8-10-2016

Developing a LIWC Dictionary: The lyrical poetry of 2PAC, Frank Waln, Litefoot, and Nataanii Means

Teresa Y. Neely
Sarah Kostelecky
Monica Etsitty Dorame
Jorge Ricardo López-McKnight

Follow this and additional works at: http://digitalrepository.unm.edu/ulls_fsp

Recommended Citation

This Presentation is brought to you for free and open access by the Scholarly Communication - Departments at UNM Digital Repository. It has been accepted for inclusion in University Libraries & Learning Sciences Faculty Publications by an authorized administrator of UNM Digital Repository. For more information, please contact disc@unm.edu.
NDLC’16
National Diversity in Libraries Conference
August 10-13, 2016, UCLA Libraries, Los Angeles, California
FROM “MY BLOCK” TO “BORN ON THE REZ”

A linguistic analysis of lyrics by 2PAC, Frank Waln, Litefoot, Nataanii Means, and Night Shield
DEVELOPING A LIWC DICTIONARY

The lyrical poetry of 2PAC, Frank Waln, Litefoot, and Nataanii Means
RESEARCHERS @
UNIVERSITY OF NEW MEXICO

Teresa Y. Neely, Associate Professor, College of the University Libraries & Learning Sciences, neely@unm.edu
Sarah Kostelecky (Zuni Pueblo), Assistant Professor, College of the University Libraries & Learning Sciences, sarahrk@unm.edu
Monica Etsitty Dorame (Navajo Nation), Library Operations Manager, College of the University Libraries & Learning Sciences, mdorame@unm.edu
Jorge Ricardo López-McKnight, Assistant Professor, College of the University Libraries & Learning Sciences, jrlm@unm.edu
ABOUT THIS PROJECT - NOW

- **Phase 1**
  - Identify Artists and Corpus Criteria
  - Categorize each song in corpus in one of the following Lyric Categories
    - My Block (2Pac) and Born on the Rez (FW)
    - Thug Life (Out on Bail; Is there Heaven for a G?; When We Ride) (2Pac) and Rez Affiliated (LF)
  - Present Findings in March 2016 and August 2016 (if proposals are accepted)

- **Phase 2**
  - Complete dataset with (8) Night Shield albums, (1) Litefoot album, and the forthcoming (1) Frank Waln album.
  - Present findings from remaining categories at National Diversity in Libraries Conference in Los Angeles, CA, August 10-13, 2016
    - Mama’s Just a Little Girl (2Pac) and My Ancestors (LF)
    - Me and My Girlfriend (Thug Passion; and Skandalouz) (2Pac) and A Song for Her (NM)

- **NEW Phase 2**
  - Develop a Linguistic Inquiry and Word Count (LIWC) Dictionary for rap and hip hop lyrics, Tupac Shakur-Litefoot-Nataanii Means-Frank Waln-Speak, Native American rappers, African American rappers, indigenous peoples, and young men in general

- **Phase 3**
  - Prepare and submit all research findings on all categories for publication in a peer-reviewed journal
WHY ARE WE DOING THIS?

Mayor Richard J. Berry: Albuquerque Is Dangerous for Homeless Natives

Natives living on Albuquerque streets experience violence at a higher rate than non-Natives. According to a 2011 study published by the University of New Mexico, 71 percent of homeless Natives in Albuquerque reported experiencing violence.

<table>
<thead>
<tr>
<th>PEOPLE KILLED IN 2016</th>
<th>RACE &amp; ETHNICITY PER MILLION</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>654</td>
<td>Native American 5.49</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Black 4.16</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hispanic/Latino 1.89</td>
<td></td>
</tr>
<tr>
<td></td>
<td>White 1.63</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asian/Pacific Islander 0.56</td>
<td></td>
</tr>
</tbody>
</table>

NEW MEXICO
Total killed: 17
Population: 2,085,109
Rank (per capita): 1 / 51
Rank (total): 11 / 51
Do Native American male rappers address socioeconomic issues in their lyrics?

How does the lyrical content of Native American male rappers compare Tupac Shakur’s lyrical content written before 1996?
THE ARTISTS

- Tupac Shakur
  - "Above all else, Tupac was an artist who represented a generation and a culture” (Hall, 2015)
    White Man’z World, Don Killuminati: 7 Day Theory, 1996

- Nataanii Means
  - “...For my generation, I can reach them through my music, through this hip hop, tell them of our people, because they never taught us in school about any of that.” (PowWows.com, 2014)
    Genocide, 2 Worlds, 2013

- Litefoot
  - “Litefoot, the Cherokee rapper who grew up in Oklahoma... is perhaps the best known Native hip-hop artist working today... Litefoot is on a mission. He sees his music as a tool for fighting ignorance about Native culture rather than a way to get rich.” (Chicago Tribune, 2004)
    Redvolution, Redvolution, 2003

- Frank Waln
  - “I have a platform now, and I have the responsibility to use it to try to help solve some of these problems. If we as Natives don’t talk about this, who will?” (Walker, 2014)
    AbOriginal, 2013
Text Analysis
- LIWC (Linguistic Inquiry and Word Count) 2015 – $$ - Computerized text analysis that uses everyday words to reveal “thoughts, feelings, personality, and motivations.”
- Categories: Function Words, Biological Processes, Time Orientation, Affect, Social, Personal Concerns, Cognitive Processes, Informal Language

Visualization
- Wordle - FREE- generates word clouds from user provided text
- Voyant Tools – FREE – visualizes text, # of words, word frequency

Software not chosen
- NVivo-$$, popular, steep learning curve
- ATLAS-$$, popular, steep learning curve
- Docuburst – FREE, user friendly, required to share uploaded data
LIWC DICTIONARY

- LIWC 2015 has 3 dictionaries included - LIWC2015, LIWC2007, LIWC2001
  - Has 125 categories including dimensions and an undetermined number of words.

- Develop Your Own dictionary and upload it to the LIWC site for sharing with others.
  - Political Issues, Turkish, Regulatory, Geography, Football Clubs, Security, Hashtags from 2016 Presidential Elections

- Review of the Literature
  - Developing simplified Chinese psychological linguistic analysis dictionary for microblog (2013)
  - Privacy dictionary: A linguistic taxonomy of privacy for content analysis (2011)
  - Developing a Tagalog LIWC ‘disaster’ dictionary for understanding mixed language social media (2014)
  - Development and evaluation of Tagalog LIWC dictionaries for negative and positive emotion (2014)
  - The Romanian version of the LIWC2001 dictionary and it's application for text analysis with Yoshikoder (2012)
  - Construction and initial validation of a dictionary for global citizen linguistic markers (2014)
Like a rolling stone: A mixed-methods approach to linguistic analysis of Bob Dylan’s lyrics (2016)

We will rock you: A diachronic corpus-based analysis of linguistic features in rock lyrics (2016)

Word usage and thematic content of song lyric analyses (2016)

Gone country: An investigation of Billboard country songs of the year across social and economic conditions in the United States (2015)

Lyrics, music, and emotions (2012)

Lyrics of national anthems and suicide rates (2011)

Tuning in to psychological change: Linguistic markers of psychological traits and emotions over time in popular U.S. song lyrics (2011)

Psychological effects of music with lyrics: A methodological study (2011)

Touched by music: Attachment-related differences in music-induced crying and lyrics (2009)

The language of lyrics: An analysis of popular Billboard songs across conditions of social and economic threat (2009)

Things we said today: A linguistic analysis of the Beatles (2008)

Using computational text analysis tools to compare the lyrics of suicidal and non-suicidal songwriters (2007)
LIWC Analysis

- Percentage of Total Words within text
- Summary Variables
  - **Analytical Thinking** – based on 8 function word dimensions – captures the degree to which people use words that suggest formal, logical, and hierarchical thinking patterns. Low analytical thinking – use language that is more narrative, focusing on the here-and-now, and personal experiences; higher analytical thinking – perform better in college and have higher college board scores.
  - **Clout** – refers to the relative social status, confidence, or leadership that people display through their writing or talking.
  - **Authenticity** – when people reveal themselves in an authentic or honest way, they are more personal, humble, and vulnerable.
  - **Emotional Tone** – includes both positive and negative emotion dimensions both in a single summary variable. The higher the number, the more positive the tone. Numbers below 50 suggest a more negative emotional tone.
WHY LIWC DIDN’T WORK FOR US

- Inadequate dictionary coverage
- Missing categories
- LIWC combined categories
- Places (California), proper names (Geronimo/Geronimo Pratt), ethnic groups (Passamaquoddy), languages (Sicangu Che), regional colloquialisms (M.O.B./off the chain), singular (Playa Haters) words vs. phrases (PlayaHaters), broad categories (social, verb, negative emotion, positive emotion), multiple categories for the same words
Born On The Rez, Color Coded LIWC Dictionary

born on the Rez
raised on the Rez
born on the Rez hoka
raised on the Rez

braids on my head hoka
hot days on the Rez
my heart stays on the Rez
ugh
my Lakota che
my Sicangu che
my people
my Lakota che
my Sicangu che

buckle up
I'll take you to a place you might've been
call it Rez
I call it evil
cause it's killing all my friends
looking back
is probably where it begins
the government kills us
for the land that you live in
White people win
Indian lose
Based on LIWC2015

Initial intent – corpus-based and to complement the LIWC

Began with the Color Code Analysis – highlights words in the LIWC Dictionary revealing the words that do not appear

Development – twofold
  - Find a word in LIWC2015 to CHANGE the Word to,
    - Rap corpus- tracks; rhymes LIWC2015-Rap (leisure category)
  - OR, Create a New Category
    - Indigenous
    - Appropriation
    - Homicide
DKL-MN2016 CATEGORIES

- Categories were created based on the way the words were used in the lyrics.
  - Categories were not pre-determined
- Has 103 categories and more than 2700 words.
- 3 combined categories (not dimensions)
  - 70WhiteManzLand – 1, 6, 9, 11, 12, 14, 15, 25, 27, 28, 30, 32, 33, 47, 54, 59, 64, 77, 86, 88, 91, 92, 94, 95, 98
  - 71Disruption – 3, 31, 40, 52, 53, 68, 73, 76, 90, 93
  - 72Casualties – 17, 49, 50, 83, 86, 94, 97, 100, 102
Intro:
Ahh, Now, everybody ready? Everybody ready?
Alright, alright, alright. Check this. Here we go.
To the Tribes, Reservations, all my Nations, check, check

Chorus:
Where my Natives with the bomb at? Heya, Heya
Who them Natives love these bomb tracks? Oh-wa, Oh-wa
You love hanging with these warriors? Hell yeah
Do these goddesses ignore you? Hell no
Where you at got the fat sounds? Heya, Heya
Where your Rez with the bomb town? Oh-wa, Oh-wa
You love them dimes with the bomb thighs? Hell yeah
You betray us for them land guys? Hell no

Verse 1:
I got a pride that burns me in me, Only Natives and Indians ever really heard of me
Burning me, surgery, what I got to do thoroughly, Who the next Indian turn on me
Big ups to those who recognize, Chin up turn your gaze to the skies
Gone bring it back, raps tracks, facts cracks, rap going stacks
Never give the new millennium, A lot of cat’s man I ain’t feeling ‘em
Same old thang, same old game, Talking loud but how many discs do you sing?
Shut ’em up quick, get the ducks sick, shoot the dice trick, Flip the wrist with the hit, damn
Got bucked, stuck, stupid, screaming mad as fuck, what?
Intro:
Ahh, Now, everybody ready? Everybody ready?
Alright, alright, alright. Check this. Here we go.
To the Tribes [20], Reservations [19], all my Nations [20], check, check

Chorus:
Where my Natives [19] with the bomb [34] at? Heya [21], Heya [21]
Who them Natives [19] love [73] these bomb [34] tracks [56]? Ohwa [21], Ohwa [21]
You love [73] hanging with these warriors [35, 36]? Hell [65, 74] yeah
Do these goddesses [35] ignore you? Hell [65, 74] no
Where you at got the fatsounds [56]? Heya [21], Heya [21]
Where your Rez [19] with the bombtown [51]? Ohwa [21], Ohwa [21]
You love them dimes [39] with the bombthighs [35]? Hell [65, 74] yeah
You betray [25, 70] us for them landguys [1, 70]? Hell [65, 74] no

Verse 1:
I got a pride [53, 71] that burns me in me, Only Natives [19] and Indians [19] ever really heard of me
Burning me, surgery, what I got to do thoroughly, Who the next Indian [19] turn on me
Bigups [54, 72] to those who recognize, Chinup [53, 68, 71] turn your gazetotheskies [53, 68, 71]
Gone bring it back, rap [56] tracks [56], facts cracks, rap [56] going stacks [56]
Never give the newmillennium [79], A lot of cats [16, 29], man [38] I ain’t feeling them
Same old thang, same old game, Talking loud but how many discs [56], do you sing [56]?
Shut them up quick, get the ducks [29] sick [100, 72], shoot [1, 70] the dicetrick [18], Flipthewristwiththehit [18], damn [65]
Got bucked, stuck, stupid, screamingmad [91, 70], as fuck [65, 80], what?
ABOUT THE CORPUS
**ORIGINAL DATASET SUMMARY**

- **4 Artists:** 2Pac, Frank Waln, Litefoot, Nataanii Means
- **Albums:** 27
- **Total Tracks:** 421
- **Unique Tracks:** 315

### Phase 1 Project Summary

<table>
<thead>
<tr>
<th></th>
<th>2Pac</th>
<th>%</th>
<th>Litefoot</th>
<th>%</th>
<th>Nataanii Means</th>
<th>%</th>
<th>Frank Waln</th>
<th>%</th>
<th>Night Shield</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albums</td>
<td>19</td>
<td>76.23%</td>
<td>7</td>
<td>23.77%</td>
<td>1</td>
<td>100%</td>
<td>0</td>
<td>0</td>
<td></td>
<td>27</td>
</tr>
<tr>
<td>Total Tracks</td>
<td>326</td>
<td>76.23%</td>
<td>76</td>
<td>23.77%</td>
<td>13</td>
<td>100%</td>
<td>6</td>
<td>100%</td>
<td></td>
<td>421</td>
</tr>
<tr>
<td>Unique Tracks</td>
<td>220</td>
<td>67.48%</td>
<td>76</td>
<td>100%</td>
<td>13</td>
<td>100%</td>
<td>6</td>
<td>100%</td>
<td></td>
<td>315</td>
</tr>
<tr>
<td>My Block</td>
<td>54</td>
<td>24.55%</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td>54</td>
</tr>
<tr>
<td>Born on the Rez</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Thug Life</td>
<td>105</td>
<td>47.27%</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td>105</td>
</tr>
<tr>
<td>Out on Bail</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Is There Heaven for a G?</td>
<td>13</td>
<td>5.90%</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td>13</td>
</tr>
<tr>
<td>When We Ride</td>
<td>11</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>Rez Affiliated</td>
<td>7</td>
<td>30.43%</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>53.84%</td>
<td>0</td>
<td>15.66%</td>
<td></td>
<td>56</td>
</tr>
<tr>
<td>Mama’s Just a Little Girl</td>
<td>4</td>
<td>1.81%</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>My Ancestors</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>17.10%</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>15</td>
</tr>
<tr>
<td>Me and My Girlfriend</td>
<td>8</td>
<td>3.63%</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>33.33%</td>
<td>8</td>
</tr>
<tr>
<td>Thug Passion</td>
<td>15</td>
<td>61.54%</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>Skandalous</td>
<td>9</td>
<td>4.05%</td>
<td>1</td>
<td>31.58%</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>A Song For Her</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>92.10%</td>
<td>1</td>
<td>7.90%</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
</tr>
</tbody>
</table>
Tupac Shakur Categories

- My Block (Better Dayz disc 2)
  - 54 titles
- Thug Life
  - When We Ride (All Eyez on Me disc 2)
  - Out on Bail (Loyal to the Game)
  - Is There Heaven for a G?
- Mama's Just a Little Girl (Better Dayz disc 1)
- Me and My Girlfriend
  - Thug Passion (All Eyez on Me disc 2)
  - Skandalouz (All Eyez On Me disc 1)
  - Me and My Girlfriend (the don killuminati)

Native American Categories

- Born on the Rez
  - (FrankWaln)
- Rez Affiliated
  - (Litefoot.Good Day to Die)
- My Ancestors
  - (Litefoot.Good Day to Die)
- A Song For Her
  - (NataaniiMeans.2Worlds)
RESEARCH QUESTIONS

Do Native American male rappers address socioeconomic issues in their lyrics?

How does the lyrical content of Native American male rappers compare Tupac Shakur’s lyrical content written before 1996?
DO NATIVE AMERICAN MALE RAPPERS ADDRESS SOCIOECONOMIC ISSUES IN THEIR LYRICS? YES!

- Total word count=41,067
- #1 ColonizedAggression - 0.93% of total (381 words)
  - genocidal, violence, war, treaties, warpath, poverty
- #25 Oppression - 0.86% (353 words)
  - Government, BoardingSchool, disease, CIB, BlackCharredTipis
- #51 Home - 0.94% (386 words)
  - Reservation, Plains, Earth, Dirtstreets, communities
- #64 Structural Racism - 1.34% (550 words)
  - CIB, BoardingSchool, Bible, poverty
HOW DOES THE LYRICAL CONTENT OF NATIVE AMERICAN MALE RAPPERS COMPARE TO TUPAC SHAKUR’S LYRICAL CONTENT WRITTEN BEFORE 1996?

- Total word count=165,433
  - #1 ColonizedAggression- 1.47% of total (2,431 words)
    - Gat, LatashaHarlins, DressedinBlueSlacks
  - #25 Oppression- 0.41% (678 words)
    - Slavery, BanAllRapMusic, AmeriKKKa
  - #51 Home- 0.61% (1,009 words)
    - Warzone, hood, block, concrete
  - #64 Structural Racism- 1.24% (2,051 words)
    - PenitentiaryTimes, PackedJails, Deadbroke

- Total word count=41,067
  - #1 ColonizedAggression- 0.93% of total (381 words)
    - genocidal, violence, war, treaties, warpath
  - #25 Oppression- 0.86% (353 words)
    - Government, disease, CIB, BlackCharredTipis
  - #51 Home- 0.94% (386 words)
    - Reservation, Plains, Earth, Dirtstreets, communities
  - #64 Structural Racism- 1.34% (550 words)
    - CIB, BoardingSchool, Bible
FUTURE DIRECTIONS

Pettijohn and Sacco, JR., “The language of lyrics: An Analysis of Popular Billboard Songs across conditions of social and economic threat,” *Journal of Language and Social Psychology* 28 (3) (2009): 297-311. **Findings**: Songs with more words per sentence, a focus on the future, and greater mention of social processes and intergroup themes were popular during threatening social and economic conditions.


Powell-Morse, “Lyric intelligence in popular music: A Ten Year Analysis,” Data Visualizations from SeatSmart,” (2005). **USED**: Readability Score- It uses writing analysis tools like the Flesch–Kincaid grade index and many others to create an average of the US reading level of a piece of text. Analyzed 225 songs in 4 different datasets, resulting in 2,000+ individual data points. **FINDINGS**: The average grade level for top ranking songs appearing on *Billboard* is the 3rd grade. And the bar is getting lower. [http://seatsmart.com/blog/lyric-intelligence/](http://seatsmart.com/blog/lyric-intelligence/).


Pettijohn and Sacco, JR., “The language of lyrics: An Analysis of Popular Billboard Songs across conditions of social and economic threat,” *Journal of Language and Social Psychology* 28 (3) (2009): 297-311. **Findings:** Songs with more words per sentence, a focus on the future, and greater mention of social processes and intergroup themes were popular during threatening social and economic conditions.

Kakde, Ghanekar, and Sindh, “‘Life’s a Song! Mining Country Music Topics Using SAS® Text Miner,’” Conference Presentation. SAS Global **USED:** SAS® Text Miner--Text Import, Text Parsing, Text Filter, Text Topic nodes; and the **FINDINGS:** Female artists are twice as likely to sing about heartbreak as compared to male artists; Percentage of songs about Country Life, Love & happiness and Childhood are almost identical for both female and male singers. http://support.sas.com/resources/papers/proceedings13/268P-2013.pdf.

Powell-Morse, “‘Lyric intelligence in popular music: A Ten Year Analysis,’” *Data Visualiziations from SeatSmart,“* (2005). **USED:** Readibility Score- It uses writing analysis tools like the Flesch-Kincaid grade index and many others to create an average of the US reading level of a piece of text. Analyzed 225 songs in 4 different datasets, resulting in 2,000+ individual data points. **FINDINGS:** The average grade level for top ranking songs appearing on Billboard is the 3rd grade. And the bar is getting lower. http://seatsmart.com/blog/lyric-intelligence/.

“Bob Dylan—analyzing 55 years of lyrics,” via Tableau Picasso (a blog about data visualization with Tableau). **RESULTS:** An interactive resource with lexical density on the y-axis and lyrical density on the x-axis. You can examine the lexical density of different albums by selecting the color in a legend. http://www.everydayanalytics.ca/2013/06/radiohead-lyrics-data-visualization-and-content-analysis.html

QUESTIONS??

- Teresa Y. Neely, Associate Professor, College of the University Libraries & Learning Sciences, neely@unm.edu
- Sarah Kostelecky (Zuni Pueblo), Assistant Professor, College of the University Libraries & Learning Sciences, sarahrk@unm.edu
- Monica Etsitty Dorame (Navajo Nation), Library Services Coordinator, College of the University Libraries & Learning Sciences, mdorame@unm.edu
- Jorge Ricardo López-McKnight, Assistant Professor, College of the University Libraries & Learning Sciences, jrlm@unm.edu
THANK YOU!