Language Processing for Predicting Suicidal Tendencies: A Case Study in Greek Poetry

Alexandros Dimitrios Zervopoulos, Evangelos Geramanis, Alexandros
Toulakis, Asterios Papamichail, Dimitrios Triantafylloy, Theofanis Tasoulas,
Katia Kermanidis

HAL Id: hal-02363833
https://inria.hal.science/hal-02363833

Submitted on 14 Nov 2019

Language Processing for Predicting Suicidal Tendencies: A Case Study in Greek Poetry

Alexandros Dimitrios Zervopoulos, Evangelos Geramanis, Alexandros Toulakis, Asterios Papamichail, Dimitrios Triantafylloy, Theofanis Tasoulas and Katia Kermanidis

Ionian University

Abstract. Natural language processing has previously been used with fairly high success to predict a writer's likelihood of committing suicide, using a wide variety of text types, including suicide notes, micro-blog posts, lyrics and even poems. In this study, we extend work done in previous research to a language that has not been tackled before in this setting, namely Greek. A set of language-dependent (but easily portable across languages) and language-independent linguistic features is proposed to represent the poems of 13 Greek poets of the 20th century. Prediction experiments resulted in an overall classification rate of 84.5% with the C4.5 algorithm, after having tested multiple machine-learning algorithms. These results differ significantly from previous research, as some features investigated did not play as significant a role as was expected. This kind of task presents multiple difficulties, especially for a language where no previous research has been conducted. Therefore, a significant part of the annotation process was performed manually, which likely explains the somewhat higher classification rates compared to previous efforts.

Keywords: NLP · Suicide · Poetry

1 Introduction

A significant amount of research has been conducted on the topic of identifying mental health illnesses through the means of text analysis. For instance, such analysis has been used to predict various psychological states [1]. Being able to detect such psychological states in their early stages, through the means of Natural Language Processing (NLP) techniques could be of vital importance, before they lead to an unfortunate conclusion for the individual.

Furthermore, a great deal of effort has gone into discerning suicidal tendencies through the writer's language use, as suicide is a leading cause of death all over the world [2]. More specifically, there have been multiple recent attempts focusing on suicide notes. Pestian et al. [3] concluded that NLP could be used to differentiate between genuine and elicited suicide notes, by achieving a higher classification rate than mental health professionals. Many have tried to tackle the problem of emotion detection in suicide notes utilizing various methods,

including entropy classification [4] and latent sequence models [5] in the pursuit of better understanding the suicidal mind.

Additionally, with the rise of the digital era and widespread use of the Internet, the phenomena of micro-blogging and social media emerged. This has resulted in an increasingly large proportion of people who use these means daily to express their feelings and opinions. Thus, there has been major interest in examining whether these can be useful in predicting suicidal ideation. Zhang et al. [6] deduced that such a task is realizable using the Chinese version of the Linguistic Inquiry and Word Count (LIWC), as well as Latent Dirichlet Allocation (LDA), on Chinese micro-blog users' data. Litvinova et al. [7] developed a mathematical model to predict suicidal tendencies based on Russian Internet texts, employing numerical rather than linguistic features. Burnap et al. [8] aimed to identify suicide related topics from posts on Twitter by training baseline classifiers, then improving upon them with an ensemble classifier.

It has been proven that suicide rates among artists, such as musicians and poets, are significantly higher than rates pertaining to the general population [9,10,11]. Lightman et al. [12] deployed Coh-Metrix and LIWC to contrast textual features of suicidal and non-suicidal songwriters. Mulholland and Quinn [13] composed a corpus consisting of songs from various English lyricists, which was comprised of a development, a training and a test set. This corpus was then used to derive lexical, syntactic, semantic class and n-gram features. These were then input into the Waikato Environment for Knowledge Analysis (Weka) to compare the performance of multiple machine learning algorithms in classifying whether or not the song was written by a suicidal lyricist. The algorithm that proved to have the highest classification rate was SimpleCart, with an overall accuracy of 70.6%. Pajak and Trzebiński [14] used LIWC to analyze Polish poems from six separate poets. Afterwards, ANOVA and logistic regression were used to extract the most prevalent features in identifying suicidal predisposition. They drew conclusions similar to those of previous works.

Perhaps one of the most influential contributions in the text analysis of poets is that of Stirman and Pennebaker [15], considering their results have been the basis of many different studies, such as [12,13,14]. The methodology they followed was the collection of 300 poems, and the study, with the aid of LIWC, of linguistic characteristics that could be distinguished between suicidal and non-suicidal poets. Furthermore, they investigated how such characteristics accord with the two most dominant suicide models, namely Durkheim's model [16], where suicide rate is linked to a society's integration level, and the hopelessness model [17], where an individual is overcome with negative emotions, such as hopelessness and helplessness, which ultimately drive them to suicide. These characteristics were then examined as to how they varied in different stages of the poets' careers. It was concluded that suicide can in fact be predicted by their language use, finding stronger support for the social integration suicide model, and that there's no significant variation over time.

The purpose of this study is to train a machine learning model, which classifies a poem by the poet's proclivity towards suicide, based on textual features.

In contrast to previous research in this area, the study will be focusing on Greek poetry of the 20th century, and aims to examine whether preceding results, derived mostly from English language works, can be verified. These include higher use of the first-person singular form, compared to the first-person plural form, more frequent use of death and sexual-related words, as well as fewer overall positive emotion words [15]. Overall, the methodology followed differs compared to previous efforts, given the lack of available tools for Greek, which resulted in the investigation of verb suffixes as features, as well as measuring in a novel manner the emotion of a poem's lyrics set in the range of $[0, 2]$, in hopes of increasing the overall reliability of the corpus' annotation. The algorithms used were for the most part examined in previous research, and have been shown to perform relatively well in these kinds of tasks.

The rest of this paper is organized as follows. Section 2 describes the collected data, as well as the process and criteria for gathering it. Section 3 presents the feature vector that represents each poem, and why the respective values were selected. In Section 4, the classification experiments are described in detail, including the algorithms used and their results. Section 5 illustrates some of the difficulties faced during the process of classifying a poet's suicidal ideation by their poem, especially in a language with a scarcity of processing tools. Finally, Section 6 presents future improvements that could be made and a conclusion of what was accomplished in this study.

2 Data Collection

A corpus of 90 poems was constructed, consisting of poems from 7 poets who committed suicide and 6 who did not. The number of poems is equally distributed between the two groups. Moreover, the number of poems from each poet ranges from 5 to 9. The vast majority of poets are male (with the exception of two female poets in the suicidal group), who lived in approximately the same period of time, i.e. in the early to mid 1900s. The rationale behind this requirement was for the poets to belong to the same phase of the Greek language history, as the writing style of Greek changed significantly during the second half of the 20th century. Specifically, Katharevousa was abolished as the official Greek language in 1976 and was replaced by Modern Greek.

It was of significant importance that the poets in the suicide group undoubtedly took their own life. Similarly, special attention was paid so that poets in the non-suicide group did not have a history of self-harm or suicide attempts. All of the above is illustrated in Table 1. The size of the poems varies between 80-300 words, with an average size of 155.35. For each poet, all poems were randomly selected, with no distinction as to the nature of their content, e.g. a poem particularly high in negative emotion was not specifically selected for a poet who committed suicide. The above criteria, combined with the overall low suicide rates of Greek poets, significantly restricted the size of the corpus.

Poet	Poems	Year of Death	Suicide
Giorgos Makris	7	1968	Yes
Kostas Karyotakis	5	1928	Yes
Alexis Traianos	5	1980	Yes
Minas Dimakis	5	1980	Yes
Katerina Gogou	8	1993	Yes
Napoleon Lapathiotis	7	1944	Yes
Koralia Andreadis	8	1976	Yes
George Seferis	9	1971	No
Tellos Agras	7	1944	No
Constantine P. Cavafy	7	1933	No
Kostis Palamas	8	1943	No
Nikolaos Kavvadias	7	1975	No
Odysseus Elytis	7	1996	No

Table 1. Composition of Corpus

The poems were stripped of all punctuation marks and accentuation using a Python script. This was required by certain parts of the annotation process described next, so that they could be automated.

3 Feature Description

Every poem is represented as a feature-value learning vector. The features that were selected for designing the vector are presented below. Some of these were based on the work of Kao and Jurafsky [18] as well as Stirman and Pennebaker [15], whereas others are novel, which are mostly applicable to the Greek language. A large number of features selected represent sums of occurence of linguistic characteristics of interest throughout the entire poem, which were normalized by the poem's length. These sum values for each poem were then normalized across the entire corpus using the well known normalization transformation of Equation 1, where v is the initial feature value, v_{min} and v_{max} are the minimum and maximum values of the feature across all poems, and v_{norm} is the resulting normalized value.

$$V_{norm} = \frac{v - v_{min}}{v_{max} - v_{min}} \tag{1}$$

Overall, the number of features that were explored reached 37 in total.

3.1 Vocabulary Features

The vocabulary feature that was chosen is the type to token ratio (TTR), which is used as an indicator of the poem's richness in vocabulary [18].

3.2 Morphosyntactic Features

The social integration model suggests that individuals who commit suicide have failed to integrate with society, so they are expected to be more self-centered,

which seems to be manifested through the use of more first-person singular words, as opposed to first-person plural words [15]. Based on that observation, the two morphosyntactic features that were selected were the count of occurrences of first-person singular and first-person plural verbs.

Base Suffix	Different Tenses				
	Singular Active Construction				
αβω	αβω	αβα	αψα	αψω	
αζω	αζω	αζα	ασα	ασω	
αινω	αινω	αινα	ηνω		
εινω	εινω	εινα	εισα	εισω	
ερνω	ερνω	ερνα	ειρα	ερα	ειρω
ευω	ευω	ευα	εψα	εψω	
ηνω	ηνω	ηνα	ησα	ησω	
ιζω	ιζω	ιζα	ισα	ισω	
ινω	ινω	ινα	ωσα	ωσω	
λλω	λλω	λλα	ιλα	ιλω	
ττω	ττω	ττα	ξα	ξω	
ωνω	ωνω	ωνα	ωσα	ωσω	
ω					
	Plural Active Construction				
αβουμε	αβουμε	αβαμε	αψαμε	αψουμε	
αζουμε	αζουμε	αζαμε	ασαμε	ασουμε	
αινουμε	αινουμε	αιναμε	ησαμε	ησουμε	
εινουμε	εινουμε	ειναμε	εισαμε	εισουμε	
ευουμε	ευουμε	ευαμε	εψαμε	εψουμε	
ηνουμε	ηνουμε	ηναμε	ησαμε	ησουμε	
ιζουμε	ιζουμε	ιζαμε	ισαμε	ισουμε	
ινουμε	ινουμε	ιναμε	σαμε	σαμεν	σουμε
λλουμε	λλουμε	λλαμε	λαμε	λουμε	
ττουμε	ττουμε	τταμε	ξαμε	ξουμε	
	Singular Passive Construction				
αμαι	αμαι				
ουμαι	ουμαι	ουμουν			
ιεμαι	ιεμαι	ιομουν			
ομαι	ομαι	ομουν			
μαι	μαι	μουν	μουνα	ηκα	ηθω
	Plural Passive Construction				
ομαστε	ομαστε	ομασταν			
μαστε	μαστε	μασταν	ουμε	ηκαμε	

Table 2. Suffixes used for the morphosyntactic features

Due to the lack of reliable morphological analysis tools for Greek, especially for the historical phase of the language targeted herein, these counts were obtained using a set of predefined verb suffixes (Table 2), that constitute person and number morphemes. The core set of these suffixes was found in [19], and was subsequently enriched by the authors of this paper. Afterwards, a Python

script was used to identify them in the poems. When any tense of a base suffix is found, it counts towards the base's occurrence, e.g. the occurrence of suffix ειρω is equivalent to that of ερνω. These bases are also used as features, to examine whether a subset of them are more prevalent in the process of the classification. All in all, these constitute 32 of the 37 features that were examined, 30 of which are the aforementioned suffixes.

3.3 Semantic Class Features

Suicidal poets are expected to deal more with negative emotion than with positive feelings, according to the hopelessness suicide model [17]. To test this, each verse was assigned a number ranging from 0 to 2 for the positive emotion it expressed, and 0 to 2 for the negative emotion. The range is overall lower than ones used in previous studies, as $[-5, 5]$ has been used for example in [20]. This was done to increase the overall reliability of the manual annotation. These numbers were then summed up for all verses of a poem, resulting in two features: one sum reflecting the positive, and one reflecting the negative emotion.

Additionally, sexual and death-related references were included as features, as has been suggested, in an ad-hoc analysis, that they can contribute to identifying suicidal tendencies [15]. For these features, the number of verses for each poem that were deemed to contain sexual or death references were summed up for each poem.

Each poem was annotated with the aforementioned semantic information by two Greek language native speakers. When a verse led to contrasting annotators' decisions, the feature value was decided upon by the majority vote of the authors of the present work.

4 Prediction Results

Each poem was represented using the aforementioned features. The Weka machine learning workbench [21] was used for running prediction experiments pertaining to poets' suicidal tendencies. The baseline for our comparisons is the prediction accuracy of 50%, as the two groups are evenly split. Since the corpus was not divided into different sets for training and testing, Weka's k-fold cross validation functionality was used, for k=5. A variety of tree and rule-based algorithms were compared as to their performance at accomplishing this task. The parameters used, which achieved the best results when running these tests, are described below. These are well-established algorithms in dealing with such tasks, JRIPPER was used in [3], C4.5 was used in [3] and [8] and finally Simple CART was used in [13].

The rule-based classifier tested was JRIPPER. This algorithm was run using 50 optimization runs, achieving a classification rate of 84.4% when making use of pruning. Afterwards, three different tree classifiers were tested, the first one being Simple CART. This resulted in an accuracy of 82.8%. Another tree based classifier used was RandomForest, which had a classification rate of 82.2%, after

having performed 10,000 iterations. The last algorithm tested was C4.5 (known as J48 in Weka) classifying 84.5% of the instances correctly, with 0.25 as the confidence factor, performing pruning.

The algorithm that eventually achieved the highest classification rate was C4.5, reaching 84.5%. Various statistics are presented in Table 3 detailing some of the tests done, and the models derived from C4.5 are presented in Fig 1 and Fig 2.

| Algorithm	Precision	Recall	F-Measure	Class
JIPPER (without pruning)	0.739	0.756	0.747	TRUE
	0.750	0.733	0.742	FALSE
JRIPPER (with pruning)	0.830	0.867	0.848	TRUE
	0.860	0.822	0.841	FALSE
Random Forest	0.800	0.800	0.800	TRUE
	0.800	0.800	0.800	FALSE
Simple CART (with pruning)	0.784	0.889	0.833	TRUE
	0.872	0.756	0.810	FALSE
Simple CART (without pruning)	0.848	0.867	0.857	TRUE
	0.864	0.844	0.854	FALSE
C4.5 (with pruning)	0.830	0.867	0.848	TRUE
	0.860	0.822	0.841	FALSE
C4.5 (without pruning)	0.787	0.822	0.804	TRUE
	0.801	0.800	0.800	FALSE

Table 3. Detailed accuracy for both classes by each algorithm
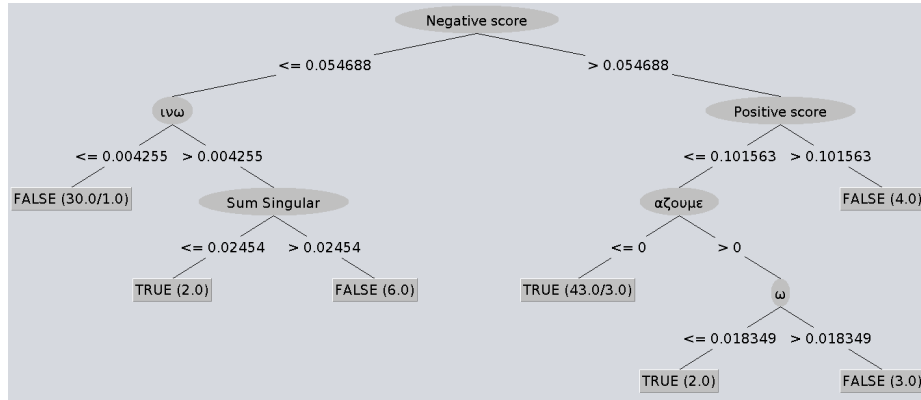
Fig. 1. The decision tree generated by C4.5 with pruning

The resulting models indicate that the overall classification result is largely based on the positive and negative emotion scores. At a first glance, and considering the researchers' lack of a professional literary background, the suffixes don't
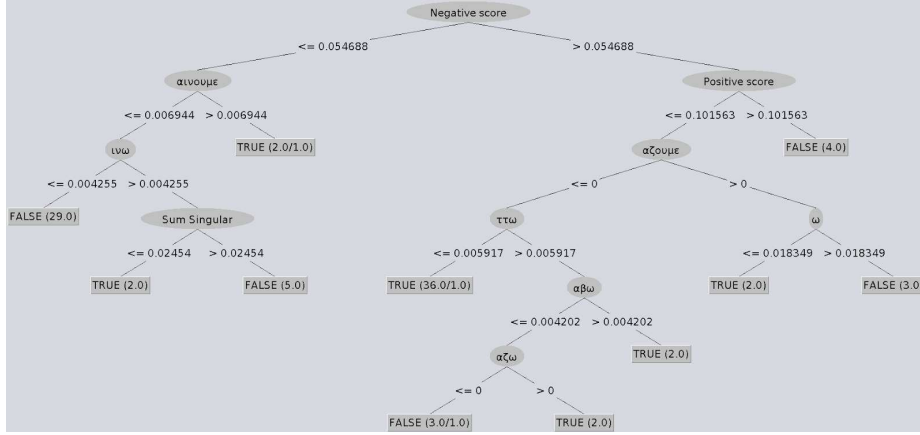
Fig. 2. The decision tree generated by C4.5 without pruning

seem to form some clear pattern which could help identify suicidal ideation. However, it is evident that the results do differ based on the occurence of singular, as opposed to plural, verb suffixes, which does support Durkheim's model.

Tree and rule-based classifiers perform fairly well and result in simple models, which is in part attributed to the manual nature of the annotation process, as a result of the lack of reliable processing tools for the Greek language, as well as to the nature of the semantic features, which are of high-level linguistic knowledge. They have proven to perform fairly well in similar research conducted in the past as well. As a basis for comparison, the results of previous studies are presented in this paragraph. Pestian et al. [3] managed to accurately distinguish between elicited and genuine suicide notes 78% of the time. Litvinova et al [7] reached a 71.5% classification rate of deciding whether Russian internet texts were suicidal or not. Mulholland and Quinn [13] achieved a performance of 70.6% in identifying suicidal tendencies of songwriters through their lyrics.

5   Discussion

Tackling the task of identifying suicidal ideation in poetry, particularly for a language where no previous research has been conducted, is certain to pose many difficulties. The construction of the corpus was largely difficult due to the strict criteria described in Section 2. It was hard to determine whether someone did not at least attempt to commit suicide in their lifetime, as it may not have necessarily become known to people outside their close social circle. Furthermore, in the case where it was ambiguous whether the cause of death was actually suicide, as is the case with Maria Polydouri, the poet was left out, which further shortened the number of candidate poems. Perhaps, one solution to this would have been to focus more on the few poets that have written a lot of poems. This does, however, introduce the risk of the classifier learning the

patterns of those particular poets and being unable to accurately classify poems by others.

Additionally, and perhaps most importantly, the process of annotating the corpus was especially demanding, since the most prevalent tools used in previous research were not available for Greek. Such tools include LIWC, the UAM CorpusTool, various word-lists, such as AFINN and ANEW, used in sentiment analysis tasks, as well as corpora used in previous studies, which could have been used for comparison. This resulted in the manual annotation of a significant portion of the selected features. It was also a major restriction in the selection of the features, as the more sophisticated a feature is, the more rigorous the process of identifying it in the text needs to be. Not adhering to the appropriate level of rigorousness introduces further potential bias into the data, and given the annotators' lack of a professional literary background, this would have been an increasingly precarious task.

6   Conclusion and Future Work

In conclusion, a corpus of poems was composed to identify suicidal ideation in Greek poetry of the 20th century. This proved to be a challenging task, since there has not been any previous work done on the language that was selected. Nonetheless, the resulting classifier, C4.5, reached an accuracy of 84.4%. The features explored were mostly semantic in nature, while also utilizing morphosyntactic features, such as verb suffixes which are language specific and have not been investigated before. The results are overall promising and the features selected are easily portable across different strategies, which should allow future studies to confirm how successful our methodology was.

The construction of a properly annotated corpus proved to be the most challenging part of this task. Therefore, the development of such a corpus would be worth looking into, as it would greatly aid future efforts in studying NLP related topics for Greek texts. It is of critical importance, as has been showcased by the attempts in this study, to pay special attention when manually annotating data, and to adhere to best practices which have been studied extensively before, for example by Janyce et al. [22].

There's also a scarcity of available tools, which are often required in NLP related tasks, due to the complex structure of the Greek language. For example, tools for lemmatization and stemming are somewhat difficult to implement, partly due to punctuation and grammar rules. This indicates it would be meaningful to spend time developing such tools, before further progress into more intensive NLP tasks is made. Considering the above, any future advancements in the area will be interesting to behold, after this first attempt has been made, to keep track of how such difficulties are handled.

References

1. J. W. Pennebaker and L. A. King, "Linguistic styles: Language use as an individual difference." *Journal of personality and social psychology*, vol. 77, no. 6, p. 1296,

1999.

2. J. M. Bertolote and A. Fleischmann, "A global perspective on the magnitude of suicide mortality," *Oxford textbook of suicidology and suicide prevention: A global perspective*, pp. 91–98, 2009.

3. J. Pestian, H. Nasrallah, P. Matykiewicz, A. Bennett, and A. Leenaars, "Suicide note classification using natural language processing: A content analysis," *Biomedical Informatics Insights*, vol. 3, p. BII.S4706, 2010. [Online]. Available: https://doi.org/10.4137/BII.S4706

4. R. Wicentowski and M. R. Sydes, "emotion detection in suicide notes using maximum entropy classification," *Biomedical informatics insights*, vol. 5, pp. BII–S8972, 2012.

5. C. Cherry, S. M. Mohammad, and B. D. Bruijn, "Binary classifiers and latent sequence models for emotion detection in suicide notes," *Biomedical Informatics Insights*, vol. 5s1, p. BII.S8933, 2012. [Online]. Available: https://doi.org/10.4137/BII.S8933

6. L. Zhang, X. Huang, T. Liu, A. Li, Z. Chen, and T. Zhu, "Using linguistic features to estimate suicide probability of chinese microblog users," in *Human Centered Computing*, Q. Zu, B. Hu, N. Gu, and S. Seng, Eds. Cham: Springer International Publishing, 2015, pp. 549–559.

7. T. A. Litvinova, P. V. Seredin, O. A. Litvinova, and O. V. Romanchenko, "Identification of suicidal tendencies of individuals based on the quantitative analysis of their internet texts," *Computación y Sistemas*, vol. 21, no. 2, 2017.

8. P. Burnap, W. Colombo, and J. Scourfield, "Machine classification and analysis of suicide-related communication on twitter," in *Proceedings of the 26th ACM conference on hypertext & social media*. ACM, 2015, pp. 75–84.

9. K. R. Jamison, "Manic-depressive illness and creativity," *SCIENTIFIC AMERICAN-AMERICAN EDITION-*, vol. 276, pp. 44–52, 1997.

10. S. D. Raeburn, "Psychological issues and treatment strategies in popular musicians: A review, part i," *Medical Problems of Performing Artists*, vol. 14, pp. 171–179, 1999.

11. S. Stack, "Suicide among artists," *The Journal of Social Psychology*, vol. 137, no. 1, pp. 129–130, 1997, pMID: 9121137. [Online]. Available: https://doi.org/10.1080/00224549709595421

12. E. J. Lightman, P. M. McCarthy, D. F. Dufty, and D. S. McNamara, "Using computational text analysis tools to compare the lyrics of suicidal and non-suicidal songwriters," in *Proceedings of the Annual Meeting of the Cognitive Science Society*, vol. 29, no. 29, 2007.

13. M. Mulholland and J. Quinn, "Suicidal tendencies: The automatic classification of suicidal and non-suicidal lyricists using nlp," in *Proceedings of the sixth international joint conference on natural language processing*, 2013, pp. 680–684.

14. K. Pająk and J. Trzebiński, "Escaping the world: Linguistic indicators of suicide attempts in poets," *Journal of Loss and Trauma*, vol. 19, no. 5, pp. 389–402, 2014. [Online]. Available: https://doi.org/10.1080/15325024.2013.794663

15. S. W. Stirman and J. W. Pennebaker, "Word use in the poetry of suicidal and nonsuicidal poets," *Psychosomatic medicine*, vol. 63, no. 4, pp. 517–522, 2001.

16. E. Durkheim, *Suicide: A study in sociology*. Routledge, 2005.

17. K. Petrie and R. Brook, "Sense of coherence, self-esteem, depression and hopelessness as correlates of reattempting suicide," *British Journal of Clinical Psychology*, vol. 31, no. 3, pp. 293–300, 1992.

18. J. Kao and D. Jurafsky, "A computational analysis of style, affect, and imagery in contemporary poetry," in *Proceedings of the NAACL-HLT 2012 Workshop on Computational Linguistics for Literature*, 2012, pp. 8–17.

19. Wictionary, https://el.wiktionary.org/wiki/, Last accessed 16 Mar 2019.

20. F. Å. Nielsen, "A new ANEW: evaluation of a word list for sentiment analysis in microblogs," *CoRR*, vol. abs/1103.2903, 2011. [Online]. Available: http://arxiv.org/abs/1103.2903

21. Weka 3 - Data Mining with Open Source Machine Learning Software in Java, https://www.cs.waikato.ac.nz/ml/weka/.

22. J. Wiebe, T. Wilson, and C. Cardie, "Annotating expressions of opinions and emotions in language," *Language resources and evaluation*, vol. 39, no. 2-3, pp. 165–210, 2005.