
IDENTIFYING SUICIDE NOTES USING FORENSIC LINGUISTICS

Yonatan Eklesia Laia

Magister Ilmu Linguistik/Universitas Warmadewa

jonathanxclusivelaiya@gmail.com

Published: Desember/2022

How to cite :

Laia, Yonatan Eklesia. (2022). Identifying Suicide Notes Using Forensic Linguistics. *IJFL (International Journal of Forensic Linguistic)*, Vol 3(2), Page. Doi: <https://doi.org/10.55637/ijfl.3.2.6335.98-104>

Abstract- Forensic linguistics is a subfield of corpus linguistics that uses linguistic information, analysis techniques, and linguistic insights to the setting of subjects such as law, criminal investigation, trial, and judicial procedure. In addition to different aspects of criminal investigation and court proceedings, its scope encompasses the domains of authorship identification (such as authorship verification, authorship profiling, and authorship attribution). The aim research is to identifying suicide notes using forensic linguistics. The research method used cumulative sum calculation strategy for elements: SL, Pron, Prep, Ver, Adj, Adv, Conj, CSS-SL, CSS-Pron, CSS-Prep, CSS-Ver, CSS-Adj, CSS-Adv, & CSS-Conj. The program read the given texts and counted the percentage of words that reflect different emotions, thinking styles, and social concerns. The result shows that from the two letters that have been analyzed above, it can be concluded that LETTER 1 has several aspects that support that the letter is original and was written by a single author. On the other hand, LETTER 2 does not have aspects that support the indicator that the letter is genuine, but all aspects analyzed strengthen the indicator that the letter is fabricated, as evidenced by the results of the analysis shown in the graph (file attached).

Keywords: *Suicide Notes, Forensic Linguistic, Linguistics*

I. INTRODUCTION

Traditionally, "forensic" refers to the use of scientific methods to crime investigation, which encompasses the legal requirements for acceptable evidence and criminal procedures. During an investigation, forensic scientists are responsible for collecting, preserving, and analyzing the evidence. Some forensic scientists travel to the scene of the crime to collect evidence, while others analyze artifacts in the laboratory. The evidence consists of numerous tangible and intangible elements, including fingerprints, hair, DNA detection, and transfusion analysis. Digital forensics, on the other hand, acquires and analyzes commercially and/or intimately utilized digital data when the data become the focus of an inquiry. The digital data are often stored on digital media and sent across a network. Digital forensics scientists collect, store, and analyze such data before submitting them to the legal system as evidence.

Professor Jan Svartvik began forensic linguistics in 1968 (Svartvik, 1968) by analyzing the writings of Timothy John Evans (a renowned murder suspect). He studied four texts with various linguistic characteristics and discovered significant inconsistencies between them. It hinted that the works were maybe not written by the same author. The International Association of Forensic Linguists (IAFL) was established in 1993, and publication of the international magazine *The Law and the International Journal of Law, Language and Discourse* began in 1994.

Forensic linguistics is a subfield of corpus linguistics that uses linguistic information, analysis techniques, and linguistic insights to the setting of subjects such as law, criminal investigation, trial, and judicial procedure. In addition to different aspects of criminal investigation and court proceedings, its scope encompasses the domains of authorship identification (such as authorship verification, authorship profiling, and authorship attribution). Forensic linguistics is the branch of the linguistics study which investigates legal text in crime. From its point of view, forensic linguistics is really

helping the court or police or law attribute in investigating criminal case (Sudjana & Fitri, 2013). Coulthard & Johnson (2010) mention that forensic linguistics ranges from courtroom discourse and legal language to plagiarism. Briefly, plagiarism is using another person's work for personal advantage without mentioning his/her name. Forensic linguistic experts proficient in plagiarism cases and copyright infringements provide evidence to determine which work is based upon another.

As forensic linguistics advances, more research on suicide notes have been undertaken. Durkheim (1951) categorized four types of suicide based on integration and regulation. The first is a "egoistic suicide," which occurs when suicide perpetrators are not incorporated into the social group and do not receive assistance from the social group. In this sort of suicide, individuals frequently feel alone or despondent when facing adversity. The second type is a "selfless suicide," which occurs when social cohesion is extremely poor. It is a sort of suicide connected with persons who have a very strong relationship with the group and whose personal identity is valued. In this sort of suicide, individuals typically disregard their own needs in favor of group objectives. When the degree of social control is too low, when people encounter turmoil or an unanticipated economic collapse, when they lose moral norms or societal common values that regulate their activities, or when they lose a lover due to death or divorce, a "anomic suicide" happens. People who commit this sort of suicide do so because they are overwhelmed by a significant shift in their lives. The fourth is a 'fatalistic suicide', which occurs owing to an overwhelming oppression or despair. People believe they have no future and that an authority restricts their freedom.

From a linguistic forensics perspective, the suicide notes exhibit typical (linguistic) characteristics. In forensic linguistics, the examination of suicide notes requires the use of all theoretical linguistics (including phonetics, phonology, morphology, syntax, semantics, pragmatics, and discourse

analysis). Shneidman's corpus has served as the foundation for the majority of the earliest linguistic studies on suicide texts. The corpus contains 66 pieces, a mix of authentic and fabricated suicide notes. The texts are evaluated linguistically using discourse analysis techniques, the usage of various auxiliary verbs (including modals), or verbs that distinguish between the genuine and the counterfeit. Olsson (2004) said, "Suicide notes generally have sentences alluding to the act of killing oneself, or the method of suicide that was undertaken" (p. 1).

Some research have utilized discourse analysis (Edelman & Renshaw, 1982) or semantic space analysis to examine suicide notes (Matykiewicz et al., 2009). Using the so-called Syntactic Language Computer Analysis, Edelman & Renshaw (1982) examined not only syntactic elements such as parts of speech (POS), but also the semantic aspects of nouns, verbs, adjectives, etc. (Pestian et al., 2008; Pestian et al., 2010). As corpus linguistics and machine learning advance, there are currently numerous attempts to apply these techniques to the statistical analysis of suicide notes. Analysis of many linguistic units, including personal pronouns, past tense verbs, nouns, and numerous semantic categories (Olsson, 2008). In recent years, there has been a discernible trend to rely on the automated corpus analysis techniques for the detection of suicide notes. Scholars use corpus-analysis methods (Shapiro, 2011) or automated machine-learning techniques in order to recognize and categorize typical suicide notes (Pestian et al., 2010). The aim research is to identifying suicide notes using forensic linguistics.

II. METHOD

The research method used cumulative sum calculation strategy for elements: SL, Pron, Prep, Ver, Adj, Adv, Conj, CSS-SL, CSS-Pron, CSS-Prep, CSS-Ver, CSS-Adj, CSS-Adv, & CSS-Conj. The program read the given texts and counted the percentage of words that reflect different emotions, thinking styles, and social concerns. Next, the suicide notes were read several times; the clauses, phrases, and sentences were

highlighted and analyzed with semantic forensic theory.

III. RESULT AND DISCUSSION

Letter 1

*Assalamualaikum warahmatullahi
wabarakatuh*

*Wasiat kepada orang yang saya cintai
karena Allah*

*Wahai Uummy ku, minta maafka kalo ada
salahku baik perilaku maupun lisanku,
jangan ki lupa senantiasa beribadah kepada
Allah dan jangan ki tinggalkan sholat.
Semoga Allah kumpulkanki di Surganya.*

*Uummy sekali lagi minya maaf ka, ku sayang
sekali tapi Allah lebih menyayangi
hambanya.*

*Makanya saya tempuh jalanku sebagai
mana jalan Nabi/Rasul Allah untuk
selamatkan ki dan bisa ki kembali
berkumpul di surga*

*Satu ji pesanku buat kita ummy, berhenti
ambil uang bank, karena uang bank itu riba
dan tidak diberkahi oleh Allah.*

*Ini ada uang simpananku 2.350.000 untuk
bayar pinjaman di bank dan itu uang
kontrak rumahku masih ada 5 bulan di
karyawan laundrynya mus. 500.000/bulan
na kontrakan ambil meri tiap bulan, simpan
ki untuk bayar pinjaman*

*Pitto, minta maaf ka kalau ada salahku dek,
baik itu lisanku maupun perbuatanku dulu.*

*Satu pesanku untuk kau dek, jaga ummy
baik-baek. Kau mami bisa jaga ummy dan
jangan juga malas-malasan sholat dan
jangan i bergaul-gaul, fokus saja bantu
ummy.*

*Istiqomah ki semua di jalan ini nah ummy,
Pitto dan keluarga ku yang saya cintai
karna Allah, semoga Allah kumpulkan ki di
surga dan semua sodarahnya dan keluarga
bapakku*

Muh. Lukman Alfariz

Letter 2

*Assalamualaikum warahmatullahi
wabarakatuh*

*Wasiat kepada orang yang saya cintai
karena Allah*

*Wahai mamaku, maafin Zakiah yang belum
pernah membalas pemberian keluarga.
Mama, ayah jangan lupa senantiasa
beribadah kepada Allah SWT dan jangan*

tinggalkan salat. Semoga Allah kumpulkan kembali keluarga di surga.

Mama, sekali lagi Zakiah minta maaf. Zakiah sayang banget sama Mama. Tapi Allah lebih menyayangi hamba-Nya, Makanya Zakiah tempuh jalan ini sebagaimana jalan Nabi/Rasul Allah untuk selamatkan Zakiah dan dengan izin Allah bisa memberi syafaat untuk Mama dan keluarga di akhirat.

Pesan Zakiah untuk Mama dan keluarga, berhenti berhubungan dengan bank (kartu kredit) karena itu riba dan tidak diberkahi Allah.

Pesan berikutnya agar Mama berhenti bekerja menjadi Dawis yang membantu kepentingan pemerintah thogut.

Pesan berikutnya untuk kaka agar rumah Cibubur jaga Dede dan mama, iadah kepada Allah, dan tinggalkan penghasilan dari yang tidak sesuai jaran islam, serta tinggalkan kepercayaan kepada orang-orang yang mengaku mempunyai ilmu, dekati ustad/ulama, tonton kajian dakwah, tidak membanggakan kafir Ahok dan memakai hijab kak. Allah yang akan menjamin rezeki kak. Maaf ya kak, Zakiah tidak bisa membalas semua pemberian kakak...

Untuk Mba Leli agar memperingatkan Mama, jaga Mama ya Mba. Untuk Bp, jangan tinggalkan ibadah solat 5 waktu, maafin ya Mba, pe kalau ada salah lisan dan lainnya. Jaga mama, ayah, dede baik-baik.

Mama, Ayah, semua lihat di samping itu adalah tingkatkan amalan. Insya Allah dengan karunia Allah amalan jihad Zakiah akan membantu memberi syafaat kepada keluarga di akhirat. Jihad adalah tertinggi dalam islam.

Inti pesan Zakiah kepada mama dan keluarga adalah agar tidak mengikuti kegiatan pemilu. Karena orang-orang yang terpilih itu akan membuat hukum tandingan Allah bersumber Alquran-Assunah.

Demokrasi, Pancasila, UUD, pemilu, berasal dari ajaran kafir yang jelas musyrik. Zakiah nasehatkan kepada mama dan keluarga agar semuanya selamat dari fitnah dunia yaitu demokrasi, pemilu dan tidak murtad tanpa sadar.

Sekali lagi maafkan Zakiah ma, ayah, kakak, Mba Leli, awi, Bpe, ka Effa, dede, Baim, Kevin, semuanya. Maafkan bila ada salah kata dan perbuatan. Semoga Allah kumpulkan kembali di surga-Nya. Amiin....
Zakiah Aini

From the results of the analysis of Letters 1 and 2, there are several points that can be indicators to determine whether the author of the letter is genuine or fabricated.

Letter 1

Consists of 13 lines of sentences. From the results of the original text analysis graph and CSS analysis, there are several things that can be considered for data analysis, including:

The letter indicator is original (written by one author), can be seen from the following graphs:

- The graph (SL-CSS SL) shows the consistency of the direction of the lines that are in the same direction and do not intersect.
- The graph (Pron - CSS Pron) shows the direction of the line in the same direction and there are no points that intersect.
- The graph (Adv - CSS Adv) shows the direction of the lines in the same direction and there are no intersecting lines.
- The graph (Conj - CSS Conj) shows the direction of the lines that are in the same direction and do not touch each other.

Letter indicator is engineering (written by more than one author)

- The graph (Prep - CSS Prep) shows the direction of the lines that intersect at the following two points:

If the Coordinate Point = (Prep, CSS Prep), then the following is the point of tangency between the two lines in the graph. The starting point {4, 2} "sentence 8" then goes down and intersects the line at point {0,1} "sentence 9". So on, from the point of the 10th and 11th sentences meet each other in one point.

- b. Graph Verb – CSS Verb shows lines that intersect at point $\{2, (2,46)\}$ "12th sentence".
- c. Graph Adj - CSS Adj shows the direction of the crossed lines and there are several coordinate points where the two lines intersect, namely the point $\{0, (0,15)\}$ "11th sentence".

Therefore, in accordance with the theory of CSS analysis, the 4 graphs above fulfill the indicators which show that the aspects being analyzed are the original versions of the letter and were written by 1 Author.

From the explanation above, it can be concluded that the results of the analysis of letter 1 are dominated by indicators that the letter is original and was written by a single author, as shown in the list below:

- Graphic Sl - CSS Sl shows indicators the letter is original and written by a single author.
- Pron Graph - CSS Pron shows an indicator that the letter is original and written by a single author.
- Prep Graph - CSS Prep shows indicators that the letter is engineered and written by more than one author.
- Graphic Verb - CSS Verb shows the indicator that the letter is engineered and written by more than one author.
- Graph Adj - CSS Adj shows indicators that the letter is engineered and written by more than one author.
- Graph Adv - CSS Adv shows the indicator that the letter is original and written by a single author
- Conj Graph – CSS Conj shows the indicator that the letter is original and written by a single author

Letter 2

Consists of 27 lines of sentences. Based on CSS analysis theory, from the results of the graphical analysis of the original text and CSS analysis of the contents of the letter 2, there are no indicators indicating that the letter is original and written by a single author, but all of them show graphics that intersect so that they do not meet the

indicators of the authenticity of the letter. Evidenced by the explanation below:

- a. The SL-CSS SL graph shows the directions of lines that intersect and intersect each other that occur at the following coordinate points:
 - Period $\{6, (19,22)\}$ "12th sentence"
 - Point $\{5, (8,19)\}$ "sentence 19" intersects the line at point $\{13, (9,04)\}$ "sentence 20".
 - Point $\{20, (16,59)\}$ "23rd sentence" rises and intersects the line at point $\{17, (21,44)\}$ "24th sentence".
- b. Pron Graph – CSS Pron shows lines that intersect at the following coordinate points:
 - Point $\{1, (0,56)\}$ "2nd sentence" rises and intersects the line at point $\{1, (1,33)\}$ "3rd sentence".
 - Point $\{0,0\}$ "sentence 18th" looks like a line that intersects and shows a downward graph.
 - The point $\{0, (0,11)\}$ "22nd sentence" goes down and intersects the line at coordinate point $\{0, (-0,11)\}$ "23th sentence".
- c. Graph Prep – CSS Prep shows the direction of the lines that are not in the same direction and intersect at the following coordinate points:
 - The point $\{4, (2,63)\}$ "11th sentence" goes down and intersects the line at coordinate point $\{0, (1,78)\}$ "12th sentence".
 - The point $\{0, (0,37)\}$ "16th sentence" rises and intersects the line at the coordinates of point $\{1, (0,52)\}$ "17th sentence".
 - The point $\{2, (1,67)\}$ "18th sentence" intersects the coordinates $\{1, (1,81)\}$ "19th sentence".
 - There is an intersection of lines at coordinates $\{1, (0,85)\}$ "26th sentence".
- d. Graph Verb – CSS Verb shows two coordinate points which become the point of contact between two lines, namely:

- Starting point {9, (7,41)} "11th sentence" then goes down and intersects the line at coordinates {1, (6,44)} "12th sentence".
- e. Graph Adj – CSS Adj is a graph that is almost consistent in the direction of the lines in the graph, but there is one coordinate point that intersects each other, namely the point {0, (0,22)} "26th sentence".
- f. Graph Adj - CSS Adj shows the direction of lines that intersect and touch each other, as shown at the following coordinate points:
 - The coordinates of {3, (2,81)} "5th sentence" which is the point of intersection of the lines.
 - The line from the starting point {0, (0,41)} "16th sentence" goes down and intersects the point {1, (0,37)} "17th sentence".
 - The line from the starting point {0, (0,30)} "19th sentence" rises and intersects the point {1, (0,26)} "20th sentence".
- g. The Conj – CSS Conj graph shows lines that are tangent to each other and intersect at the following coordinate points:
 - The coordinates of {2, (1,93)} "10th sentence" are the intersection points between the lines and give an upward trend graph. then descends at the coordinate point {2, (1,85)} "20th sentence" which at this point the two lines cross each other with the rising trend of the graph.

IV. CONCLUSION

From the explanation above, it can be concluded that the results of the analysis of two Letter do not meet the indicators where the letter is original and written by a single author, but indicates engineering and was written by more than one author. as shown in the list below:

- Graphic SL - CSS SL shows indicators that the letter is engineering and was written by more than one author.
- Graphic Pron - CSS Pron shows indicators that the letter is

engineering and was written by more than one author.

- Prep Graph - CSS Prep shows an indicator of a letter being engineered and written by more than one author.
- Graphic Verb - CSS Verb shows the indicator of the letter is engineering and written by more than one author.
- Graph Adj - CSS Adj shows the indicators of the letter being engineered and written by more than one author.
- Graph Adv - CSS Adv shows indicators that the letter is engineering and written by more than one author.
- Conj Graph - CSS Conj shows indicators that the letter is engineering and was written by more than one author.

Thus, from the two letters that have been analyzed above, it can be concluded that LETTER 1 has several aspects that support that the letter is original and was written by a single author. On the other hand, LETTER 2 does not have aspects that support the indicator that the letter is genuine, but all aspects analyzed strengthen the indicator that the letter is fabricated, as evidenced by the results of the analysis shown in the graph (file attached).

Based on the results of forensic analysis using CSS analysis on the two attached letters, it can be concluded that:

Advantages:

- The advantage of CSS is that it can show the results of the analysis of the authenticity of letters in LETTER 1 marked by the use of the subject I, I (first pronoun) which indicates that the letter is objective. so that from the graph that is displayed it is more dominant that the aspect that fulfills the letter indicator is original.
- CSS analysis on letter 2 as a whole shows that the letter is engineering by proving the lines in the graph that intersect and intersect. in this case, CSS analysis can provide evidence that the letter is fabricated.
- From a technical point of view, CSS analysis does not make indicators of

conclusions outside of language such as context in communication and social, but instead focuses on aspects of the language itself such as part of speech, lexical, and even grammar.

Deficiency:

- CSS analysis still cannot fully prove that aspects in one letter are genuine, there are still several other aspects that do not show these indicators.
- CSS analysis can be very complicated and quite difficult when faced with a lot of data, this is because CSS techniques are still manual, especially in analyzing sentences into word class classifications.

REFERENCES

- Coulthard, M., & Johnson, A. (2010). *Handbook of forensic linguistics*. Routledge.
- Durkheim, E. (1951). *Suicide*. The Free Press.
- Edelman, A., & Renshaw, L. (1982). Genuine versus simulated suicide notes: An issue revisited through discourse analysis. *Suicide and Life-Threatening Behavior*, 12(2), 103–113.
- John Olsson. (2008). *Forensic Linguistics: An Introduction To Language, Crime and the Law* (2nd ed.). Bloomsbury Publishing.
- Matykiewicz, P., Wlodzislaw, D., & Pestian, J. (2009). Clustering semantic spaces of suicide notes and newsgroups articles. *Proceedings of the Workshop on Current Trends in Biomedical Natural Language Processing*, 179–184.
- Olsson, J. (2004). *Forensic linguistics: An introduction to language, crime and the law*. Continuum.
- Pestian, J., Matykiewicz, P., Grupp-Phelan, J., A., L., Combs, J., & Kowatch, R. (2008). Using natural language processing to classify suicide notes. *Proceedings of the Workshop on Current Trends in Biomedical Natural Language Processing (BioNLP'08)*, 96–9
- Pestian, J., Nasrallah, H., Matykiewicz, P., Bennett, A., & Leenaars, A. (2010). Suicide note classification using natural language processing: A content analysis. *Biomedical Informatics Insights*, 3, 19–28.
- Shapero, J. (2011). *The language of suicide notes*. University of Birmingham.
- Sudjana, E. T. S., & Fitri, N. (2013). Kurt Cobain's Suicide Note Case: Forensic Linguistic Profiling Analysis. *International Journal of Criminology and Sociological Theory*, 6(4), 217–227.
- Svartvik, J. (1968). *The Evans statements: A case for forensic linguistics*. Gothenburg. University of Gothenburg Press.