

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Using Computational Text Analysis Tools to Compare the Lyrics of Suicidal and Non-Suicidal Songwriters

Permalink

<https://escholarship.org/uc/item/0dh4553j>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 29(29)

Authors

Lightman, Erin J.
McCarthy, Philip M.
Dufty, David F.
[et al.](#)

Publication Date

2007-01-01

Peer reviewed

Using Computational Text Analysis Tools to Compare the Lyrics of Suicidal and Non-Suicidal Songwriters

Erin J. Lightman (elightman@mail.psyc.memphis.edu)

Department of Psychology, 365 Innovation Dr.
Memphis, TN 38152 USA

Philip M. McCarthy (pmccarthy@mail.psyc.memphis.edu)

Department of Psychology, 365 Innovation Dr.
Memphis, TN 38152 USA

David F. Dufty (d.dufty@mail.psyc.memphis.edu)

Department of Psychology, 365 Innovation Dr.
Memphis, TN 38152 USA

Danielle S. McNamara (d.mcnamara@mail.psyc.memphis.edu)

Department of Psychology, 365 Innovation Dr.
Memphis, TN 38152 USA

Abstract

Recent research in cognitive science and psycholinguistics has provided evidence that extreme behaviors such as suicide may be predicted by analyzing an individual's writing style. This study seeks to extend research of this kind to a new and less traditional written genre: song lyrics. We applied the computational tools Coh-Metrix and Linguistic Inquiry and Word Count to a corpus of 35 songs each by 16 different songwriters in order to investigate how differences between suicidal and non-suicidal songwriters' lyrics reflect previous research on suicide. The results of our study suggest that suicidal songwriters use words of lower concreteness, fewer words, more future verbs, and fewer death-themed words in their lyrics. Overall, our results contrast notably with previous research, providing a new perspective on deriving psychological information from non-traditional texts.

Keywords: Coh-Metrix; cohesion; coherence; difficulty; text analysis; suicide; lyrics; songwriters.

Introduction

To study an individual's cognitive processes and personality traits, we often look at that individual's use of language. The psychological implications of language use have been studied with computational tools in a variety of mediums, ranging from classical literature and poetry (e.g., Stirman & Pennebaker, 2001) to transcripts from therapy sessions (e.g., Stone et al., 2000). Lyrics from music however, have rarely been considered in this context. While clear distinctions can be made between song lyrics and more conventional forms of text, such as textbooks and novels, the relationship between song lyrics and poetry has been widely acknowledged (Guerra, in press). As such, this study seeks to investigate the linguistic features of song lyrics to study musicians, particularly in terms of suicidal behavior.

Destructive behaviors such as suicide are considered to be a greater risk for musicians than for the general

population (Raeburn, 1999; Stack, 1997). The current study's design draws from theoretical models and research on suicide using two diverse but complimentary computerized textual analysis tools: Coh-Metrix (Graesser et al., 2004) and Linguistic Inquiry and Word Count (LIWC, Pennebaker & Francis, 1999). The analyses produced by these computational tools are potentially capable of revealing aspects of the cognitive and affective processes involved in suicidal thought.

Over the past 40 years, research using computational text analysis tools has made substantial developments in the fields of cognitive science and psycholinguistics (Jurafsky & Martin, 2000). For example, Newman et al. (2003) used LIWC to predict whether an individual is telling a truthful or fabricated story. And Pennebaker and Stone (2002) used LIWC to explore the relationship between language and aging using a large and diverse corpus of texts, including novels, transcripts of conversations, and classical poetry.

Along with personality and demographic traits, research has addressed a variety of clinical problems including suicide (McDermott & Porter, 1989; Stirman & Pennebaker, 2001). For instance, Thomas and Duszynski (1985) found text analysis to be an effective predictor of suicide among medical students. And Silverman and Will (1986) computationally analyzed the work of Sylvia Plath, a famous poet who committed suicide, and found that she shifted from using a somewhat traditional form of writing to a more personal, expressive form over the years as she neared suicide¹.

Much of this research has been based on psychoanalytic theories of suicide, such as Durkheim's social integration/disengagement model. Durkheim's model

¹ However, McCarthy et al. (2006) used Coh-Metrix to analyze stylistic changes in writing throughout the careers of three well-known authors finding that each of them significantly shifted in style over their careers although none of them committed suicide.

suggests that suicidal individuals tend to become detached from social life as a result of failing to effectively integrate into society (Durkheim, 1951; Petrie & Brook, 1992). Related theories involving disengagement (Farberow & Marmor, 1996; Prezant & Neimeyer, 1988) suggest that suicidal individuals withdraw from social relationships, try to detach from the sources of their pain, and become more self absorbed. More traditional theories of suicide (e.g., Beck et al., 1979) are based on hopelessness models and suggest that suicide takes place during prolonged periods of sadness and desperation. The overwhelming feeling of hopelessness and helplessness, along with the inclination to think of the world in terms of absolutes, leads a depressed individual to conclude that suicide is the only way to stop the pain (Stirman & Pennebaker, 2001).

The models of suicide described above provided a base for Stirman and Pennebaker (2001), who investigated how the writing styles and characteristics of writers who have killed themselves differ from writers who have not. Stirman and Pennebaker applied LIWC to a corpus of 300 poems written by suicidal or non-suicidal poets, and found that suicidal poets used more first-person singular pronouns than the non-suicidal poets. No other significant results emerged, but directions of some categories agreed with their predictions. The suicidal poets in their corpus used *communication words* (e.g., *talk, share, listen*) less often and used *death-themed words* (e.g., *dying, coffin, funeral*) more often as they neared the ends of their lives.

Building on the work of Stirman and Pennebaker, we assessed song lyrics by suicidal and non-suicidal songwriters with two computational tools. The first tool, Coh-Metrix (Graesser et al., 2004) computes over 700 indices of textual cohesion, difficulty and conceptual characteristics (Graesser et al., 2004). The system evaluates texts using Latent Semantic Analysis (LSA, Landauer & Dumais, 1997; Landauer et al., 2006), lexicons, pattern classifiers, part-of-speech taggers, syntactic parsers, shallow semantic interpretations, and other components developed in computational linguistics research (Jurafsky & Martin, 2000). Coh-Metrix incorporates a large number of lexicons including Celex (Baayen, Piepenbrock, & Van Rijn, 1993), WordNet (Miller et al., 1990), the MRC Psycholinguistic Database (Coltheart, 1981), and lexicons that estimate the frequency of word use. These lexicons allow Coh-Metrix to assess each word on various components such as number of syllables, abstractness, imagery, ambiguity, frequency of usage, and age of acquisition (Graesser et al., 2004). In the output, Coh-Metrix offers a linguistic representation of practically any given text.

The second computational tool we used in this study is Linguistic Inquiry and Word Count (LIWC, Pennebaker, Francis, & Booth, 2001). LIWC was developed in response to a need for an efficient and effective method for examining the many cognitive, emotional, structural, and process components in writing and speech (Pennebaker et al., 2001). To capture these concepts, the text analysis program uses a word count strategy whereby LIWC

searches a text for more than 2300 words or word stems within its dictionary. These search words are categorized within the tool into over 70 linguistic dimensions, including standard language categories (e.g., articles, prepositions, pronouns), psychological processes (e.g., positive and negative emotion words), relativity-related words (e.g., time, verb tense, spatiality), and content dimensions (e.g., death, sex, home, occupation). The output reveals how much these categories and subcategories of words are represented in a text (Pennebaker, Mehl, & Niederhoffer, 2003).

Five LIWC dimensions and five Coh-Metrix indices were selected to investigate the writing styles and characteristics of suicidal and non-suicidal songwriters. These choices were based on previous research concerning the relationship between psychological health and language use (e.g., Stirman & Pennebaker, 2001; McDermott & Porter, 1989), and on theoretical models of suicide (e.g., Durkheim, 1951; Petrie & Brook, 1992). The following are brief descriptions of the 10 indices/dimensions and rationale for their predictive power. More detailed explanations can be found for Coh-Metrix in Graesser et al. (2004) and for LIWC in Pennebaker et al. (2001).

Coh-Metrix Indices

Argument Overlap This index measures aspects of cohesion. The argument overlap index that we used, *adjacent*, refers specifically to referential similarities such as *table/tables* and *it/table*, in adjacent sentences (Kintsch, 1998). Using both Coh-Metrix and LIWC, Phillips-Stoll and Schober (2006) reported that currently depressed college students write less cohesively than never depressed and formerly depressed college students. Considering the widely acknowledged relationship between depression and suicide (Dieserud et al., 2001), we predicted that the lyrics by suicidal songwriters would be less cohesive than the lyrics by non-suicidal songwriters, as indicated by lower argument overlap values.

Latent Semantic Analysis (LSA) LSA is a statistical algorithm that compares the contextual similarity between words, sentences, paragraphs, and complete texts, reflecting that text's cohesiveness (Landauer et al., 2006). LSA, like argument overlap, is a measure of cohesion. LSA differs from argument overlap in that values are produced for relative similarities. For example, *pen/pens* generates a higher value than *pen/paper*, which in turn generates a higher value than *pen/octopus*. Based on our prediction that suicidal songwriters would write less cohesively, we predicted that suicidal artists' lyrics would have lower LSA values than non-suicidal artists.

Word Concreteness This index measures a word's concreteness based on human ratings. For example, words such as *garden* and *door* score more highly than words such as *envy* or *despair*. Disengagement theories have suggested that suicidal individuals attempt to escape their source of

pain (e.g., Durkheim, 1951) and detach from the real, concrete world, which would suggest that suicidal individuals might think in less concrete terms. We therefore predicted that this escapism instinct would be indicated by less concrete language.

Tense/Aspect Ratio This index measures an aspect of temporal cohesion: the degree to which sentences are linked by time relations. A popular hypothesis on temporality and suicide is that suicidal individuals experience distorted time perception in which the present seems never-ending (Neuringer & Harris, 1974; Twenge et al., 2003). We predicted that this focus on the present would be reflected by lower *tense/aspect ratio*, meaning that tense/aspect would shift less often between representations of past, present, and future.

Word Count This index is simply the number of words in a text. Phillips-Stoll and Schober (2006) suggest that depressed individuals, when told to write about something, tend to write less than non-depressed individuals. Thus, we predicted that the songs written by suicidal songwriters would generally have fewer lyrics.

Linguistic Inquiry and Word Count Dimensions

First, it should be noted that some of these dimensions required more than one measure to provide a comprehensive representation. Although 5 dimensions were selected, 9 LIWC measures were incorporated.

References to Self and Others To assess self absorption and relationships with others, we used three measures: *first person singular pronouns*, *total first person references*, and *references to other people* (first person plural, second person, and third person pronouns). Various studies have found differences in pronoun use between people as a function of depression. For instance, Bucci and Freedman (1981) examined speech samples of five elderly depressed subjects and found a higher frequency of first person singular pronouns and a lack of second and third person pronouns. They interpreted these results as a limited ability or desire to connect with others. Drawing from this literature, we predicted that lyrics by suicidal songwriters would more first person pronouns and all self references as well as fewer references to other people.

Emotion Words The words that address affective or emotional processes can be divided into two categories: *positive emotion words*, such as *happy* and *pretty*, and *negative emotion words*, such as *enemy* and *worthless*. Because of the close relationship between suicidal behavior and depression (Dieserud et al., 2001), and because depression is often characterized by negative affect, we predicted that suicidal artists would use fewer *positive emotion words* and more *negative emotion words* than non-suicidal artists. However, we note that Stirman and Pennebaker (2001) found no significant differences in either category for suicidal and non-suicidal poets.

Communication Words These are words that signify communication, such as *talk*, *share*, and *converse*. Based on Durkheim's (1951) social integration/disengagement model, we predicted that the suicidal songwriters would use fewer *communication words* than their non-suicidal counterparts.

References to Time We measured the incorporation of verbs referred to as future tense and words related to time in general. Cognitive aging literature has suggested that perception of time is associated with social and emotional adjustment. This theory indicates that individuals with socially restricted goals think of themselves as having limited futures, and individuals with larger networks of friends have more future-oriented perspectives of time (Carstensen, Isaacowitz, & Charles, 1999; Danner, Snowdon, & Friesen, 2001). Applying this concept to the social disengagement aspect of Durkheim's model (1951), we predicted that suicidal songwriters would use fewer *future-tense verbs*, such as *shall* or *will*, than non-suicidal songwriters. Drawing from the idea that suicidal individuals perceive time as interminable (Neuringer & Harris, 1974), we predicted that suicidal artists would also focus more heavily on time, and thus, use more *time-related words*.

Death-Themed Words This category includes 29 words related to death and dying, such as *dead*, *burial*, and *coffin*. Preoccupation with death seems to be an understood element of suicidal thought (e.g., Durkheim, 1951). Hence, we predicted that suicidal songwriters would use more *death-themed words* than non-suicidal songwriters.

Method

Two groups of song lyrics were collected. We selected eight songwriters who committed suicide for the first group. Each suicidal artist was matched with a songwriter who has not shown signs of suicidal intentions (i.e., no reported attempts). The songwriters were matched as closely as possible for genre of music, date of birth, degree of fame, nationality, mental status, presence or absence of addiction, and education.

The artists were matched in order to control for variables other than suicidal action; during statistical analysis, the groups were evaluated as independent samples. Our corpus does not include female musicians, because few of the female musicians who have committed suicide wrote extensively. All artists were native English speakers and represent diverse regions of the United States and the United Kingdom. We were careful to select a wide range of music encompassing traditional genres such as classic rock, country, and folk music as well as more recently developed genres such as punk and grunge (see Table 1).

Each artist is represented by the lyrics of 35 songs obtained from websites devoted to that artist or his band. Because some of the websites did not specify when choruses or phrases were repeated, many of the songs were listened to in order to correct any omission errors. Lyrics for each song were processed through both Coh-Metrix 2.0 and LIWC.

Table 1: Suicidal songwriters, their non-suicidal matches and general information.

| Suicidal Songwriter | Counterpart | Genre | Nationality | Birth | Death | Cause of Death |
|---------------------|-----------------|-------------|---------------|-------|-------|---------------------|
| Stuart Adamson | Paul Weller | punk/rock | England | 1958 | 2001 | hanging |
| Adrian Borland | Mark Smith | rock | England | 1957 | 1999 | crushed by train |
| Kurt Cobain | Chris Cornell | grunge | United States | 1967 | 1994 | gunshot to the head |
| Ian Curtis | David Byrne | punk | England | 1956 | 1980 | hanging |
| Tom Evans | Joey Molland | rock/pop | England | 1947 | 1983 | hanging |
| Pete Ham | George Harrison | rock/pop | England | 1947 | 1975 | hanging |
| Shannon Hoon | Scott Weiland | alternative | United States | 1967 | 1996 | overdose: heroin |
| Phil Ochs | Tom Paxton | folk | United States | 1940 | 1976 | hanging |

Results

Of the 14 selected indices/measures, 4 revealed significant differences in paired samples *t*-test (see Tables 2 and 3). As shown in Table 2, the Coh-Metrix analyses indicated that the lyrics by the *non-suicide* group incorporated more *concrete* language than the lyrics by the *suicide* group. The mean difference between groups was -4.73, and the 95% confidence interval for the estimated population mean difference was between -8.33 and -1.13. An independent *t*-test showed that the differences between conditions was significant ($t(279) = -2.59, p = 0.01$, two-tailed).

Coh-Metrix analyses also indicated that the average *word count* per song was larger for the non-suicidal songwriters. The mean difference between groups was -13.8, and the 95% confidence interval for the estimated population mean difference was between -26.94 and -0.66. An independent *t*-test showed that the difference between groups was significant ($t(279) = -2.07, p = 0.04$).

Table 2: Coh-Metrix indices in terms of Argument Overlap (AO), Latent Semantic Analysis (LSA), Concreteness, Tense/Aspect Ratio, and Word Count for Suicidal and Non-Suicidal Songwriters.

| | suicide | non-suicide |
|--------------|-----------------|-----------------|
| AO | 0.21 (0.15) | 0.23 (0.16) |
| LSA | 0.19 (0.14) | 0.21 (0.15) |
| concreteness | 328.68 (20.17)* | 333.41 (24.17)* |
| tense/aspect | 0.89 (0.11) | 0.88 (0.11) |
| word count | 193.40 (79.13)* | 207.20 (80.25)* |

Notes: standard deviations are in parentheses; * $p < 0.05$

As shown in Table 3, LWIC analyses indicated that *future-tense verbs* appeared more frequently in the lyrics by suicidal songwriters. The mean difference between conditions was 0.39 and the 95% confidence interval for the estimated population mean difference was between 0.03 and 0.75. An independent *t*-test showed that the difference between conditions was significant ($t(279) = 2.15, p = 0.03$). The lyrics by non-suicidal songwriters also contained more

death-themed words. The mean difference between conditions was -0.18, and the 95% confidence interval for the estimated population mean difference was between -0.32 and -0.03. An independent *t*-test showed that the difference between groups was significant ($t(279) = -2.41, p = 0.02$).

Table 3: LIWC Dimensions in terms of Pronoun Use, Emotion Words, Communication Words, Time Words, Future Tense Verbs, and Death-Themed Words for Suicidal and Non-suicidal Songwriters.

| | suicide | non-suicide |
|----------------------|--------------|--------------|
| 1st person singular | 7.53 (5.37) | 7.49 (6.63) |
| all 1st person | 8.90 (5.31) | 8.40 (6.42) |
| references to others | 7.71 (4.84) | 7.97 (5.00) |
| positive emotions | 2.60 (2.59) | 2.99 (3.50) |
| negative emotions | 2.22 (2.47) | 2.03 (2.26) |
| Communication | 1.68 (2.02) | 1.71 (2.15) |
| time words | 4.36 (3.38) | 4.11 (3.55) |
| future verbs | 1.94 (2.42)* | 1.55 (2.00)* |
| death words | 0.25 (0.65)* | 0.43 (1.02)* |

Notes: standard deviations are in parentheses; * $p < 0.05$

Discussion

In this study we used the computational tools Coh-Metrix and LIWC to investigate how differences between suicidal and non-suicidal songwriters' lyrics reflect previous research and theoretical models on suicide. The corpus consisted of 560 songs, half written by suicidal songwriters and half written by non-suicidal songwriters. Our focus on songwriters addresses the higher than average rate of suicide for musicians.

The Coh-Metrix results partly adhered to our original predictions, although the results of this study did not strongly replicate previous research on suicidal behavior using computational text analysis. The *non-suicide* group's higher use of concrete words was in line with our predictions and could possibly lend support to disengagement models of suicide (Durkheim, 1951; Petrie

& Brook, 1992). In addition to corresponding with our predictions, the lower word frequency for the *suicide* group was congruent with Phillips-Stoll and Schober's (2006) research on depressed college students. However, whereas the suicidal songwriters used more abstract language and wrote songs with fewer lyrics, the *argument overlap*, *LSA*, and *tense/aspect ratio* did not yield significant results.

The values for the five LIWC dimensions did not adhere to our predictions. Although there was a trend of suicidal songwriters using more *self-references* and fewer *references to others*, the distinctions were not significant. There were also no significant differences in frequency of *emotion words*, *references to time*, and *communication words*. The values for *future-tense verbs* and *death-themed words* were not in agreement with our predictions and contrast notably with previous research. While most psychoanalytic literature suggests that suicidal individuals are preoccupied with thoughts of death and the interminability of the present, the *suicide* group used significantly more *future-tense verbs* and fewer *death-themed words* than the *non-suicide* group.

The suicidal group's lower usage of *death-themed words* might be partially explained by the therapeutic benefits of writing. Therapeutic writing is based on the idea that writing about negative emotions and events can be cathartic, and has been heavily implemented into counseling and research (e.g., Smyth, 1998). The eight suicidal songwriters in the corpus were matched with eight non-suicidal songwriters based on several criteria including mental health and addiction. Therefore, it is a reasonable assumption that most or all of these individuals experienced depression on some level. We cannot say that writing more about death and dying relates to the non-suicidal songwriters' choices to live, but the results might suggest a healthier and more progressive mental state.

Along with *death-themed words*, the positive results for *future-tense verbs* and null results for *argument overlap*, *LSA*, *positive* and *negative emotion words*, *first person pronouns*, *references to others*, and *communication words* did not support previous research on poetry and writing samples. These dissimilarities may be the outcome of distinctions between song lyrics and other forms of writing.

There are several factors that may explain why the medium of song lyrics might have yielded results different from poetry and other forms of writing. First, people other than the credited songwriters, such as band members and record producers, may have shaped the lyrics during the writing and editing processes: Whereas most traditional writers work independently of co-authors, many songwriters have band members (who have not committed suicide) with whom they collaborate in the writing process. Second, while there is strong evidence that computational tools like Coh-Metrix and LIWC can interpret mental states from text, the tools are not designed to recognize subtleties such as irony. Irony is often prevalent in music; a melancholy song can have positive lyrics set to sorrowful melodies and timbre. The abstractness of songwriting might make it impractical

to analyze lyrics with computational tools. Third, there are important contextual differences between music and traditional writing. A poem or novel can be studied through its text alone, but music is arguably best presented through auditory means. Perhaps elements of songwriters' mental states cannot be interpreted through their lyrics alone because all of music's components must be considered.

Results may have also been influenced by limitations in the corpora. The lengths of the individual songs may have affected the results. Most of the text files consisted of fewer than 200 words, and some of the computational algorithms that we used are more reliable with larger text samples. In addition, our corpus consisted of sixteen songwriters. This rather low number is due to our constraints of carefully matching the suicidal and non-suicidal song-writers. Nonetheless, these sample size constraints may account for the limited effect sizes. Our future analyses will attempt to include a larger and more diverse corpus. Our goal is to collect corpora for additional songwriters, incorporating more diverse genres, time periods, and nationalities. We will also collect corpora by poets and novelists in order to further examine the measures we used in this study and replicate results of previous studies. Future research will also investigate stylistic changes in writing throughout the careers of suicidal songwriters.

Coh-Metrix and LIWC have emerged as valuable tools in cognitive science. Given the scope of these resources, it may be possible to improve upon existing models of suicide and maybe even create new models. Our goal for this research is to find elements of language that differ between suicidal and non-suicidal songwriters. These differences could provide insight into the mental states behind suicide, and thus offer potential in developing preventative treatments for musicians, and perhaps, the general population.

Acknowledgments

This research was supported by the Institute for Education Sciences (IES R305G020018-02). Opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the IES. The authors would also like to acknowledge the contributions made to this paper by Art Graesser, Max Louwerse, and Zhiqiang Cai.

References

- Baayen, R.H., Piepenbrock, R., & van Rijn, H. (Eds.) (1993). The CELEX lexical database [CD-ROM]. University of Pennsylvania, Philadelphia: Linguistic Data Consortium.
- Beck, A.T., Rush, A.J., Shaw, B.F., Emery, G. (1979). *Cognitive therapy of depression*. New York: Guilford Press.
- Bucci, W., & Freedman, N. (1981). The language of depression. *Bulletin of the Menninger Clinic*, 45, 334-358.

- Carstensen, L.L., Isaacowitz, D.M., & Charles, S.T. (1999). Taking time seriously: A theory of socioemotional selectivity. *American Psychologist*, *54*, 165-181.
- Coltheart, M. (1981). The MRC psycholinguistic database quarterly. *Journal of Experimental Psychology*, *33A*, 497-505.
- Danner, D., Snowdon, D., & Friesen, W. (2001). Positive emotions in early life and longevity: Findings from the Nun Study. *Journal of Personality and Social Psychology*, *80*, 814.
- Dieserud, G., Roysamb, E., Ekeberg, O., & Kraft, P. (2001). Toward an integrative model of suicide attempt: A cognitive psychological approach. *Suicide and Life-Threatening Behavior*, *31*, 153-168.
- Durkheim, E. (1951). *Suicide*. New York: Free Press.
- Farberow, N.L., & Marmor, K. (1996). Short-term group psychotherapy with suicidal patients. In M. Rosenbaum (Ed.) *Handbook of short-term therapy groups. The master work series*. Northvale: Jason Aronson Inc. (pp 337-355).
- Graesser, A.C., McNamara, D.S., Louwerse, M.M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, *36*, 193-202.
- Guerra, B. (in press). Music is poetry. *Middle School Language Arts Lessons*. Retrieved November 6, 2006, from <http://www.youthfirst.org/winners/music-poetry.htm>.
- Jurafsky, D., & Martin, J.H. (2000). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Upper Saddle River, NJ: Prentice-Hall.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, UK: Cambridge University Press.
- Landauer, T.K., & Dumais, S.T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*, 211-240.
- Landauer, T., McNamara, D.S., Dennis, S., & Kintsch, W. (Eds.) (2006). *Handbook of Latent Semantic Analysis*. Mahwah, NJ: Erlbaum.
- McCarthy, P.M., Lewis, G.A., Dufty, D.F., & McNamara, D.S. (2006). Analyzing writing styles with Coh-Metrix. In *Proceedings of the Florida Artificial Intelligence Research Society International Conference (FLAIRS)*. Menlo Park, California: AAAI Press.
- McDermott, J.F., & Porter, D. (1989). The efficacy of poetry therapy: A computerized content analysis of the death poetry of Emily Dickinson. *Psychiatry*, *52*, 462-468.
- Miller, G.A., Beckwith, R., Fellbaum, C., Gross, D., & Miller, K. (1990). *Five papers on WordNet* (Technical Report). Princeton University, Cognitive Science Laboratory.
- Neuringer, C., & Harris, R.M. (1974). The perception of the passage of time among death-involved hospital patients. *Life-Threatening Behavior*, *4*, 240-254.
- Newman, M.L., Pennebaker, J.W., Berry, D.S., & Richards, J.M. (2003). Lying words: Predicting deception from linguistic style. *Personality and Social Psychology Bulletin*, *29*, 665-675.
- Pennebaker, J.W., & Francis, M.E. (1999). *Linguistic inquiry and word count (LIWC)*. Mahwah, NJ: Erlbaum.
- Pennebaker, J.W., Francis, M.E., & Booth, R.J. (2001). *Linguistic Inquiry and Word Count (LIWC): LIWC2001*. Mahwah: Lawrence Erlbaum Associates.
- Pennebaker, J.W., Mehl, M.R., & Niederhoffer, K.G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology*, *54*, 547-577.
- Pennebaker, J.W., & Stone, L.D. (2002). Words of wisdom: language use across the lifespan. *Journal of Personality and Social Psychology*, *85*, 291-301.
- Petrie, K., & Brook, R. (1992). Sense of coherence, self-esteem, depression and hopelessness as correlates of reattempting suicide. *British Journal of Clinical Psychology*, *31*, 293-300.
- Phillips-Stoll, A.K., & Schober, M.F. (2006, July). *Textual cohesion in writing by depressed, formerly-depressed, and never-depressed students*. Presentation at the 16th Annual Meeting of the Society for Text and Discourse, Minneapolis, MN.
- Prezant, D.W., & Neimeyer, R.A. (1988). Cognitive predictors of depression and suicide ideation. *Suicide Life Threatening Behavior*, *18*, 259-264.
- Raeburn, S. D. (1999) Psychological issues and treatment strategies in popular musicians: A review, part 1. *Medical Problems of Performing Artists*, *14*, 4.
- Silverman, M.A., & Will, N.P. (1986). Sylvia Plath and the failure of emotional self-repair through poetry. *Psychoanalytic Quarterly*, *55*, 99-129.
- Smyth, J.M. (1998). Written emotional expression: Effect sizes, outcome types, and moderating variables. *Journal of Consulting & Clinical Psychology*, *66*, 174-184.
- Stack, S. (1997). Suicide among artists. *The Journal of Social Psychology*, *137*, 129-130.
- Stirman, S.W., & Pennebaker, J.W. (2001). Word use in the poetry of suicidal and non-suicidal poets. *Psychosomatic Medicine*, *63*, 517-522.
- Stone, A., Smyth, J., Kaell, A., & Hurewitz, A. (2000). Structured writing about stressful events: Exploring potential psychological mediators of positive health effects. *Health Psychology*, *19*, 619-624.
- Thomas, C.B., & Duszynski, K.R. (1985). Are words of the Rorschach predictors of disease and death? The case of "whirling." *Psychosomatic Medicine*, *46*, 201-211.
- Twenge, J.M., Catanese, K.R., & Baumeister, R.F. (2003). Social exclusion and the deconstructed state: Time perception, meaninglessness, lethargy, lack of emotion, and self-awareness. *Journal of Personality and Social Psychology*, *85*, 409-423.