

# Things We Said Today: A Linguistic Analysis of the Beatles

Keith J. Petrie  
The University of Auckland

James W. Pennebaker  
University of Texas

Borge Sivertsen  
University of Bergen

The musical group the Beatles had a profound impact on Western culture during their time together between 1960 and 1970. The three songwriters, John Lennon, Paul McCartney, and George Harrison, together and separately evolved in their lyrical styles over time. Using a new generation of computer text analyses, the Beatles' lyrics were analyzed to address how the group changed as a unit over time, how the various members changed in their writing styles, and the overlap in lyrical styles from one composer to the next. Overall, the Beatles' lyrics became darker, more psychologically distant, and less immediate over time. Paul McCartney's lyrical style proved to be more variable and broad ranging than either Lennon or Harrison. Using latent semantic analyses, Harrison's lyrics were more influenced by Lennon than by McCartney. Finally, the lyrics jointly written by Lennon and McCartney were mathematically more similar to Lennon's linguistic styles than McCartney's.

*Keywords:* Beatles, linguistic analysis, latent semantic analysis, lyrics

In early 1963, the Beatles released the song "Please Please Me," which immediately climbed to the top of the British charts. By the time the group disintegrated in 1970, they had recorded over 200 songs and had influenced music, politics, fashion, and culture like no other group before or since. Because of their popularity and the murder of John Lennon in 1980, the history of the Beatles has become a cultural story or narrative. In this study, we examine whether a psychological analysis of the Beatles lyrics provides insight into changes within the group over time and enables a comparison of characteristics of the main songwriters and their influence in the group.

The success of the Beatles was powered by their two lead songwriters, Paul McCartney and John Lennon. A number of well-known early Beatle songs were the result of a close collaboration between them such as "From Me to You," "She Loves You," and "Love Me Do." However, over time the two wrote separately with the other occasionally contributing part of a melody or phrase to help complete a song. While both Lennon and McCartney grew up in working-class areas of Liverpool, they came from quite different family backgrounds, which undoubtedly shaped their later song writing.

Paul McCartney was the eldest son of an accomplished jazz band musician. After Paul's mother died of breast cancer when he was 14, his father gave him a guitar, which Paul played constantly. Despite the sudden loss of his mother, Paul's family background was a warm, stable, and happy one. Paul has generally been described as being outgoing and popular with both teachers and his classmates (Salewicz, 1986). On the other hand, John Lennon had a more traumatic upbringing. After returning from the war, John's father left his wife and John was eventually given by his mother to her sister and her husband to raise as an only child. John only saw his mother occasionally, and she was later killed in a motor accident when he was 17. A bright and rebellious child, John was often in trouble at school for questioning his teachers' authority and disrupting classes. Like McCartney, the death of his mother prompted John to immerse himself in music, resulting in his starting his own band, which eventually became the Beatles.

McCartney was generally thought of as the melodic tunesmith of the Lennon–McCartney pairing, as he was responsible for many of the enduring standards such as "Yesterday," "Michelle," and "Let it Be." McCartney's style is often characterized as optimistic, wistful, and harmonious (Everett, 1999). His songs also cover an eclectic range of styles. McCartney has also been characterized psychologically as being less introspective, more exuberant, more resilient to external stress, and less focused on his own distress than Lennon (Brog, 1995).

Lennon is often thought to be the more intellectual, political, and cynical of the Beatles. He is also seen as writing more creative and thoughtful lyrics than McCartney. Whereas McCartney often wrote about everyday situations, Lennon had a more global or even cosmic perspective, penning songs such as "Revolution," and "Across the Universe." Emotionally, Lennon's music is often described as less optimistic and more melancholic than McCartney with songs such as "Help!," "You've Got to Hide Your Love

---

Keith J. Petrie, Department of Psychological Medicine, The University of Auckland, Auckland, New Zealand; James W. Pennebaker, Department of Psychology, University of Texas; Borge Sivertsen, Department of Clinical Psychology, University of Bergen, Bergen, Norway.

Preparation of this manuscript was aided by a grant from the National Institutes of Health (MH52391), a contract from the Army Research Institute (W91WAW-07-C-0029), and by the statistical help of R. Sherlock Campbell.

Correspondence should be addressed to Keith J. Petrie, Department of Psychological Medicine, University of Auckland, Private Bag 92019, Auckland, New Zealand. E-mail: kj.petrie@auckland.ac.nz

Away," "Nowhere Man," and the song written about his mother, "Julia." As Lennon explained: "He [McCartney] provided a lightness, an optimism, while I would always go for the sadness, the discords, the bluesy notes" (Sheff & Golson, 1981, p. 148).

Although not as productive or well known as Lennon and McCartney, George Harrison also contributed an increasing number of songs to the Beatles' albums over time. George Harrison was the youngest of four children of a working-class south Liverpool family. Harrison was also the youngest of the Beatles, and looked up to Lennon who was three years ahead at school. Known as the quiet member of the Beatles, his introspective nature led him to try song writing and his first song "Don't Bother Me" appeared on the Beatles' second album. Over time, Harrison became more prolific and successful as a songwriter, penning "Here Comes the Sun" and the standard "Something" on the Abbey Road album. Over the career of the Beatles, Harrison became more interested in Indian music and philosophy, which influenced his later song writing.

The career of the Beatles can be divided into three main periods of relatively similar productivity. The first stage started from the establishment of the band until their attaining worldwide popularity in 1964. During this period the Beatles toured extensively in the United Kingdom and overseas. Songs were written and recorded quickly and predominantly dealt with simple emotions and romantic love (Cook & Mercer, 2000). The years 1965 to 1967 represent the most creative and significant output of the group. During this period the band abandoned touring and focused more on developing new music ideas in the studio. The band's output during this period included the albums "Help," "Rubber Soul," "Revolver," and "Sgt Pepper's Lonely Heart Club Band." The songs over this period were more complex and dealt with a greater variety of topics than romantic love (Cook & Mercer, 2000). The final stage of the Beatles career was from 1968 to 1970 and started with the Beatles' last trip together to visit the Maharishi Mahesh Yogi. The trip ended in disappointment but provided creative stimulation for the "White Album" which came out later in 1968. The albums "Let it Be" and "Abbey Road" were also produced during this period. While the output during this period remained high, the Beatles themselves were working more individually and relations in the band deteriorated until the band finally folded in 1970.

The goal of the current project was to apply two relatively new computerized text analysis methods to the lyrics of the Beatles. By analyzing the individual words in the collected songs written by the group members, we tracked the development of their music over time and compared the relative contributions of Lennon, McCartney, and Harrison.

Very broadly, two general strategies are available to analyze word usage—a word count method and a word pattern analysis (Pennebaker, Mehl, & Niederhoffer, 2003). A simple word count method calculates the percentage of different word categories within any given text. Most computerized approaches measure language categories such as standard grammatical units (personal pronouns, prepositions) or psychologically derived linguistic dimensions (e.g., emotion words, achievement-related words). Word count strategies are based on the assumption that the words people use convey psychological information over and above their literal meaning and independent of their semantic context. In the current study, we relied on the text analysis program, Linguistic Inquiry and Word Count (LIWC, Pennebaker, Booth, & Francis, 2007).

The second text analysis approach is a word pattern analysis which has emerged from the artificial intelligence community. Rather than exploring text "top down" within the context of previously defined psychological content dimensions or word categories, word pattern strategies mathematically detect "bottom-up" how words covary across samples of text (Foltz, 2003). One particularly promising strategy is latent semantic analysis (LSA) (e.g., Landauer & Dumais, 1997), which is akin to a factor analysis of individual words. By establishing the factor structure of word use within a large number of writing samples, it is possible to learn how any new writing samples are similar to one another. Traditionally, this technique has been used to determine the degree to which two texts are similar in terms of their content. In the analysis of lyrics, for example, LSA can establish the degree that any two authors tend to draw on similar content themes. Recently, this logic has been extended to determining how similar any text samples are in terms of their writing style—through the use of the analysis of function words, including pronouns, preposition, and articles (Campbell & Pennebaker, 2003). In the current study, LSA can provide a window into the similarity of any two songwriters' linguistic styles as well as content.

Drawing on both deductive and inductive methods, the present investigation addressed the following questions:

- To what degree was there a shift in emotional tone and cognitive analysis over the lifetime of the Beatles—for example, to what degree did the band become more negative and complex?
- How did the Lennon–McCartney songs (which were jointly written) differ from those written solely by Lennon and McCartney as well as George Harrison?
- To what degree are the differences among the writers' lyrics a function of their content or their linguistic styles?

## Method

### *Selection of Songs*

The lyrics of Beatles songs were downloaded from two Web sites ([www.lyricsdownload.com/beatles-lyrics.html](http://www.lyricsdownload.com/beatles-lyrics.html); <http://www.beatles-lyricsarchive.com/>). A random sample of lyrics from songs on both websites was compared and no major discrepancies were found. Only songs written and performed by the Beatles were included in the final data set. Where the song was a Lennon–McCartney composition, the composer of the song—Lennon, McCartney, or Lennon and McCartney—as determined by using a number of available reference materials (e.g., Everett, 1999; Hertsgaard, 1995). Songs composed by Richard Starkey (Ringo Starr) were removed, because there were too few songs to adequately analyze, as were songs that contained fewer than 50 words. The final database of 185 songs included 78 by Lennon, 67 by McCartney, 25 by Harrison, and 15 by the Lennon–McCartney collaborations.

### *Text Analysis Procedure*

The lyrics from each song were converted to conventional American English spelling. Phrases or choruses that were repeated three times or more within a song were deleted—allowing for only a single repetition. For the LIWC analyses, each song was converted to an individual text file. For each song, then, LIWC calculated the percentage of total words in the file that reflected

each of four conceptual linguistic categories: emotional tone, cognitive dynamics, social/identity processes, and time orientation. Twelve of the 18 individual language dimensions associated with these four categories have been used extensively in previous studies (Pennebaker & King, 1999; Pennebaker & Stone, 2003). One additional word count item, sexual words (e.g., sex, breast, love) were added because of its relevance to the current topic. Finally, two of the dimensions, “immediacy” and “making distinctions” are factor-analytically derived from the original Pennebaker and King (1999) work. Texts high in immediacy tend to use short words, present tense, first person singular pronouns, and discrepancy (would, should, could) words. The Making Distinctions category is associated with negations (no, not), exclusive words (except, but), tentative words (perhaps, maybe), discrepancy words, and a low number of inclusive words (and, with).

As noted above, LSA provides a way by which to analyze the similarity of word use from one text sample to another. Recall that the word pattern analyses can focus on both the content of the lyrics (which tend to be based on nouns and regular verbs) and, separately, on the linguistic style of writing (which depends on common function words, including pronouns, prepositions, articles, conjunctions, and auxiliary verbs). In order to establish a basis of comparison between any song lyrics, it is necessary to establish a comparison group of lyrics or, in LSA terms, a semantic space. To do this, approximately 100 of the top-rated songs listed on the yearly United States Billboard Charts for each year between 1962 and 1972 (inclusive) were collected. All songs by the Beatles were excluded, resulting in a total of 1,040 songs. Using these lyrics as a base, two semantic spaces were constructed. The content space, which excluded commonly used words, was made up of 3,120 words and was ultimately reduced to 298 factors, using singular value decomposition. The linguistic style semantic space was comprised of the 588 most commonly used words and resulted in 52 separate factors. As described below, LSA analyses were computed on both style and content dimensions to assess the consistency of each author’s works from song to song and to establish the similarity between pairs of authors.<sup>1</sup>

## Results

Four basic questions were addressed. The first deals with how the Beatles’ lyrics changed over time. The second addressed how the word usage of Lennon, McCartney, Harrison, and Lennon–McCartney differed. The final two questions examined how the degree to which the composers’ lyrics were similar to each other in terms of content and in terms of style.

### Lyrics Over Time

The LIWC analyses for each of the 185 songs were initially broken into three time periods: 1960–1964, 1965–1967, and 1968–1970, and subjected to simple between-subjects one-way analyses of variance (ANOVAs). In the cases where the overall ANOVA was significant, separate linear and quadratic effects were computed. As seen in Table 1, several interesting and important effects emerged.

Partially consistent with popular views, the Beatles’ lyrics tended to become more negative in emotional tone as well as psychologically distant over the course of their career. Indeed, their songs dropped sharply in their sexual and love-related con-

Table 1  
*Linguistic Analyses of the Beatles’ Lyrics, Collapsed Across Composer*

	1960–1964	1965–1967	1968–1970	Effects <sup>a</sup>
Words per song	166.1	164.3	151.2	
Emotional tone				
Positive emotion	6.00	4.79	4.91	
Negative emotion	2.82	1.54	1.92	O, L <sup>#</sup> , Q
Sexual words	2.65	1.47	1.17	O, L
Social/identity				
Social processes	18.34	17.70	15.37	O <sup>#</sup> , L
First person singular	13.60	8.75	7.04	O, L, Q <sup>#</sup>
First person plural	0.48	0.73	0.90	
Cognitive processes				
Cognitive mechanisms	18.17	16.45	15.17	O, L
Words > 6 letters	6.44	7.52	9.17	O, L
Articles	4.11	5.44	5.81	O, L
Making distinctions <sup>b</sup>	0.57	0.12	–0.59	O <sup>#</sup> , L
Time orientation				
Past tense	3.39	3.05	2.73	
Present tense	16.70	15.61	13.77	O <sup>#</sup> , L
Future tense	2.98	2.13	1.48	O, L
Immediacy <sup>b</sup>	1.49	0.09	–1.27	O, L
N	49	73	63	

*Note.* All linguistic categories are expressed as percentage of total words, except words per song, making distinctions, and immediacy.

<sup>a</sup> Effects refer to the significance levels of between-subjects analyses of variance (ANOVAs) that are significant at  $p < .05$ , where O = overall ANOVA (3, 182 *df*), L = linear effect (1, 182 *df*), and Q = quadratic effect (1, 182 *df*). <sup>#</sup> indicates that the reported effect  $p < .09$ .

<sup>b</sup> The categories making distinctions and immediacy are factor scores computed on the weightings reported by Pennebaker & King (1999). Making distinctions is a marker of the degree to which individuals use words that are associated with making distinctions among categories and has been linked to greater cognitive work. Immediacy is a factor associated with use of present tense, simple words, first-person singular, and low usage of articles and suggests the state of living in the present.

tent. The most significant change was the degree to which self-referencing dropped. The degree to which their songs included social referents (e.g., social words) also decreased over time. Drops in social/identity words has frequently been linked to greater psychological distancing and less concern with social relationships (e.g., Pennebaker et al., 2003).

The markers of cognitive processes paint a more nuanced picture. Overall cognitive mechanism words include an array of words that suggest that the author(s) is attempting to understand or think through an object, event, or some feature of personality or relationship. A factor analytically derived dimension, making distinctions, which has been linked to higher level thinking (e.g., the tendency to make distinctions between what is in a category vs.

<sup>1</sup> This pattern comparison problem is akin to trying to determine whether two entities—let’s say marbles—are similar or different. If one has never seen a marble and is given two marbles that differ slightly in weight, coloration, scratches, and so on, one would conclude that the two marbles were completely different. If, however, the judge saw dozens of different shapes and sizes of marbles, he or she might conclude that the two original marbles were moderately similar. Indeed, if the judge were asked to evaluate the similarity of the two marbles in comparison to watermelons, kangaroos, and toasters, he or she would have to admit that the marbles were virtually identical.

what is not), also tended to decline over time. Whereas markers of cognitive processing are dropping over the years, there is a corresponding increase in the Beatles' use of big words and articles. Articles (a, an, the) are used when people are referring to concrete nouns in impersonal ways. Similarly, the use of large words is often a marker of intellectualization and emotional distancing (cf., Cohn, Mehl, & Pennebaker, 2004).

Consistent with this pattern was the degree to which the Beatles' time orientations switched. That is, early in their career, they were very present-oriented, as well as future-oriented. Through the course of their career the Beatles wrote less about the present and the future and the immediacy factor dropped accordingly over time. Immediacy, which is based on the usage of present tense, small words, first-person singular, and low usage of articles, is considered to be a marker of the degree to which people are living in the moment (from Pennebaker & King, 1999; Cohn et al., 2004).

### *Lyric Comparisons of the Composers*

With the exception of a few songs written by Ringo Starr, all of the Beatles' lyrics can be classified as having been written by John Lennon, Paul McCartney, the Lennon–McCartney duo, and George Harrison. Songs by the Lennon–McCartney duo were written in such a way that biographers and the writers themselves agree that both were instrumental in putting lyrics together. It should be noted that a one-way ANOVA on the year of the release of the songs by composer indicated a significant composer effect,  $F(3, 181) = 4.75, p < .01$ , such that Lennon–McCartney jointly written songs were released earlier than Harrison and McCartney songs, but not Lennon songs (mean year: Lennon = 1966.0, McCartney = 1966.7, Lennon/McCartney = 1965.2, Harrison = 1967.4).

As is apparent in Table 2, Lennon and McCartney have distinctive differences in their word usage. When writing together, Lennon–McCartney produced lyrics with a highly positive emotional tone that were written predominately in the present tense. While the lyrics seem simpler, as evidenced by the number of words with less than six letters, the collaborative songs are more concerned with social processes and feature a larger number of sexual words when compared with songs that were written separately.

Songs penned by Lennon are typically higher in negative emotion than McCartney compositions. The Lennon lyrics are higher in cognitive mechanism words, which suggests the author is reflecting on or trying to make sense of events. On the other hand, McCartney lyrics are less concerned with living in the moment and are characterized by a much greater focus on a collective orientation, as reflected in a significantly greater use of words such as "us" and "we." Despite this, the number of words with greater than six letters is higher in McCartney compositions indicating that although the songs may be less intellectual they are more lyrically complex and varied than Lennon songs.

George Harrison lyrics have some characteristics similar to Lennon songs. They are also higher in cognitive mechanism words reflecting an intellectual approach and a search for understanding. These features of his song writing probably reflect Harrison's developing interest in Eastern philosophy and religion.

### *Consistency of Content and Style Within Composer*

For both the content and style semantic spaces, two general analyses were computed. The first simply compared each authors' songs

with each other. For example, John Lennon composed 78 songs on his own. Using LSA, each song was individually compared with his other 77 songs, resulting in 3,003 separate comparisons. For each song comparison, then, a similarity coefficient (similar to a correlation) was computed for both lyric content and style. In the Lennon case, the 3,003 coefficients were averaged revealing a number that depicted how similar his own songs were to each other—with the higher the number reflecting similar content and styles.

As depicted in Table 3, the similarity coefficients for both style and content were compared between the four author sets. Most striking is that the songs by McCartney were most varied (and least redundant) compared to the other solo authors. George Harrison, on the other hand, penned lyrics that overlapped considerably from song to song both in style and content.

### *Comparison of Composers' Content and Style*

LSA also allows us to compare the similarity of any pairs of authors as well. Such analyses can help determine how any two authors shared lyrical approaches. For each pair of authors, each song lyric for one composer was compared with every song lyric by the other composer. So, for example, each of Lennon's 78 songs were sequentially compared with each of McCartney's 67 songs, thus resulting in 5,226 similarity coefficients for both style and content. The average of these coefficients are depicted in Table 4. Two patterns of interest emerge. First, the Lennon–McCartney jointly written songs are more similar to Lennon's linguistic content and style than to McCartney's. In other words, McCartney's contributions as solo author are a greater departure from the Lennon–McCartney lyrics than are Lennon's. The second intriguing pattern is that George Harrison's lyrics were far more similar to Lennon's content and style than to McCartney's.

Overall, the LSA approach provides a different picture than the word count strategy. Paul McCartney's lyrics can be differentiated from Lennon's in being less redundant from song to song and more distant from the Lennon–McCartney lyrics. These effects are apparent for both the actual content of the songs as well as the linguistic styles in which they are written. Finally, it is worthwhile to note that the word use patterns of the lyrics of George Harrison were highly correlated with those of Lennon but not McCartney.

## Discussion

The analysis of the lyrics of the Beatles revealed a number of interesting psychological issues over the career of the band. Songs from the early years of the Beatles were characterized by positive emotion with many tunes concerning the joys of new romantic relationships. Their early records maintained a largely positive emotional tone until the album "Help!" which was released in 1965. The emotional content of songs continued after this period to be more psychologically distant and melancholic and less positive with many songs looking back to happier times. In later years the Beatles writing had become less situated in the present and future than their early lyrics. Along with the emotional changes, their lyrics became more complex and intellectual over time. While early songs were related to personal experiences and feelings, later songs were more often written about other people (e.g., "She's leaving home," "Lady Madonna," and "Get Back"). The orientation of the lyrics also changed from imme-



Table 2  
Word Use Comparison of Lyrics Written by Lennon, McCartney, Lennon–McCartney, and Harrison

	Lennon	McCartney	Lennon–McCartney	Harrison	Effects <sup>a</sup>
Words per song	164.2	156.6	173.6	150.2	
Emotional tone					
Positive emotion	4.83 <sub>b</sub>	4.96 <sub>b</sub>	7.90 <sub>a</sub>	4.96 <sub>b</sub>	
Negative emotion	2.43 <sub>a</sub>	1.49 <sub>b</sub>	2.17 <sub>ab</sub>	1.96 <sub>ab</sub>	
Sexual words	1.43 <sub>a</sub>	1.62 <sub>a</sub>	3.54 <sub>b</sub>	1.52 <sub>a</sub>	O <sup>#</sup>
Social/identity					
Social processes	17.09 <sub>b</sub>	17.11 <sub>b</sub>	22.22 <sub>a</sub>	13.85 <sub>b</sub>	O
First person singular	9.61	8.44	10.96	10.73	
First person plural	0.35 <sub>a</sub>	1.37 <sub>b</sub>	0.18 <sub>b</sub>	0.48 <sub>ab</sub>	O
Cognitive processes					
Cognitive mechanisms	16.16 <sub>a</sub>	15.37 <sub>a</sub>	17.21 <sub>ab</sub>	19.95 <sub>b</sub>	O
Words > 6 letters	7.72 <sub>ab</sub>	8.45 <sub>a</sub>	5.85 <sub>b</sub>	7.64 <sub>ab</sub>	
Making distinctions <sup>b</sup>	0.26 <sub>ab</sub>	-0.58 <sub>a</sub>	-0.87 <sub>a</sub>	1.36 <sub>b</sub>	O
Time orientation					
Past tense	3.11	2.97	3.84	2.45	
Present tense	15.84 <sub>ab</sub>	13.66 <sub>a</sub>	17.55 <sub>b</sub>	16.46 <sub>ab</sub>	O <sup>#</sup>
Future tense	2.19	2.32	1.23	2.02	
Immediacy <sup>b</sup>	-0.33	-0.24	1.20	0.94	
N	78	67	15	25	

Note. All linguistic categories are expressed as percentage of total words, except words per song, making distinctions, and immediacy.  
<sup>a</sup> Effects refer to the significance levels of between-subjects analyses of variance (ANOVAs) that are significant at  $p < .05$ , where O = overall ANOVA (3, 182 *df*), L = linear effect (1, 182 *df*), and Q = quadratic effect (1, 182 *df*). <sup>#</sup> indicates that the reported effect  $p < .09$ . Means with difference subscripts for each variable are significantly different ( $p < .05$ , two-tailed) based on contrasts using the mean-square error term.  
<sup>b</sup> The categories making distinctions and immediacy are factor scores computed on the weightings reported by Pennebaker and King (1999). Making distinctions is a marker of the degree to which individuals use words that are associated with making distinctions among categories and has been linked to greater cognitive work. Immediacy is a factor associated with use of present tense, simple words, first-person singular, and low usage of articles and suggests the state of living in the present.

diate emotional experience to a more emotionally distanced approach that was often trying to understand people or situations.

As the text analyses suggest, John Lennon and Paul McCartney had distinctive lyrical approaches. Lennon, perhaps because his music relied on a more typical blues structure, used more negative emotion in his lyrics and the data reveal that he was more focused on his own personal distress. Lennon songs are indicative of someone trying to reflect or understand his own negative experiences or situations (e.g., “Help!,” “I’m a Loser,” “Don’t Let Me Down,” and “Ballad of John and Yoko”). On the other hand, the data shows McCartney lyrics to be more intellectually and lyrically complex than Lennon’s. McCartney songs also cover a broader range of themes (e.g., “Back in the USSR,” “Paperback Writer,” and “When I’m Sixty Four”) and are more likely to take the viewpoint of others. Possibly in keeping with

his more stable family background, McCartney songs have more references to close collective relationships than Lennon compositions (e.g., “We can Work it Out,” “Two of Us,” “Things we Said,” “Sgt Peppers Lonely Heart Club Band”).

The LSA analyses provided evidence that John Lennon had the most influence over the band members. Songs written jointly by Lennon and McCartney show more linguistic similarities to Lennon’s lyrics than McCartney’s. Most of the jointly written tunes were written early in the band’s career and Lennon’s influence may be due to the fact that he started the band that preceded the Beatles—“The Quarrymen”, was older and more experienced in performing than McCartney. Lennon’s influence was also quite apparent on the songwriting of George Harrison—both in terms of style and substance—in comparison with Paul McCartney. According to biographers of the Beatles, Harrison, the youngest, Beatle always looked up to Lennon from the early days of the band (Hertsgaard, 1995).

Table 3  
Content and Style Consistency by Composer: Latent Semantic Analysis (LSA) Coefficients

	Lennon	McCartney	Lennon–McCartney	Harrison
Content	.119 <sub>a</sub>	.105 <sub>b</sub>	.119 <sub>ab</sub>	.144 <sub>c</sub>
Style	.504 <sub>a</sub>	.480 <sub>b</sub>	.476 <sub>ab</sub>	.575 <sub>c</sub>
N/pair basis	78/3003	67/2211	15/105	25/300

Note. A higher number means the composer’s language usage in any given song is correlated with word usage in his other songs. Within each row, numbers with different subscripts differ at  $p < .05$  (two-tailed *t*-tests). The N/pair basis row refers to the number of songs written by each composer (N) and the number of correlations on which the consistency coefficients were based (pair basis).

Table 4  
Mean Content and Style Similarity Coefficients Across Composers

	Lennon	McCartney	Lennon–McCartney	Harrison
Lennon	—	.108 <sub>v</sub>	.120 <sub>xy</sub>	.127 <sub>z</sub>
McCartney	.487 <sub>a</sub>	—	.101 <sub>w</sub>	.116 <sub>x</sub>
Lennon–McCartney	.488 <sub>a</sub>	.466 <sub>d</sub>	—	.129 <sub>yz</sub>
Harrison	.538 <sub>c</sub>	.509 <sub>b</sub>	.511 <sub>b</sub>	—

Note. Numbers above the diagonal refer to content coefficients; below the diagonal are style coefficients. Coefficients with different subscripts are significantly different at  $p < .05$  (two-tailed tests).

The findings of this study contrast with some of the popular stereotypes of the Beatles. The first is the commonly held view of Lennon as the more intellectual songwriter and McCartney as the sentimental tunesmith. As Everett (1999) notes, "McCartney is seen as the sentimentalist, nonintellectual working-call craftsman who counts his pay in smiles and moves on to the next project, toiling to get every note just right" (p. 10). In fact, the linguistic evidence shows that, while McCartney lyrics have less negative emotional words than Lennon's, McCartney's songs are more intellectually complex and cover a far wider range of perspectives and themes. Lennon's songs tend to more self-focused and higher in levels of negative emotion.

Our linguistic analyses also provide insights into the group dynamics of the Beatles. It is clear that John Lennon had more of an influence on the band's music, particularly in the early stages of the group where his ideas had a greater sway on compositions, and on George Harrison's writing. The changes in the group from a close-knit positive unit to a collection of individuals working independently are reflected in the change of lyrics over time. The early songs are full of immediate positive energy and about close personal relationships while the later material is less positive and immediate and more reflective and complex. This latter finding is consistent with a previous analysis of Beatles songs that showed an increasing novelty and complexity in the group's lyrics and music over time (West & Martindale, 1996).

Beyond the story of the Beatles, the current project suggests a methodology by which to track groups and individuals over time to better understand their psychological dynamics. Psychobiography is certainly not a new endeavor. However, the application of a new generation of text analytic tools is. Computerized word counting strategies have been applied to the analysis of political speeches, novels, poetry, and lyrics (Chung & Pennebaker, 2007; Hart, 2001; Stirman & Pennebaker, 2001; Weintraub, 1989). More recent methods, such as LSA, promise additional ways by which to think of artistic creations. Taken together, we are entering a new world of language analysis that promises to revolutionize the ways we can use people's words to understand their psychological and social states.

## References

- Brog, M. A. (1995). The phenomena of Pine's "four psychologies": Their contrast and interplay as exhibited in the Beatles' "White Album." *American Journal of Psychotherapy, 49*, 385-404.
- Campbell, R. S., & Pennebaker, J. W. (2003). The secret life of pronouns: Flexibility in writing style and physical health. *Psychological Science, 14*, 60-65.
- Chung, C. K., & Pennebaker, J. W. (2007). The psychological functions of function words. In K. Fiedler (Ed.), *Social communication* (pp. 343-359). New York: Psychology Press.
- Cohn, M. A., Mehl, M. R., & Pennebaker, J. W. (2004). Linguistic markers of psychological change surrounding September 11, 2001. *Psychological Science, 15*, 687-693.
- Cook, G., & Mercer, N. (2000). From me to you: Linguistic austerity to profligacy in the songs of the Beatles. In I. Inglis (Ed.), *The Beatles: Popular music and society*. London: MacMillan.
- Everett, W. (1999). *The Beatles as musicians*. New York: Oxford University Press.
- Foltz, P. W. (2003). Quantitative cognitive models of text and discourse processing. In A. C. Graesser, M. A. Gernsbacher, & S. R. Goldman (Eds.), *Handbook of discourse processes* (pp. 487-523). Mahwah, NJ: Erlbaum.
- Hart, R. P. (2001). Redeveloping DICTION: Theoretical considerations. In M. West (Ed.), *Theory, method, and practice in computer content analysis*. New York: Ablex.
- Hertsgaard, M. (1995). *A day in the life: The music and artistry of the Beatles*. London: Macmillan.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review, 104*, 211-240.
- Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). *Linguistic inquiry and word count: LIWC 2007*. Austin, TX: LIWC.
- Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology, 77*, 1296-1312.
- Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology, 54*, 547-577.
- Pennebaker, J. W., & Stone, L. D. (2003). Words of wisdom: Language use over the lifespan. *Journal of Personality and Social Psychology, 85*, 291-301.
- Salewicz, C. (1986). *McCartney*. New York: St Martin's Press: New York.
- Sheff, D., & Golson, G. B. (1982). *The Playboy interviews with John Lennon and Yoko Ono*. New York: Penguin Group.
- Stirman, S. W., & Pennebaker, J. W. (2001). Word use in the poetry of suicidal and nonsuicidal poets. *Psychosomatic Medicine, 63*, 517-522.
- Weintraub W. 1989. *Verbal behavior in everyday life*. New York: Springer
- West, A., & Martindale, C. (1996). Creative trends in the content of Beatles lyrics. *Popular Music and Society, 20*, 1-8.

Received January 29, 2008

Revision received June 16, 2008

Accepted June 16, 2008 ■