	O/E	E/E	E/O	O/O
I13	N/N	O/O	N/O	E/E	N/O	N/N	O/O	N/N	E/E	E/E	E/E	O/O	O/O	N/E	N/E	N/O	N/N	E/O	N/O	E/E	E/E	N/O	E/O	O/O	O/O	O/N	O/O	O/O	E/O	E/O	E/O	E/O
I14	N/N	N/E	O/N	E/E	N/O	N/N	N/N	O/N	O/E	O/E	E/E	O/O	O/O	E/E	E/E	N/N	O/O	E/O	O/O	O/E	O/E	O/O	E/E	O/N	N/N	E/O	E/E	O/E	O/E	O/O	O/O	O/O
I15	N/N	N/E	N/O	E/E	N/O	N/O	E/O	N/O	E/E	E/O	E/E	O/N	E/E	E/E	O/E	N/N	O/O	E/E	O/O	E/E	O/E	E/E	E/E	O/O	N/O	N/N	O/O	E/E	O/O	E/E	E/O	E/E
I16	N/O	N/E	O/E	E/E	N/O	N/O	N/O	N/O	E/O	O/O	E/E	E/O	O/O	O/E	E/E	N/N	N/N	E/E	N/E	E/E	E/E	N/E	E/E	N/N	O/E	O/N	N/E	E/E	O/O	E/O	E/E	E/E
I17	N/O	N/O	O/O	O/O	N/O	N/O	O/N	N/N	E/O	N/O	N/O	N/N	N/O	O/O	O/E	N/O	N/O	N/O	N/N	O/O	N/O	N/O	N/O	N/N	N/O	N/N	N/O	E/O	O/O	E/O	O/O	O/O
DSA diagnosis	N/N	N/E	N/N	E/E	N/N	N/N	N/N	N/N	E/E	E/E	E/E	E/N	E/E	E/E	E/E	N/N	N/N	E/E	N/N	E/E	E/E	N/E	E/E	N/N	N/N	N/N	N/E	E/E	E/E	E/E	E/O	E/E
Final diagnosis	N	E	E	E	N	N	E	N	E	E	E	N	E	E	E	N	N	E	N	E	N	E	E	N	N	N	N	E	E	E	E	E

Table 3: Rating of patients on individual items (rater 1/rater 2). I: Item; P: Patient; E: Epilepsy/in favor of epilepsy; N: PNES/in favor of PNES. Items with agreement are highlighted in blue (light blue when rating was uncertain; navy blue when both raters agreed in favor of PNES; intermediate blue when both raters agreed in favor of epilepsy).