Article:

Pevy, N. orcid.org/0000-0001-5263-2753, Christensen, H. orcid.org/0000-0003-3028-5062, Walker, T. orcid.org/0000-0002-2583-7232 et al. (1 more author) (2023) Differentiating between epileptic and functional/dissociative seizures using semantic content analysis of transcripts of routine clinic consultations. Epilepsy & Behavior, 143. 109217. ISSN 1525-5050

https://doi.org/10.1016/j.yebeh.2023.109217

Epilepsy & Behavior 143 (2023) 109217

Contents lists available at ScienceDirect

Epilepsy & Behavior

journal homepage: www.elsevier.com/locate/yebeh

Differentiating between epileptic and functional/dissociative seizures using semantic content analysis of transcripts of routine clinic consultations

Nathan Pevy [a,*], Heidi Christensen [b], Traci Walker [c], Markus Reuber [d]

[a] Department of Neuroscience, The University of Sheffield, United Kingdom
[b] Department of Computer Science, The University of Sheffield, United Kingdom
[c] Division of Human Communication Sciences, The University of Sheffield, United Kingdom
[d] Academic Neurology Unit, University of Sheffield, United Kingdom

ARTICLE INFO

Article history:
Received 10 February 2023
Revised 2 April 2023
Accepted 4 April 2023
Available online 27 April 2023

Keywords:
Epilepsy
Functional Seizures
Differential Diagnosis
Semantic
Automated Analysis of Speech
Clinical

ABSTRACT

The common causes of Transient Loss of Consciousness (TLOC) are syncope, epilepsy, and functional/dissociative seizures (FDS). Simple, questionnaire-based decision-making tools for non-specialists who may have to deal with TLOC (such as clinicians working in primary or emergency care) reliably differentiate between patients who have experienced syncope and those who have had one or more seizures but are more limited in their ability to differentiate between epileptic seizures and FDS. Previous conversation analysis research has demonstrated that qualitative expert analysis of how people talk to clinicians about their seizures can help distinguish between these two TLOC causes. This paper investigates whether automated language analysis - using semantic categories measured by the Linguistic Inquiry and Word Count (LIWC) toolkit - can contribute to the distinction between epilepsy and FDS. Using patient-only talk manually transcribed from recordings of 58 routine doctor-patient clinic interactions, we compared the word frequencies for 21 semantic categories and explored the predictive performance of these categories using 5 different machine learning algorithms. Machine learning algorithms trained using the chosen semantic categories and leave-one-out cross-validation were able to predict the diagnosis with an accuracy of up to 81%. The results of this proof of principle study suggest that the analysis of semantic variables in seizure descriptions could improve clinical decision tools for patients presenting with TLOC.

1. Introduction

Patients who experience Transient Loss of Consciousness (TLOC) are frequently asymptomatic on presentation to health services. This means that investigation and treatment are largely guided by a clinician taking and interpreting the history available from the patient (and TLOC witnesses if possible) [1]. The diagnostic process, therefore, relies heavily on the expertise of the history-taker. Given that patient typically first present with TLOC in non-specialist medical settings such as primary and emergency care, it is not surprising that the median initial misdiagnosis rates between epilepsy, syncope, and functional/dissociative seizures (FDS), the three most common causes of TLOC, is around 20% across different studies (range 2–71%) [2]. Patients who receive the wrong initial diagnosis may be referred for inappropriate tests and receive the wrong treatment. Because syncope may be caused by potentially lethal arrhythmias and uncontrolled epilepsy is associated with a risk of Sudden Unexpected Death in Epilepsy (SUDEP), early diagnostic errors and diagnostic delay can be fatal.

The initial management of patients presenting with TLOC could be improved by reliable decision-making aids or stratification tools guiding non-specialists to provide management that is appropriate for the cause of TLOC. However, to date, only a limited number of potential clinical decision tools have been proposed for this setting and none have been shown to differentiate reliably between the three common causes of TLOC [3]. Historical features have been shown to aid the differentiation between syncope and seizures [4,5,6]. Therefore, the overall performance of these tools is

Abbreviations: TLOC, Transient Loss of Consciousness; FDS, functional/dissociative seizures; LIWC, Linguistic Inquiry and Word Count; PWE, People with epilepsy; PWFDS, People with FDS; RBF, Radial Basis Function; SVM, Support Vector Machine; VA, Virtual agent.

* Corresponding author at: Regent Court, The University of Sheffield, 211 Portobello Street, Sheffield S1 3JD, United Kingdom.

E-mail address: n.pevy@sheffield.ac.uk (N. Pevy).

https://doi.org/10.1016/j.yebeh.2023.109217
1525-5050/

particularly diminished by the challenging differentiation between people with epilepsy (PWE) and people with FDS (PWFDS) [4].

We hypothesized that it may be possible to improve the ability of questionnaire-based stratification tools to distinguish between PWE and PWFDS by incorporating an automated analysis of the language used by patients answering questions about their seizures. Previous studies based on the qualitative analysis of interactions between patients with seizures and clinicians by experts in conversation analysis identified a number of differences in diagnostic value. Differences have been observed at the levels of language use (linguistic level), topic choice (topical level), and relation to turn-taking behavior (interactional level). For example, at the linguistic level, PWE and PWFDS display different amounts of formulation effort while describing their seizures (including more hesitations (''um''), repetitions (''I I felt''), restarts (''It was like I felt confused''), and reformulations of descriptions) [7,8]. Their accounts also tend to differ in the use of metaphors that are used to describe seizure experiences [9], diagnostic labels [10], and the type of third-party references [11]. At the topical level, they vary in the extent to which they describe subjective symptoms. Interactional observations include the degree of resistance patients may exhibit when they are asked to focus on specific memorable single seizure experiences (such as the first, last, or worst seizure) [7,8]. Expert evaluations based on linguistic profiles capturing these observations (and no additional clinical data about the patient's seizure disorders) achieved an accuracy of 80–90% in the differentiation of epilepsy and FDS [12–17]. These findings highlight the potential of automated language analysis in the differentiation between PWE and PWFDS, an idea supported by previous research using semi-automated methods to detect differences in formulation effort in seizure descriptions provided by patients in these two diagnostic groups [18].

Recordings of routine medical encounters represent an effective way to explore the feasibility of predicting TLOC diagnoses using an automated analysis of language because the interaction often includes a lot of information about the patient, their medical history, and their seizure(s). During the history-taking phase of the medical interactions [19], the doctor attempts to understand who the patient is, how the patient's behavior may be related to the medical condition, what the illness is, how it has developed, and whether there are other illnesses present [20]. This interaction may give rise to other linguistic differences between PWE and PWFDS based upon the different aetiologies of and experiences associated with each health condition: FDS are automatic and uncontrolled responses to emotions, thoughts, sensations, or situations perceived as threatening [21] and PWFDS report higher levels of general psychopathology [22], are more likely to experience panic symptoms during a seizure [22] and catastrophize life experiences [23] than PWE. This means that the history-taking phase of clinical interactions can provide a lot of information that is suitable for exploring whether there are linguistic variations in the spoken responses of patients that can be incorporated into an automated analysis of language focussing on seizure descriptions.

1.1. Linguistic Inquiry and Word Count

Linguistic Inquiry and Word Count (LIWC) is an application that processes text and measures the proportion of words that correspond to different semantic categories [24]. The term ''semantic'' denotes a sub-branch of linguistics related to meaning in language. In this paper, ''semantic'' is used to describe language differences that center around the use of different words, and semantic categories describe categories that contain multiple words that share similar or related meanings based on some arbitrary concept. The words in each semantic category were generated using previous research. This research involved the evaluation of the validity and reliability of each category by a panel of judges and by exploring the frequency of category activations and similarity with human ratings across a broad range of text documents [25–27]. The application has been used to compare semantic differences between people with and without various psychiatric conditions, for example, social anxiety, borderline personality disorder, depression, and Alzheimer's disease [28–31]. Moreover, the application of LIWC to semi-structured interviews with PWE and PWFDS found that PWE used significantly more instances of ''she/he'', ''we'' and family references compared to PWFDS [32]. These findings demonstrate the potential of LIWC to detect linguistic patterns associated with a specific health condition.

1.2. Aim

The objective of this paper is to explore how effectively semantic categories from the LIWC application can predict a diagnosis of epilepsy or FDS when applied to the history-taking phase of routine seizure clinic encounters. The focus was on semantic categories that align with the linguistic differences observed in previous differential diagnostic research.

2. Method

2.1. Data

The dataset consisted of 58 manually transcribed recordings of encounters involving patients and neurologists in a routine seizure clinic setting. The recordings were originally collected for previously published conversation analysis studies [11,33]. The neurologists in one group of interviews took part in a training program aiming to enhance their ability to pick up interactional and linguistic differential diagnostic features during their clinic interactions with patients. During the training, they were instructed to ask participants about their first, most recent, and worst seizure, and encouraged not to interrupt patients during their narratives [33]. The neurologists in the second group had received no instructions [11]. Recordings were included in the analysis if a final medical diagnosis had been confirmed by a review of all clinical data by an epileptologist or the diagnosis had been confirmed by the video-EEG recording of a typical seizure. Patients were only included if their final diagnosis was one of epilepsy (N = 37) or FDS (N = 21). Two stages of the medical encounters, establishing the reason for the visit and the history taking [19] were manually extracted from the whole interaction before the doctor started talking about the diagnosis.

2.2. Linguistic Inquiry and Word Count

The recordings of the doctor-patient interactions were manually transcribed. A manually created algorithm was used to extract the text that corresponded to all patient turns during the target subsection of the interaction. The most recent version of the LIWC application has a dictionary of almost 6400 words across 93 different semantic categories [34]. We followed a hypothesis-driven approach based on previous research focussing on the differentiation of patients' epileptic seizures or FDS to identify 21 potentially relevant semantic LIWC categories and used these categories to build an automatic classification algorithm.

We included five semantic categories that measured the frequency of social words (''We'', ''She/He'', ''Family'', ''Social'', and ''Affiliation'') because previous research has demonstrated group differences for ''We'', ''She/He'' and ''Family'' [32] and PWFDS are more likely to make catastrophizing third party references, whereas PWE are more likely to make normalizing third party references [11]. The categories "Risk" and "Reward" were included to detect differences in anxiety, catastrophization, and avoidance [23,35]. "Cause" was selected because people with FDS have previously reported that seizures cause greater disruption to their everyday life compared to people with epilepsy [36]. Seven categories were included to measure differences in emotive language ("Emotional tone", "Affect", "Positive emotions", "Negative emotions", "Anxiety", "Anger", and "Sad") to capture differences associated with levels of general psychopathology [21] and any associations between emotional processing and experiences and FDS [21,35,36]. The categories "Focus Present" (measuring the number of present tense verbs) and "Quantifiers" (words that express quantity) were included because PWFDS display an increased tendency to talk about seizures in general rather than focussing on the description of a single seizure experience in the past tense, for example "in most of my seizures" and "in all of my seizures" [7,8]. Two categories ("Certainty" and "Tentativeness") were used to measure varying levels of certainty associated with descriptions, for example, increased tentativeness during the description of subjective symptoms for PWE and increased tendency to make absolute negations for PWFDS [7,8]. Finally, the two categories were included to measure different metaphoric conceptualizations ("Space" and "Power") frequently observed in descriptions of seizures [9] and the increased focus on the circumstances and consequences of seizures ("Space") for PWFDS [9].

We conducted group comparisons for each semantic category to gain insight into which categories may be the most effective for future research. The Shapiro-Wilk test was used to test the normality of each variable, and the Levene test was used to test for homogeneity of variance. Group differences for each semantic category were calculated using an independent T-Test or Mann-Whitney U Test depending on whether the variables were normally distributed and if the samples had equal variances. No correction was made for multiple comparisons because of the exploratory aim of this study. Corrections for multiple comparisons would have increased the risk of making a type 1 error, which could prevent future researchers from exploring variables that may improve the predictive performance of machine learning models.

2.3. Demographic and general speech differences

A chi-squared test of independence was performed to investigate whether there was a significant difference in gender between PWE and PWFDS. There was no significant difference between these variables $X^2 = (1, N = 58) = 0.431$, $p = 0.511$.

A Mann-Whitney U test was used to evaluate whether there was a difference in the word count and total number of unique words per transcript for PWE and PWFDS. We found that PWFDS spoke more words in general (median = 1445) compared to PWE (median = 1115), $U = (1, N = 58) = 275$, $p < 0.05$. Moreover, PWFDS spoke more unique words (median = 351) compared to PWE (median = 271), $U = 230.5$, $p < 0.01$.

2.4. Classification

Different machine learning models are often more effective on different data types and classification tasks. We explored the classification performance of multiple machine learning classifiers to identify the algorithms most suited to the differentiation between people with epilepsy and FDS using an automated analysis of language because the identification of the best-performing algorithm could guide future research. The classification performance of the semantic categories was evaluated using five different machine learning models trained using the sci-kit learn toolkit in Python [37]: Random Forest [38], Support Vector Machine with either a linear or Radial Basis Function (RBF) kernel [39], Logistic Regression, and K-Nearest Neighbour [40]. Each model was trained using "leave-one-out" cross-validation and a nested search for the optimum hyper-parameters to prevent overfitting [41]. The importance of each feature for the best-performing machine learning models was determined using an ablation analysis where each feature was removed, and the classification accuracy of the model was recalculated. Features that resulted in the largest decrease in classification accuracy were considered the most important.

3. Results

3.1. Comparison of the semantic categories

There were significant between-group differences in 11 of the 21 LIWC variables (Table 1). The semantic categories with a significant group differences were "Negative emotions", "Emotional tone", "Quantifiers", "Focus present", "Sad", "Reward", "Anger", "Family", "Power", "Cause", and "Affiliation".

3.2. Classification performance

The results of the classification analysis demonstrate a large degree of variation between the five classifiers (Fig. 1). The best performance was demonstrated with the three non-linear classifiers, which were the K-Nearest Neighbour classifier (accuracy = 81%), the support vector machine with an RBF kernel (accuracy = 77.6%), and the Random Forest algorithm (accuracy = 69%). The two classifiers that use a linear operation performed less effectively (Support Vector Machine with a linear kernel, accuracy = 67.2%, and the Logistic Regression algorithm, accuracy = 62.1%). All classifiers were better at identifying individuals with FDS because they demonstrated a greater specificity (70.3–83.8%) than sensitivity (42.9–76.2%).

3.3. Most important features

The most important features were calculated for the K-Nearest Neighbour model and the SVM model with the RBF kernel. The top nine most important features determined from the ablation analysis were: "Focus present tense", "Emotional tone", "Tentativeness", "Quantifiers", "Reward", "Social", "Affect", "We", and "He/She". The subsequent four features ("Positive emotion", "Family", "Cause", and "Affiliation") had the same importance score. Furthermore, four out of the top nine features were among the features within our analysis showing significant group differences in the single-item comparisons, excluding "Tentativeness", "Social", "Affect", "We", and "He/She". The support vector machine with an RBF kernel appeared to be the most stable machine learning algorithm because there was no change in the accuracy of the model when 13 out of the 21 features were removed independently, suggesting that this model is less reliant on individual features, whereas k-nearest neighbors demonstrated the most and largest changes due to the removal of features (Fig. 2).

4. Discussion

This paper explored the feasibility of differentiating between spoken accounts of epileptic and functional/dissociative seizures (FDS) by comparing semantic differences between the words patients use in routine medical encounters. Using 21 semantic categories measured using the Linguistic Inquiry and Word Count (LIWC) application, we were able to accurately predict the diagnosis of epilepsy or FDS in up to 81% of cases. The best classification performance was achieved using the K-Nearest Neighbour and SVM-RBF algorithms. These findings support previous research

Table 1
Means and standard deviations (parametric) or medians and interquartile ranges (non-parametric) of the percentage of words per semantic category for PWE and PWFDS. The t-value and p-value are reported for each group comparison unless otherwise specified.

Semantic category	PWE	PWFDS	T value	P-value
Negative emotions †	1.04 (1.17)	1.73 (0.57)	M = 210	P < 0.01
Emotional tone †	28.8 (28.35)	21.95 (7.83)	M = 216	P < 0.01
Quantifiers †	1.7 (0.91)	2.19 (0.4)	M = 229	P < 0.01
Focus present †	10.5 (3.54)	12.94 (2.64)	M = 230.5	P < 0.01
Sad †	0.21 (0.32)	0.38 (0.27)	M = 231.5	P < 0.01
Reward	1.035 (0.61)	1.399 (0.44)	T = −2.356	P < 0.05
Anger †	0.05 (0.2)	0.18 (0.23)	M = 271.5	P < 0.05
Family †	0.26 (0.37)	0.35 (0.69)	M = 272	P < 0.05
Power	1.189 (0.55)	1.489 (0.38)	T = −2.181	P < 0.05
Cause	1.184 (0.66)	1.568 (0.59)	T = −2.179	P < 0.05
Affiliation †	0.52 (0.52)	0.71 (0.33)	M = 285.5	P < 0.05
Space	5.517 (1.74)	6.254 (1.05)	T = −1.736	P = 0.08
Social	5.638 (2.07)	6.7 (1.97)	T = −1.879	P = 0.07
Risk †	0.39 (0.39)	0.48 (0.41)	M = 286.5	P = 0.05
We †	0.06 (0.2)	0.14 (0.29)	M = 291	P = 0.05
SheHe †	0.72 (1.1)	0.9 (1.09)	M = 289	P = 0.05
Affect	2.76 (0.91)	3.09 (0.64)	T = −1.431	P = 0.16
Positive Emotions †	1.54 (0.63)	1.4 (0.42)	M = 313.5	P = 0.11
Anxiety †	0.29 (0.46)	0.31 (0.38)	M = 318.5	P = 0.13
Certain	1.62 (0.77)	1.489 (0.59)	T = 0.663	P = 0.51
Tentativeness †	2.78 (1.34)	2.65 (1.47)	M = 388	P = 0.5

† - Mann Whitney U, median, and interquartile ranges are reported when variables were not normally distributed or the homogeneity of variance assumption was violated.

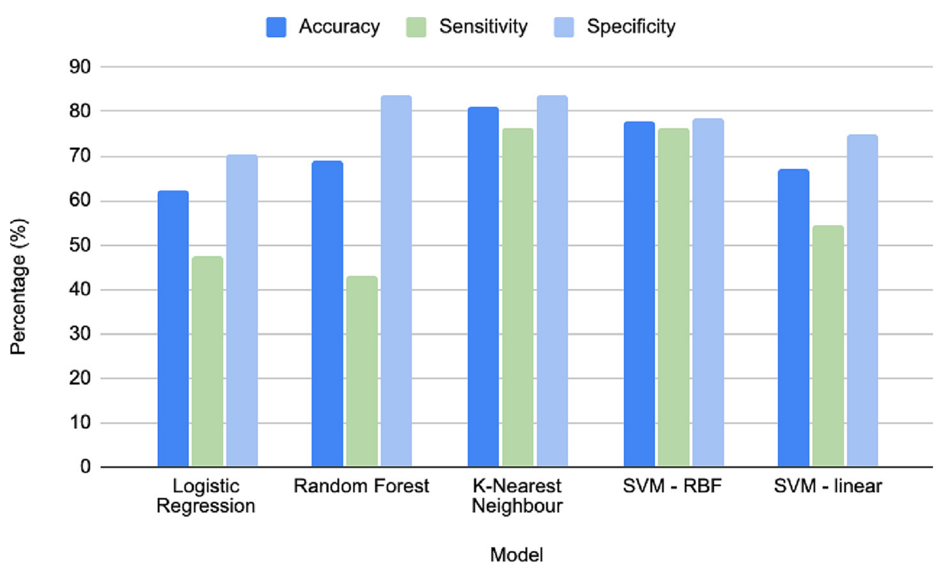

Fig. 1. A comparison of the performance (accuracy, sensitivity, and specificity) of each of the machine learning algorithms using all 21 semantic categories.

suggesting that linguistic features extracted from spoken seizure descriptions could be used for diagnostic predictions [18] and suggest that semantic word categories could be effective features to incorporate into an automated analysis of spoken descriptions of seizures and the medical history [18].

In addition to exploring the predictive performance of the semantic categories, we explored the contribution of independent features to determine the most effective features for the classification task. Although we observed a significant group difference for 11 of the 21 semantic categories, only four of these variables were in the top ten performing features in the two best-performing models. This demonstrates that semantic categories may be effective for predicting the diagnosis without distinct group differences, potentially due to complex relationships between the different semantic categories that can be captured by these machine learning algorithms. Future researchers may wish to use these findings to select a lower number of features to incorporate into their model.

A comparison of different machine learning algorithms demonstrated that the non-linear models (K-Nearest Neighbours, Support Vector Machine with RBF kernel, and Random Forest) outperformed the models that used a linear decision boundary or hyperplane (Logistic Regression and Support Vector Machine with linear kernel). This finding suggests the presence of non-linear group differences in language between people with epilepsy or FDS. Future research should consider this when designing similar machine-learning pipelines.

This research could be advanced by exploring the potential of a fully automated system where people speak with a virtual agent (VA) and their responses are automatically analyzed in real-time to generate a diagnostic prediction. A similar VA is currently being developed as a stratification tool for patients presenting with memory problems [41]. The present proof-of-principle analysis provides a baseline for how well a computer may analyze the spoken language used by patients during the history taking in routine medical consultations. Future research may explore whether better

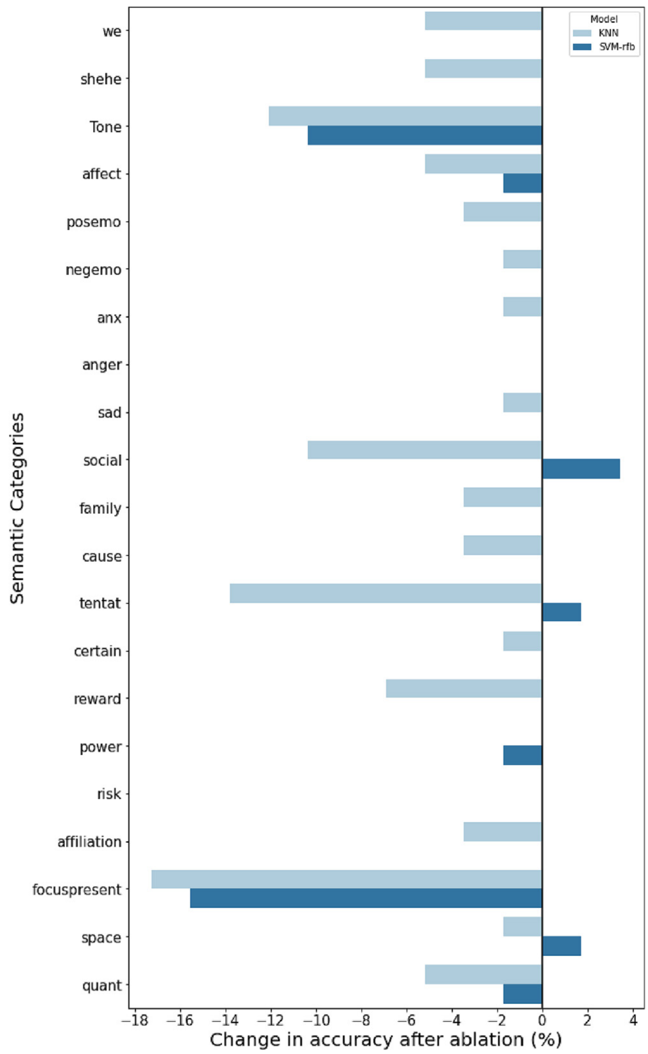

Fig. 2. The change in classification accuracy (x-axis) for each classification model (hue) when each feature (y-axis) is removed from the analysis independently.

diagnostic differentiation could be achieved if spoken seizure descriptions from patients are sampled in more standardized ways, for instance by a VA.

An interesting observation from this analysis is that PWFDS typically said more in their interactions with clinicians than people with epilepsy. This finding was unexpected because previous research has found that PWFDS typically provide less detailed descriptions of their seizures and are more likely to use complete negations instead of describing their seizure experiences more precisely (e.g. statements like "I don't remember anything") [7,8]. One potential explanation is that the wider history-taking procedure also involves conversations about other areas of the patient's health, for example, conversations about the consequences of seizures, the impact they have on the patient's life, potential causes of the seizure, and information about previous medical interactions and other comorbid health conditions. This finding may be important for designing a fully automated system because it demonstrates that PWFDS can have a lot to say during medical interactions and that questions focussing on what happened during the seizure may not capture all the variations in the responses that allow these semantic categories to perform effectively in the machine learning models.

4.1. Limitations

One limitation of this analysis is that the LIWC categories were designed to measure these semantic concepts broadly and are not tailored to seizure consultations. They are not able to measure important semantic categories associated with seizure descriptions, for example, the label used for the chief complaint [10] or descriptions of subjective symptoms [7,8]. They contain many keywords unrelated to seizure consultations. It may be possible to generate semantic categories that are customized for seizure consultations that are better able to differentiate between PWE or PWFDS, but future research would need a larger dataset to detect the broad range of words used for these semantic categories. A second limitation is that the questions that the neurologist asked the patient were not standardized. Although some of the neurologists had received instructions about what questions to ask patients as part of the original research project [33], these instructions did not apply to the whole history-taking procedure and the neurologists whose consultations were studied in the other project had received no instructions [11]. Future research should explore semantic differences in interactions where every participant is asked the same question because this may change what words people use in the interaction. A third limitation of this analysis is that it focuses on independent words and does not consider the wider context of keywords within the talk. The LIWC may not be as effective at identifying semantic constructs compared to human raters because people are able to label a whole segment of text as corresponding to a construct, whereas LIWC only detects the keywords from that segment [27]. There are more complex, nonlinear, machine learning algorithms that can process a segment of text rather than a single word, for example, recurrent neural networks with long short-term memory [42], that could be used to overcome these limitations. However, these methods typically require larger datasets to be effectively used. Finally, this analysis uses manual transcripts instead of automatic speech recognition. Although this has allowed us to test the proof-of-principle of this method, automatic speech recognition would be required for an automatic stratification tool and a small proportion of words will be misidentified due to the associated word error rate, which may change the predictive accuracy of this method.

5. Conclusion

These findings evidence that semantic differences in the contributions of patients with epilepsy or FDS during the history-taking phases of medical consultations can be identified automatically and may have differential diagnostic value. In this proof of principle study, these differences predicted the diagnosis with a good level of accuracy using non-linear machine learning classifiers. Our results support previous qualitative and quantitative research demonstrating that language can be analyzed to predict the cause of seizures and provides further evidence to suggest that automated analysis of language could improve the challenging differentiation between epilepsy and FDS. These findings are relevant for researchers and clinicians who are aiming to develop clinical decision tools for TLOC. Our findings suggest that incorporating methods of automatically collecting and evaluating information related to TLOC from patients may support the identification of the different health conditions associated with TLOC. These clinical decision tools could help to stratify patients presenting with TLOC to ensure that they are referred to the most appropriate further specialist assessment and treatment pathways and reduce the risk of diagnostic delay and misdiagnosis.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors gratefully acknowledge the support of this study by Epilepsy Research UK in the form of a PhD studentship, grant number 160296-1.

References

[1] Plug L, Reuber M. Making the diagnosis in patients with blackouts: it's all in the history. Pract Neurol 2009;9(1):4–15.
[2] Xu Y, Nguyen D, Mohamed A, Carcel C, Li Q, Kutlubaev MA, et al. Frequency of a false positive diagnosis of epilepsy: a systematic review of observational studies. Seizure 2016;41:167–74.
[3] Wardrope A, Newberry E, Reuber M. Diagnostic criteria to aid the differential diagnosis of patients presenting with transient loss of consciousness: A systematic review. Seizure 2018;61:139–48.
[4] Wardrope A, Jamnadas-Khoda J, Broadhurst M, Grünewald RA, Heaton TJ, Howell SJ, et al. Machine learning as a diagnostic decision aid for patients with transient loss of consciousness. Neurol: Clin Pract 2020;10(2):96–105.
[5] Sheldon R, Rose S, Ritchie D, Connolly SJ, Koshman ML, Lee MA, et al. Historical criteria that distinguish syncope from seizures. J Am Coll Cardiol 2002;40(1):142–8.
[6] Schwabe M, Howell SJ, Reuber M. Differential diagnosis of seizure disorders: a conversation analytic approach. Soc Sci Med 2007;65(4):712–24.
[7] Schwabe M, Reuber M, Schondienst M, Gulich E. Listening to people with seizures: how can linguistic analysis help in the differential diagnosis of seizure disorders? Commun Med 2008;5(1):59.
[8] Plug L, Sharrack B, Reuber M. Seizure metaphors differ in patients' accounts of epileptic and psychogenic nonepileptic seizures. Epilepsia 2009;50(5):994–1000.
[9] Plug L, Sharrack B, Reuber M. Seizure, fit or attack? The use of diagnostic labels by patients with epileptic or non-epileptic seizures. Appl Linguis 2010;31(1):94–114.
[10] Robson C, Drew P, Walker T, Reuber M. Catastrophising and normalising in patient's accounts of their seizure experiences. Seizure 2012;21(10):795–801.
[11] 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(1):139–44.
[12] Cornaggia CM, Gugliotta SC, Magaudda 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(4):598–604.
[13] 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.
[14] Yao Y, Ma W, Markus R, Lu Q, Huang Y, Zhou X, et al. Conversation analysis in differential diagnosis between epileptic seizure and psychogenic nonepileptic seizure. Chin J Neurol 2017:266–70.
[15] Biberon J, de Liège A, de Toffol B, Limousin N, El-Hage W, Florence AM, et al. Differentiating PNES from epileptic seizures using conversational analysis on French patients: A prospective blinded study. Epilepsy Behav 2020;111:107239.
[16] Beghi M, Cornaggia I, Diotti S, Erba G, Harder G, Magaudda A, et al. The semantics of epileptic and psychogenic nonepileptic seizures and their differential diagnosis. Epilepsy Behav 2020;111:107250.
[17] Pevy N, Christensen H, Walker T, Reuber M. Feasibility of using an automated analysis of formulation effort in patients' spoken seizure descriptions in the differential diagnosis of epileptic and nonepileptic seizures. Seizure 2021;91:141–5.
[18] Robinson JD. An interactional structure of medical activities during acute visits and its implications for patients' participation. Health Commun 2003;15(1):27–59.
[19] Cassell EJ. Talking with patients, Volume 1: The theory of doctor-patient communication. MIT Press; 1985.
[20] Brown RJ, Reuber M. Psychological and psychiatric aspects of psychogenic non-epileptic seizures (PNES): a systematic review. Clin Psychol Rev 2016;45:157–82.
[21] Rawlings GH, Jamnadas-Khoda J, Broadhurst M, Grünewald RA, Howell SJ, Koepp M, et al. Panic symptoms in transient loss of consciousness: frequency and diagnostic value in psychogenic nonepileptic seizures, epilepsy and syncope. Seizure 2017;48:22–7.
[22] Whitfield A, Walsh S, Levita L, Reuber M. Catastrophising and repetitive negative thinking tendencies in patients with psychogenic non-epileptic seizures or epilepsy. Seizure 2020;83:57–62.
[23] Pennebaker JW, Chung CK. Expressive writing, emotional upheavals, and health; 2007.
[24] Pennebaker JW, Francis ME, Booth RJ. Linguistic inquiry and word count: LIWC 2001, vol. 71, no. 2001. Mahway: Lawrence Erlbaum Associates; 2001. p. 2001.
[25] Pennebaker JW, Francis ME. Cognitive, emotional, and language processes in disclosure. Cogn Emot 1996;10(6):601–26.
[26] Alpers GW, Winzelberg AJ, Classen C, Roberts H, Dev P, Koopman C, et al. Evaluation of computerized text analysis in an Internet breast cancer support group. Comput Hum Behav 2005;21(2):361–76.
[27] Anderson B, Goldin PR, Kurita K, Gross JJ. Self-representation in social anxiety disorder: Linguistic analysis of autobiographical narratives. Behav Res Ther 2008;46(10):1119–25.
[28] Rosenbach C, Renneberg B. Remembering rejection: Specificity and linguistic styles of autobiographical memories in borderline personality disorder and depression. J Behav Ther Exp Psychiatry 2015;46:85–92.
[29] Holtzman NS. A meta-analysis of correlations between depression and first person singular pronoun use. J Res Pers 2017;68:63–8.
[30] Shibata D, Wakamiya S, Kinoshita A, Aramaki E. Detecting Japanese patients with Alzheimer's disease based on word category frequencies. In: Proceedings of the Clinical Natural Language Processing Workshop (ClinicalNLP). p. 78–85.
[31] Cardeña E, Pick S, Litwin R. Differentiating psychogenic nonepileptic from epileptic seizures: A mixed-methods, content analysis study. Epilepsy Behav 2020;109:107121.
[32] Jenkins L, Cosgrove J, Ekberg K, Kheder A, Sokhi D, Reuber M. A brief conversation analytic communication intervention can change history-taking in the seizure clinic; 2015.
[33] Pennebaker JW, Boyd RL, Jordan K, Blackburn K. The development and psychometric properties of LIWC2015; 2015.
[34] Dimaro LV, Dawson DL, Roberts NA, Brown I, Moghaddam NG, Reuber M. Anxiety and avoidance in psychogenic nonepileptic seizures: The role of implicit and explicit anxiety. Epilepsy Behav 2014;33C:77–86.
[35] Rawlings GH, Brown I, Stone B, Reuber M. Written Accounts of Living With Epilepsy or Psychogenic Nonepileptic Seizures: A Thematic Comparison. Qual Health Res 2018;28(6):950–62.
[36] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine learning in Python. J Mach Learn Res 2011;12:2825–30.
[37] Breiman L. Random forests. Mach Learn 2001;45(1):5–32.
[38] Cristianini N, Shawe-Taylor J. An introduction to support vector machines and other kernel-based learning methods. Cambridge University Press; 2000.
[39] Altman NS. An introduction to kernel and nearest-neighbor nonparametric regression. Am Stat 1992;46(3):175–85.
[40] Vabalas A, Gowen E, Poliakoff E, Casson AJ. Machine learning algorithm validation with a limited sample size. PLoS One 2019;14(11):e0224365.
[41] Mirheidari B, Blackburn D, Walker T, Reuber M, Christensen H. Dementia detection using automatic analysis of conversations. Comput Speech Lang 2019;53:65–79.
[42] Sundermeyer M, Schlüter R, Ney H. LSTM neural networks for language modeling. Thirteenth annual conference of the international speech communication association, 2012.