

## Research

### \*Corresponding author

Sheila Brownlow, PhD

Professor and Chair

Department of Psychology

Catawba College

Salisbury, NC 28144, USA

E-mail: [sbrownlo@catawba.edu](mailto:sbrownlo@catawba.edu)

Volume 1 : Issue 1

Article Ref. #: 1000SBRPOJ1104

### Article History

Received: July 21<sup>st</sup>, 2016

Accepted: August 12<sup>th</sup>, 2016

Published: August 17<sup>th</sup>, 2016

### Citation

Beach JC, Brownlow S, Greene MD, Silver NC. The "I"s have it: Sex and social status differences on Twitter. *Soc Behav Res Pract Open J*. 2016; 1(1): 17-21. doi: [10.17140/SBRPOJ-1-104](https://doi.org/10.17140/SBRPOJ-1-104)

### Copyright

©2016 Brownlow S. This is an open access article distributed under the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

# The "I"s have it: Sex and Social Status Differences on Twitter

Jennifer C. Beach<sup>1</sup>; Sheila Brownlow, PhD<sup>1\*</sup>; Melissa D. Greene<sup>1</sup>; N. Clayton Silver, PhD<sup>2</sup>

<sup>1</sup>Department of Psychology, Catawba College, Salisbury, NC 28144, USA

<sup>2</sup>Department of Psychology, University of Nevada, Las Vegas, Las Vegas, NV 89154, USA

## ABSTRACT

We examined the use of "I" in Tweets posted by 50 famous people during a one-month window. The linguistic inquiry and word count (LIWC) was implemented to determine the percentage of 140-character Tweets that used the personal pronoun "I". Our findings showed patterns typically seen in natural speech. Specifically, women used the self-referent focus "I" significantly more often than men did, and lower status (operationalized as the number of Twitter followers) people used "I" significantly more often than those with higher status. Men of low social status used significantly more "I" language than did women with lower social status, but women and men of higher social status used "I" equally. Our findings suggest that social status may alter sex-linked communication, with the use of an informal and friendly style that includes self-referencing by men of lower social status, perhaps in an effort to engage more people and enhance their own status through more self-referencing.

**KEYWORDS:** Personal pronoun use; Twitter; Linguistic Inquiry and Word Count (LIWC).

**ABBREVIATIONS:** LIWC: Linguistic Inquiry and Word Count; ANOVA: Analysis of variance; IRB: Institutional Review Board.

## INTRODUCTION

Because people usually focus on content words in conversation and written text, they may miss the significance of function words that more accurately signal people's motivations, personality, and psychological states.<sup>1</sup> For example, Barack Obama's November 12, 2012 Tweet *Four more years*, one of the most shared and re-tweeted posts ever, may tell followers that the re-election was confirmed. However, that Tweet does not have the same meaning as *I won* or *We won*, which would have told followers something different about the person sending the Tweet while conveying the same content.

The LIWC analyzes text and classifies word use according to different parts and types of speech, using a built-in dictionary of approximately 6000 words.<sup>2</sup> The ability of the LIWC program to accurately identify psychological states of language users through their language has been validated in several ways, using many types of writing samples in multiple contexts.<sup>3</sup> The LIWC program categorizes different aspects of word use that tap social processes (e.g., relationships), affect (e.g., positive and negative emotion), cognitive mechanisms (e.g., causality, discrepancy), and linguistic dimensions, such as personal pronouns. It is these linguistic "function" words that actually say more about a speaker's emotional state, cognitive sophistication, and demographic background than do content words.<sup>1</sup>

Pennebaker<sup>1</sup> argued that pronoun use is an important social marker for age, sex, and social status; its use also provides insight into a speaker's psychological state. Although "I" is used more than any other word, comprising 3.64% of our speech<sup>1</sup> its use does not signal narcissism.<sup>4-6</sup> However, the use of "I" is common among those who are anxious, insecure, neurotic,

or self-conscious in a situation.<sup>7,8</sup> In contrast, use of “you” in social media is a marker of people who are open and conscientious.<sup>9</sup> First-person pronouns are more common in the language of people of lower status,<sup>10</sup> and they also decrease judgments of a speaker’s competence.<sup>11</sup>

Enhanced vocal use of “I” is more common in women, because of higher levels of interpersonal focus in comparison to men.<sup>12,13</sup> This sex difference in the use of first-person singular pronouns is seen across a wide variety of contexts.<sup>14</sup> Younger people are more likely to use first-person singular,<sup>12,15</sup> but older people use fewer first-person pronouns, more future tense verbs, and their word and sentence length increases as they age.<sup>16</sup> Moreover, “I” use is prevalent in the dynamic style of speaking that focuses on using narrative or storytelling to convey points, with fewer complex words that show cognitive mechanisms.<sup>17</sup>

Research using methods other than the LIWC as its primary analytical tool, nevertheless, supports the LIWC findings. Specifically, the open vocabulary method used by Eichstaedt and colleagues<sup>18,19</sup> demonstrated a clear relationship between the content of speech (i.e., words used) in social media and personality characteristics,<sup>18</sup> as well as health and well-being.<sup>19</sup> Several of these characteristics differ reliably across age and sex, confirming that women and younger people show openness and the use of articles in their speech. Therefore, content words, as well as function words provide clues as to people’s motivations and psychological states<sup>15</sup> in addition to their age and sex.

Given that much communication is in a short, written form such as texts, Facebook posts, and Snapchat captions, linguistic patterns contained in various types of planned writing may also be seen in online writing. One such online source is Twitter, a social media site that is open-access (i.e., every Tweet still exists and is easily accessed, regardless of whether you are on Twitter), and which allows only 140-characters per Tweet, forcing users to be succinct in their expression. Reuters<sup>20</sup> estimated that 500 million Tweets are sent each day. These communications afford an opportunity to examine self-referencing in a natural, daily setting among people who have a large number of followers. Therefore, our study examined how short language bursts reveal differences in self-reference through personal pronoun use, according to popularity (a measure of social status), age, and sex of actors. Specifically, we predicted that women and actors with a lower social status on Twitter would demonstrate more personal pronoun (“I”) use in their Tweets.

**METHODS**

**Sample**

Tweets from a 30-day period (July 27, 2015 to August 26, 2015) were taken from 50 different famous people. Men and women actors (*n*=25 each) who were ranked on [www.ranker.com](http://www.ranker.com) as the “most popular” in 2015 were targeted for our sample if they had an active (i.e., having tweeted in the past 30 days at least one

time) Twitter account and used Twitter in English. We chose the most popular 25 women actors and 25 most popular men actors from separate ranking lists. Because famous people sometimes have others pretend to be them, we verified the Twitter account by either checking the blue circle on the account to look for a white checkmark (which means that the Twitter corporation verified to whom it belonged), or we went to the target person’s website and located the Twitter account link. The final sample included 2128 Tweets.

**Determining Social Status**

We took a median split of number of followers, our operationalization of social status, for each Twitter user (*Mdn*=1,063,500 followers, *SD*=3,610,545), and their age (*Mdn*=45, *SD*=14.19). The rationale for using a median split (as opposed to dividing the sample into thirds to examine potential quadratic effects) was to avoid potentially discrepant and significantly small within cell sample sizes that could severely diminish power. A *t*-test confirmed that the median split for number of followers yielded two groups (low, high) differing significantly in the number of followers, *t*(48)=4.63, *p*<.0001. However, there were no statistically significant differences in the number of followers as a function of sex, *t*(48)=1.76, *p*>.05, or age *t*(48)=.70, *p*>.05. The final sample of Tweets classified by user sex, age, and social status is seen in Table 1.

	Number of tweets
Sex	
Men	1025
Women	1103
Age	
<45 years	1393
>45 years	735
Social status <i>via</i> followers	
<1,063,500	1053
>1,063,500	1075

**Note:** Age and social status divisions were determined by median split of the sample of Twitter users.

**Table 1:** Tweets by user sex, age, and social status.

**Dependent Measure**

Each person’s Tweet was copied and pasted into a separate word document and entered into the LIWC software, using its standard dictionaries that have been validated by the comparison of judges analysis of written text.<sup>2</sup> We then calculated the percentage of “I” first person personal pronouns in each Tweet.

**RESULTS**

**Preliminary Analysis**

We first determined if there were outliers in the percentage of “I” tweets. Using a criterion of *z*=3.29, *p*<.001, 34 tweets from 16 different people (18 low social status and 16 high social status;

21 men and 13 women) were observed. If one or two persons dominated the outliers, then they would be removed from the analysis. However, because the outliers were somewhat uniform in distribution across people and believing that people may phrase their tweets differently and quickly, we opted to include all tweets given that only 1.5% were outliers. Next, we found that age *via* the median split did not influence “I” use,  $t(2126)=.89$ ,  $p>.05$ , therefore we excluded age from subsequent analysis.

**Sex Differences**

Percentage of “I” language was examined in a  $2 \times 2$  (sex  $\times$  social status) between-subjects analysis of variance (ANOVA). The results produced a statistically significant main effect of sex,  $F(1, 2124)=4.05$ ,  $p<.045$ , with an observed power of .52. As seen in Table 2, women used significantly more “I” language than did men.

**Social Status Differences**

As shown in Table 2, those with lower social status (i.e., fewer followers) used a significantly higher percentage of “I” language compared to their peers with more followers,  $F(1, 2124)=26.14$ ,  $p<.0001$ . The observed power was .99.

**Sex  $\times$  Social Status**

A statistically significant sex  $\times$  social status interaction was obtained,  $F(1, 2124)=3.90$ ,  $p<.049$ , with an observed power of .51. As illustrated in Figure 1, tests of simple effects revealed

that men with few followers used “I” significantly more than did men with many followers  $F(1, 2124)=25.11$ ,  $p<.0001$ . A similar pattern was seen among women. That is, women with fewer followers used significantly more self-focused language than did those with many followers  $F(1, 2124)=4.92$ ,  $p<.028$ . More importantly, men with few followers showed a significantly greater percentage of “I” in their Tweets than did women with few followers,  $F(1, 2124)=7.95$ ,  $p<.006$ , but that among those with many followers (i.e., high social status) there were no statistically significant sex differences in “I” language,  $F(1, 2124)=.01$ ,  $p>.05$ .

**Additional Caveat**

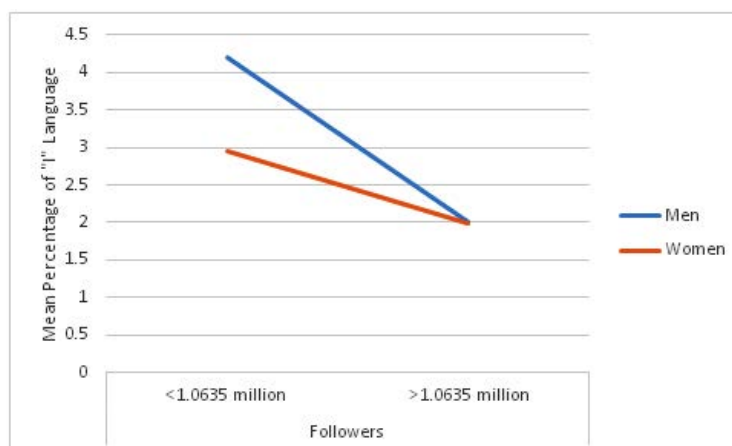
The cell means from the foregoing analysis are shown in Table 2. We note that the standard deviations of the “I” percentage are twice as large as the means, suggesting a lack of normality in the distribution, which was confirmed by a Shapiro-Wilk test ( $W=.565$ ,  $df=2128$ ,  $p<.0001$ ). Although this might result in a loss of power, all hypothesis tests were still statistically significant.

**DISCUSSION**

Our results showed that those of lower social status in our sample of famous Twitter users used “I” more than did those of higher status, and that women did so more than men, but that social status obviated sex differences in “I” language. Men with lower social status used more self-referencing than did women with lower social status, but among high-status (i.e., heavily-followed)

	Sex		Total
	Men	Women	
Followers			
Above Median	2.00 (SD=4.81) [884]	1.99 (SD=4.37) [189]	2.00 (SD=5.49)
Below Median	4.19 (SD=6.60) [141]	2.96 (SD=5.28) [914]	3.12 (SD=4.74)
Total	2.30(SD=5.15)	2.79(SD=5.15)	

The number in brackets represents the number of tweets for the particular cell.  
**Table 2:** Percentage of “I” use as a function of sex and social status (number of followers).



**Figure 1:** Percentage of “I” language according to sex and social status.

Twitter users, no sex differences were seen. It is noted, however, that there was quite a disparity in the number of tweets within each level of sex and social status group. Although our findings that women and low-status people used “I” more is consistent with the literature<sup>1,10,12,14</sup> with regard to pronoun use, the data also suggested that self-referencing among people of high status may reflect a different communication style than the typical linguistic profile. Specifically, only men with lower social ranking used *more than* the typical amount of “I” language 3.64%,<sup>1</sup> and everyone else used less. Whether this outcome was a function of medium (i.e., 140-character bursts), target sample (50 most famous celebrities), or both, is in question. Nonetheless, our data suggest that examination of language patterns might include social status as a variable that may mitigate typical sex- and age-related patterns.

One possible explanation for high status Twitter users writing “I” less often is that its use may be a function of more formal, complete language expressions, and a lack of comfort with the medium, despite the social popularity (*via* followers) of the Twitter user. Twitter users with a higher social status may have used their following as a platform to Tweet about various causes, thus removing themselves from their Tweets. A second explanation is that higher-status users with many followers may have been more personally (or media) secure.<sup>7</sup> Alternately, the use of “I” is more common in a storytelling or dynamic communication style that uses narrative, rather than linguistic devices with complex cognitive and causal mechanisms.<sup>17</sup> Therefore, lower social status may have led to users to affect a chattier and informal style designed to garner followers. The downside, however, is that such personable language may have reinforced the Twitter user’s lack of social status and called his or her competence into question.<sup>10,11</sup>

Although women used “I” more than did men, our results showed no meaningful differences in “I” use between men and women actors with many followers, a finding that is in contrast to documented literature using multiple samples indicating that women use the first person-pronoun more often than do men.<sup>1,3</sup> Although many LIWC studies use populations not limited to college-aged students,<sup>8,18</sup> most include a preponderance of persons in the college-age range whereas our sample did not. This suggests that social status, as it manifests itself in a spontaneous medium, may be more essential to expression than sex. In regard to age, differentiations between “older” and “younger” was a function of the sample and, on a more practical level, somewhat artificial, as our Twitter users (ranging from 22 to 84-years old) were generally middle-aged. Thus, it is not surprising that we found no evidence of age-related differences in “I” use.

Although our findings shed light on the use of first-person pronouns in quick language bursts, the sample we used does not reflect the typical sample of Twitter users. We operationalized social status according to the number of Twitter followers, which in our sample ranged from around 25,000 to over two

million, whereas the average number of Twitter followers for any given user is 208.<sup>21</sup> Moreover, Twitter posts may change in mood over the course of a day,<sup>22</sup> becoming less positive (and probably less “I” oriented) as the day goes on. Finally, it is possible that although the accounts belonged to the persons in question, the Tweets themselves may have been written by publicists posting on social media for the famous persons.

A second limitation of our study is that the Twitter users were primarily American, and mostly White. Thus, our findings are generated from a culture where the use of first-person pronouns and self-focus predominate, in contrast to other more collectivist cultures. It is important, therefore, to examine similar sex and status-related language differences in people other cultures and among a sample that is more inclusive of persons of different races and ethnicities.

In sum, this study of “I” use in Twitter posts by popular actors revealed that language use on social media mirrors some aspects of natural language use, but in other aspects it may be different from that which is written or spoken. Moreover, Twitter language may be different from language on social media that affords longer thoughts to be posted, such as Facebook. Further research might include other measures of social status outside of followers to verify, whether other measures of status qualify “I” use patterns that are a function of age and sex. Our results do, however, contribute to the body of research that provides insight as to how public utterances may reveal private selves and motivations.

#### ACKNOWLEDGEMENTS

The authors are grateful to Justin Adkins and Erich Guebert for their assistance with data collection and study suggestions.

#### NOTES

Authors did not add median split of age as a two-level factor in the design along with social status and sex because there were no Tweets by older women with many followers and only 25 tweets from older men with few followers.

#### CONFLICTS OF INTEREST

The authors have no disclosure or competing interests and this research was unfunded.

#### CONSENT

There is no consent required for research with public documents, and research used public-domain material (Twitter), which does not require Institutional Review Board (IRB) approval.

#### REFERENCES

1. Pennebaker J. *The Secret Life of Pronouns: What Our*

- Words Say About Us. New York, NY, USA: Bloomsbury Press/ Bloomsbury Publishing; 2011.
2. Pennebaker JW, Chung CK, Ireland M, Gonzales A, Booth RJ. The development and psychometric properties of LIWC2007. *LIWC*. Web site. <http://www.liwc.net/LIWC2007LanguageManual.pdf>. Accessed July 20, 2016
  3. Tausczik Y, Pennebaker J. The psychological meaning of words: LIWC and computerized text analysis methods. *J Lang Soc Psychol*. 2010; 29(1): 24-54. doi: [10.1177/0261927X09351676](https://doi.org/10.1177/0261927X09351676)
  4. Carey A, Brucks M, Mehl M, et al. Narcissism and the use of personal pronouns revisited. *J Pers Soc Psychol*. 2015; 109(3): e1-e15. doi: [10.1037/pspp0000029](https://doi.org/10.1037/pspp0000029)
  5. Golbeck J. Negativity and anti-social attention seeking among narcissists on Twitter: A linguistic analysis. *First Monday*. 2016; 21(3): 7. doi: [10.5210/fm.v0i0.6017](https://doi.org/10.5210/fm.v0i0.6017)
  6. DeWall C, Buffardi L, Bonser I, Campbell W. Narcissism and implicit attention seeking: Evidence from linguistic analyses of social networking and online presentation. *Pers Individ Dif*. 2011; 51(1): 57-62. Web site. <http://iscience.deusto.es/2011/03/28/dewall-c-n-buffardi-l-e-bonser-i-campbell-w-k-in-press-narcissism-and-implicit-attention-seeking-evidence-from-linguistic-analyses-of-social-networking-and-online-presentation-per/>. Accessed July 20, 2016
  7. Hirsh J, Peterson J. Personality and language use in self-narratives. *J Res Pers*. 2009; 43(3): 524-527. Web site. [http://individual.utoronto.ca/jacobhirsh/publications/Hirsh\\_Peterson\\_2009\\_JRP.pdf](http://individual.utoronto.ca/jacobhirsh/publications/Hirsh_Peterson_2009_JRP.pdf). Accessed July 20, 2016
  8. Pennebaker J, King L. Linguistic styles: Language use as an individual difference. *J Pers Soc Psychol*. 1999; 77(6): 1296-1312. doi: [10.1037/0022-3514.77.6.1296](https://doi.org/10.1037/0022-3514.77.6.1296)
  9. Golbeck J, Robles J, Edmondson M, Turner K. Predicting personality from Twitter. *2011 IEEE international conference on privacy, security, risk, and trust, and IEEE International conference on social computing*. 2011; 149-156. Web site. <https://www.semanticscholar.org/paper/Predicting-Personality-from-Twitter-Golbeck-Robles/8a6dca55d8fabd15bec20eb25a5614172de164a7/pdf>. Accessed July 20, 2016
  10. Kacewicz E, Pennebaker J, Davis M, Jeon M, Graesser A. Pronoun use reflects standings in social hierarchies. *J Lang Soc Psychol*. 2014; 33(2): 125-143. doi: [10.1177/0261927X13502654](https://doi.org/10.1177/0261927X13502654)
  11. Berry D, Pennebaker J, Mueller J, Hiller W. Linguistic bases of social perception. *Pers Soc Psychol Bull*. 1997; 23(5): 526-537. doi: [10.1177/0146167297235008](https://doi.org/10.1177/0146167297235008)
  12. Pennebaker J. How author's forgettable words reveal their personality and social behavior. *Information Design Journal*. 2013; 20(3): 267-272. Web site. <http://web.b.ebscohost.com/abstract?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=1876486X&AN=97429540&h=eV1PY1uMYKHfyGUsNf384IHC1oA%2fUw3sRtJRPb6CB2BXx3t7fOVuyQ6AFbhFdOXWvkJTSwJkGTQEmYSPEJdo0Q%3d%3d&crl=f&resultNs=AdminWebAuth&resultLocal=ErrCrlNotAuth&crlhashurl=login.aspx%3fdirect%3dtrue%26profile%3dehost%26scope%3dsite%26authtype%3dcrawler%26jrnl%3d1876486X%26AN%3d97429540>. Accessed July 20, 2016
  13. Poole M. Social class, sex, and linguistic coding. *Lang Speech*. 1979; 22(1): 49-67. doi: [10.1177/002383097902200104](https://doi.org/10.1177/002383097902200104)
  14. Newman ML, Groom CJ, Handelman LD, Pennebaker JW. Gender differences in language use: An analysis of 14,000 text samples. *Discourse Processes*. 2008; 45(3): 211-236. doi: [10.1080/01638530802073712](https://doi.org/10.1080/01638530802073712)
  15. Kern M, Eichstaedt J, Seligman M, et al. From 'Sooo excited!!!' to 'So proud': Using language to study development. *Dev Psychol*. 2014; 50(1): 178-188. doi: [10.1037/a0035048](https://doi.org/10.1037/a0035048)
  16. Pennebaker J, Stone L. Words of wisdom: Language use over the life span. *J Pers Soc Psychol*. 2003; 85(2): 291-301. doi: [10.1037/0022-3514.85.2.291](https://doi.org/10.1037/0022-3514.85.2.291)
  17. Pennebaker J, Chung C, Frazee J, Lavergne G, Beaver D. When small words foretell academic success: The case of college admissions essays. *PLoS ONE*. 2014; 9(12). doi: [10.1371/journal.pone.0115844](https://doi.org/10.1371/journal.pone.0115844)
  18. Schwartz HA, Eichstaedt JC, Kern ML, et al. Personality, gender, and age in the language of social media: The open vocabulary approach. *PLoS ONE*. 2013; 8(9). doi: [10.1371/journal.pone.0073791](https://doi.org/10.1371/journal.pone.0073791)
  19. Park G, Schwartz H, Seligman M, et al. Automatic personality assessment through social media language. *J Pers Soc Psychol*. 2015; 108(6): 934-952. doi: [10.1037/pspp0000020](https://doi.org/10.1037/pspp0000020)
  20. Reuters. Twitter incorporated company profile. 2016. Web site. <http://in.reuters.com/finance/stocks/companyProfile?symbol=TWTR.N>. Accessed July 20, 2016
  21. Muruganandam C. Average number of Twitter followers. 2016. Web site. <http://yourescapefrom9to5.com/average-number-of-twitter-followers-is-208-infographic>. Accessed July 20, 2016
  22. Golder S, Macy M. Diurnal and season mood vary with work, sleep, and daylength across diverse cultures. *Science*. 2011; 333(6051): 1878-1881. doi: [10.1126/science.1202775](https://doi.org/10.1126/science.1202775)