

Philosophy, Intercourse, and Suicide: Language in a Suicide Diary with Literary Overtones

Cynthia Whissell

Department of Social Sciences, Laurentian University, Greater Sudbury, Ontario, Canada
Email: cwhissell@laurentian.ca

How to cite this paper: Whissell, C. (2026). *Philosophy, Intercourse, and Suicide: Language in a Suicide Diary with Literary Overtones*. *Psychology*, 17, 855-863. <https://doi.org/10.4236/psych.2026.178042>

Received: June 10, 2026

Accepted: August 8, 2026

Published: August 11, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). <http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

This study analyzes patterns of emotional tone and word usage in *Diary of a Suicide*, a 21,000-word document written by W. E. Baker in 1912-1913 that contains extensive literary, political, and philosophical references. The frequency of diary entries decreased over time, a pattern likely associated with depression and demotivation. Sentiment analysis using the Dictionary of Affect in Language indicated unusually high proportions of both Sad and Pleasant words, reflecting a combination of depressed and paradoxically positive emotional tones previously observed in suicide-related writings. Linguistic analysis showed that the diarist overused first-person singular pronouns and underused second- and third-person pronouns, suggesting strong self-focus. He also avoided many common English function words and employed long, complex sentences. The most overused words relative to everyday English were *philosophy*, *intercourse*, and *suicide*. Overall, the diary's language exhibits characteristics typical of both literary writing and documents associated with suicide.

Keywords

Suicide Diary, Emotion, Language

1. Introduction

A document of approximately twenty-one thousand words, presented as the diary of Wallace E. Baker, who died by suicide, was published in 1913. It appeared in *The Glebe*, a modernist and anarchist literary magazine based in New York, under the title *Diary of a Suicide* (Baker, 1913). Baker's body was discovered on a beach near Brooklyn, and evidence in the foreword of the published diary suggests that the text was modified prior to publication. According to the foreword, the author removed certain passages to protect individuals mentioned in the diary, and additional editorial changes might have been made by the magazine's editors. With

the exception of six brief experimental entries written in March 1912, the diary takes the form of a literary discussion, raising questions about its authenticity as a suicide-related document.

Several details point to Baker's literary and philosophical orientation. He was found with a torn photograph of the dramatist August Strindberg in his pocket, and he referenced Nietzsche four times in the diary. His final sentence reflects his mindset: "The majority is always wrong, and the minority of supermen and degenerates—Zolas, Ibsens, etc.—must band together and overthrow the whole damn system which drives the best, the most sincere and honest to suicide or starvation" (Baker, 1913). Not all readers consider the diary to be a genuine expression of suicidal thinking. The Loyal Books (n.d.), which offers free public domain books in readable format, refers to the diary as a novel, and congratulates the author on his skill in representing a suicidal mind-set.

The present research examines whether *Diary of a Suicide* exhibits linguistic and emotional characteristics commonly associated with suicide-related writings. The analyses focus on the diary's language, temporal pattern of production, emotional undertones, and vocabulary. These analyses involve counting words and entries and assigning emotional scores to words using an established dictionary of affective meanings (Whissell, 2009). Because these methods rely on quantitative rather than qualitative approaches, they reduce the degree of subjectivity in interpretation. Given that no definitive form of suicide diary exists, this study contributes to the growing body of quantitative research describing the emotional tone and linguistic style of writings associated with suicide (e.g., Lester, 2020; Pennebaker & Stone, 2004; Arowosegbe & Oyelade, 2023; Ishfaq & Sultan, 2026).

2. Method

Diary of a Suicide was downloaded from Project Gutenberg for analysis (Project Gutenberg, n.d.). The publisher-prepared foreword was excluded. The diary contained 20,841 words across 46 entries dated from January 26, 1912, to September 27, 1913; Baker's body was discovered on September 28, 1913. Entries 1 - 32 were written in Havana, whereas entries 33 - 46 were written in the United States, beginning in Chicago and ending in New York.

The frequency of entries was counted across the 21 months during which the diary was kept (January 1912 to September 1913). Emotional tone was assessed using the Dictionary of Affect in Language (DAL; Whissell, 2009), a sentiment analysis tool that assigns emotional scores to individual words along two bipolar dimensions: Pleasantness-Unpleasantness and Activation-Passivity. These scores are based on ratings collected in previous research, with transformed averages of 50 for each dimension and a possible range of 125 points. For example, the word suicide received a Pleasantness score of 12 and an Activation score of 104, indicating that it was perceived as emotionally unpleasant and highly active. In contrast, hope received a Pleasantness score of 119 and an Activation score of 54, indicating high pleasantness and average activation. The DAL successfully

matched and scored 93% of the diary's words (19,358 of 20,841).

Although the Dictionary of Affect in Language (DAL) has been successfully used to analyze historical texts written centuries ago, the gap between when a text was written and when it is evaluated remains a limitation. Because DAL scores reflect emotional ratings from late 20th-century readers, they capture how modern audiences interpret a text rather than how its original contemporary readers would have experienced it. An additional limitation is that the DAL cannot score every word. In this study, 7% of the words were unscored—a portion that is typically omitted during analysis. The DAL generally covers about 90% of a text's vocabulary, a rate significantly higher than that of other sentiment analysis techniques.

Extreme scores on each emotional dimension—defined as those falling within the first or fourth quartile when words in the Dictionary were weighted for frequency of occurrence—were used to classify words into eight emotional categories (see **Table 1**). Extreme words in the document could, and in fact did (8.2% of the time), appear in more than one extreme category. In addition, the diary's vocabulary was compared to word-frequency norms derived from a corpus of everyday English containing approximately one-third of a million words (words from televised new reports, students' essays, minutes of meetings, newspapers, and so on; Whissell, 1998) to identify words used at unusually high or low rates.

Table 1. Eight categories of emotional words defined in terms of their extreme pleasantness and activation scores.

Word Category	Words Included in the Category	Examples
Pleasant	Words with scores in the top quartile of the Pleasantness Dimension	Summer, Hope
Cheerful	Words with scores in the top quartile for both the Pleasantness and the Activation dimension	Children, Love
Active	Words with scores in the top quartile for the Activation dimension	Run, Enthusiasm
Tough	Words with scores in the top quartile for the Pleasantness dimension and the bottom quartile for the Activation dimension	Battle, Nervous
Unpleasant	Words with scores in the bottom quartile for the Pleasantness dimension	Blame, Sleepless
Sad	Words with scores in the lowest quartile for both the Pleasantness and the Activation dimension	Impossible, False
Passive	Words with scores in the lowest quartile for the Activation dimension	Down, Tired
Gentle	Words with scores in the highest quartile for the Pleasantness dimension but the lowest quartile for the Activation dimension	Comfort, Ease

Note: The lowest 25% of scores in a distribution belong to the first (bottom) quartile; the next 25% belong to the second quartile, the next 25% to the third quartile, and the final 25% to the fourth (highest) quartile.

3. Results and Discussion

3.1. Pattern of Diary Entries across Time

Baker, the presumed author, did not write consistently throughout the 21-month period covered by the diary. Instead, entries were produced in four bursts of decreasing size (see **Figure 1**). These bursts were separated by months in which no entries were written. The first burst spanned January to November 1912 (11 months, 32 entries), the second January to March 1913 (3 months, 10 entries), the third June to July 1913 (2 months, 3 entries), and the final burst consisted of a single entry written in September 1913. Both the duration of the bursts (11, 3, 2, and 1 months) and the number of entries within them (32, 10, 3, and 1) decreased exponentially over time.

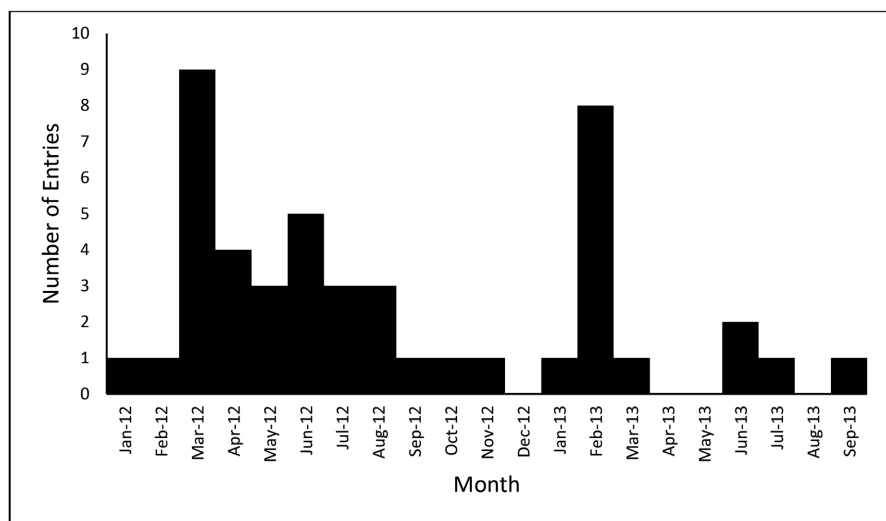


Figure 1. The distribution of diary entries in four bursts of decreasing size across 21 months.

This pattern might suggest a progressive decline in the author's motivation to write. Motivation appears strongest at the beginning of the diary, when entries were frequent and sustained, and weakest at the end, when only one final entry was produced. The wave-like fluctuations between bursts further support this interpretation. If the pattern were projected forward, the next burst would be expected to contain zero entries, consistent with the cessation of writing at the time of Baker's death. Caution is warranted when interpreting these findings, as the reasons behind writing frequency are inferred and alternative explanations exist. For instance, a decline in entries could stem from textual editing rather than diminished writing activity. Alternatively, it might reflect a shift toward a more positive mood rather than a more negative one.

Psychological research commonly associates depression with reduced motivation (Farokhi, 2024). Although Farokhi argues that individuals with depression may be motivated in different ways—such as toward managing internal negativity—the observable behavioral outcome often resembles demotivation. The de-

creasing frequency and duration of diary entries align with this interpretation, suggesting that depression and diminishing energy likely contributed to the observed decline in writing activity over time.

3.2. Use of Emotionally Extreme Words

All words scored by the Dictionary of Affect in Language (DAL) were classified as belonging or not belonging to each of the eight extreme emotional categories described earlier (Pleasant, Cheerful, Active, Tough, Unpleasant, Sad, Passive, and Gentle; see [Table 1](#)). The frequency of these words in the diary was compared with their frequency in a normative corpus of everyday English containing more than one-third of a million words ([Whissell, 1998](#)). Results are shown in [Figure 2](#). The dotted line represents the normative value (100%), and the solid line represents the diary's values expressed as percentages of the norm.

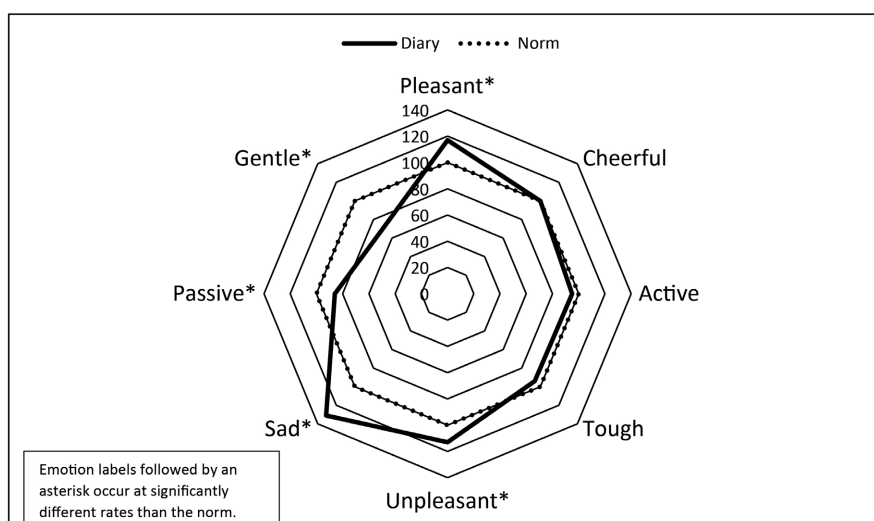


Figure 2. The use of eight types of emotional words (solid line) in comparison to their normative use (dotted line).

Several categories of extreme emotional words appeared more frequently in the diary than in the normative sample, including Sad, Unpleasant, and Pleasant words. Other categories, such as Passive and Gentle words, appeared less frequently. Categories marked with an asterisk in [Figure 2](#) indicate significant differences between the diary and the normative sample ($p < 0.001$, two-tailed z test comparing a sample proportion to a population proportion).

The reduced use of Passive and Gentle words suggests a departure from a linguistic “comfort zone”, which is consistent with the emotionally charged nature of suicide-related writing. The elevated use of Sad and Unpleasant words aligns with previous findings showing that negative emotional language is common in suicide-related documents (e.g., [Kim et al., 2019](#); [Pennebaker & Stone, 2004](#)). The increased use of Pleasant words parallels the “paradoxical positivity” described by [Leavitt et al. \(2021\)](#), who argue that writing about the decision to die by suicide

may temporarily reduce distress and lead to the use of more positive emotional language.

Overall, the diary contains unusually high proportions of both Sad and Pleasant words, reflecting a combination of depressed emotional tone and paradoxical positivity—patterns commonly observed in suicide-related writings.

3.3. Individual Words Employed at Unusually High and Low Rates

The diary contained 3310 unique words. A two-tailed z test comparing each word's frequency in the diary to its frequency in a normative corpus of everyday English (Whissell, 1998) was conducted for all individual words. Hundreds of words differed significantly from the normative sample at $p < 0.05$ (two-tailed). **Table 2** presents the 20 most overused and 20 most underused words, ranked by absolute z value. The analysis of word frequencies was exploratory: in order to maintain control of experiment-wise alpha (3310 tests were conducted), only extreme test results significant at $p < 0.000001$ are discussed.

Table 2. Words used more often and less often than normal in *Diary of a Suicide*.

Top 20 Words Used More Often than Normal ($z > 20$, $p < 0.00001$)	Top 20 Words Used Less Often than Normal ($z < -4.56$, $p < 0.00001$)
Philosophy	You
Intercourse	He
Suicide	His
Passion	Her
Passions	They
Weakness	The
I	She
My	We
Necessity	Was
Temperament	Are
Moods	Were
Enthusiasm	When
Genius	Like
Despair	Just
Myself	What
Ideals	No
Lately	Said
Literature	Their
Nervousness	About
Ambition	There

Note: The words listed in either column are illustrative: they were those with the highest absolute z values, indicating the most pronounced deviation from normal usage patterns. There were more words with significant usage effects.

The most overused words included both emotionally negative terms (e.g., *despair*, *suicide*) and emotionally positive terms (e.g., *enthusiasm*, *passion*), reflecting the simultaneous presence of Pleasant and Unpleasant emotional tones noted earlier. The word *intercourse*, which appears near the top of the overused list, was generally used to refer to sexual behavior. Three first-person singular pronouns—*I*, *my*, and *myself*—appeared at higher-than-normal rates, a pattern characteristic of both diary writing and suicide-related texts. In contrast, many second- and third-person pronouns (e.g., *he*, *you*, *her*), common English function words (e.g., *there*, *when*), and several forms of the verb *to be* (*was*, *are*, *were*) were used significantly less often than expected. Even the most common word in English, *the*, appeared at only 76% of its normative rate.

Sentence length, measured using Microsoft Word's Grammar function, averaged 24.6 words per sentence, placing the diary on the boundary between “fairly difficult” and “difficult” readability levels (Siteimprove Help Center, 2022). The combination of long sentences and reduced use of common function words likely contributes to the diary's literary tone. The Dictionary of Affect includes an additional rating scale for words: the scale for Imagery. Very concrete or highly imaged words can easily be pictured or envisioned (for example, book, face, or chair). Very abstract or poorly imaged words cannot be easily pictured (for example, justice, immediate, or perfection). The normative average for this scale is 50. The diary had a low Imagery mean of 46.11, which was significantly below the normative mean ($t = -15.95$, two-tailed $p < 0.001$, $d = 0.12$). The language of the diary was abstract. This would add further to its complexity.

Taken together, the word-frequency patterns mirror findings from previous research: the diarist overused both emotionally positive and emotionally negative words, relied heavily on first-person singular pronouns, and avoided many common English function words. The three most overused words—*philosophy*, *intercourse*, and *suicide*—underscore the diary's thematic and emotional focus. The diary included both abstract words and long sentences: these are indicators of linguistic complexity.

3.4. Changes across Time

As noted earlier, the number of diary entries declined substantially over the 21-month period during which the diary was written. In contrast, changes in emotional tone across time were minimal. A correlation analysis using month (1 - 21) as the temporal variable and individual words as the unit of analysis revealed a small but significant increase in the use of Tough words over time ($r = 0.03$, $p < 0.001$). Three additional analyses examined temporal trends in the use of the words *suicide* (including *suicides*), *intercourse*, and *philosophy*, all of which were used at unusually high rates in the diary. Relationships were again examined at the level of months. Monthly mentions of *suicide* increased slightly across months ($r = 0.04$, $p < 0.001$), whereas mentions of *philosophy* decreased slightly ($r = -0.01$, $p = 0.03$). No temporal trend was observed for the third highly overused word,

intercourse. Although statistically significant, these trends accounted for less than 1% of the variance.

Overall, the passage of time was associated with three very weak linguistic shifts: a small increase in Tough words, a slight rise in references to suicide, and a decrease in references to philosophy. These patterns are consistent with the emotional trajectory implied by the diary's content and the author's eventual death by suicide.

4. Conclusion

This study presented results from several quantitative analyses of Baker's (1913) *Diary of a Suicide*, focusing on its linguistic style and emotional undertones. Two primary conclusions emerge from the findings. First, the diary exhibits a distinctly literary style, characterized by long and complex sentences, abstract language, reduced use of common English function words, and frequent references to prominent literary figures of the early twentieth century. Second, the diary demonstrates emotional and linguistic patterns consistent with those found in suicide-related writings. The decreasing frequency of entries over time suggests that the author became increasingly demotivated as the date of his death approached. Although suicide-related writings vary widely across individuals and circumstances, the present research confirms several patterns reported in previous studies and contributes additional quantitative evidence to the field.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- Arowosegbe, A., & Oyelade, T. (2023). Application of Natural Language Processing (NLP) in Detecting and Preventing Suicide Ideation: A Systematic Review. *International Journal of Environmental Research and Public Health*, 20, Article 1514. <https://doi.org/10.3390/ijerph20021514>
- Baker, W. E. (1913). *Diary of a Suicide*. *The Glebe*, 1, 7-83.
- Farokhi, H. (2024). Why People with Depression Appear to Be Demotivated. *Philosophical Psychology*, 39, 1186-1211. <https://doi.org/10.1080/09515089.2024.2443642>
- Ishfaq, A. & Sultan, N. (2026). Language of Stress: A Corpus-Based Study to Detect Early Signs of Suicide through Lexical Choice. *Journal of Applied Linguistics and TESOL (JALT)*, 9, 862-880. <https://doi.org/10.63878/jalt2310>
- Kim, K., Choi, S., Lee, J., & Sea, J. (2019). Differences in Linguistic and Psychological Characteristics between Suicide Notes and Diaries. *The Journal of General Psychology*, 146, 391-416. <https://doi.org/10.1080/00221309.2019.1590304>
- Leavitt, J., Hong, J. H., & Walker, R. L. (2021). Paradoxical Positivity: Suicide Notes Use Less Distressed Language than Blogs about Depression, Suicidal Thoughts, and Even Cooking. *Suicide and Life-Threatening Behavior*, 51, 1005-1014. <https://doi.org/10.1111/sltb.12786>
- Lester, D. (2020). The Diary of a Suicidal Youth: A Linguistic Analysis. *Journal of Biomed-*

- ical Research & Environmental Sciences*, 1, 141-143. <https://doi.org/10.37871/jbres1132>
- Loyal Books (n.d.). *Diary of a Suicide*.
<https://www.loyalbooks.com/book/diary-of-a-suicide-by-wallace-e-baker>
- Pennebaker, J. W., & Stone, L. D. (2004). What Was She Trying to Say? A Linguistic Analysis of Katie's Diaries. In D. Lester (Ed.), *Katie's Diary* (pp. 54-79). Routledge.
- Project Gutenberg (n.d.). *The Glebe 1913/11 (Vol. 1, No. 2): Diary of a Suicide by Wallace E. Baker*. <https://www.gutenberg.org/ebooks/49261>
- Siteimprove Help Center (2022). *Readability: Why Are Long Sentences over 20 Words?* <https://help.siteimprove.com/support/solutions/articles/80000447968-readability-why-are-long-sentences-over-20-words->
- Whissell, C. (1998). A Parsimonious Technique for the Analysis of Word-Use Patterns in English Texts and Transcripts. *Perceptual and Motor Skills*, 86, 595-613E.
<https://doi.org/10.2466/pms.1998.86.2.595>
- Whissell, C. (2009). Using the Revised Dictionary of Affect in Language to Quantify the Emotional Undertones of Samples of Natural Language. *Psychological Reports*, 105, 509-521. <https://doi.org/10.2466/pr0.105.2.509-521>