Text Mining and Digital Humanities: Quantitative Analysis of African American Poetry

Taylohr Brown, Diamond Jenkins, Walter Quiller, Researchers. Seretha Williams, BSJ, MA, PhD, Faculty Advisor.

CURS Summer Scholars Program, Georgia Regents University, Augusta, Georgia

GRU GEORGIA REGENTS UNIVERSITY

INTRODUCTION

"Text Mining and Digital Humanities: Quantitative Analysis of African American Poetry" uses quantitative and qualitative analysis to formulate research questions about African American poetry. In this project, we use text-mining software to determine whether distinctive word patterns can be used to quantify the characteristics of African American poetry. For the purposes of this study, we rejected the notion that Black poetry is defined as poetry written by black authors. Instead, we argue, the distinctions in black poems should be specific enough to be classified in a separate category from other kinds of literature such as American literature, and we assert the definition of black poetry should not reduce "blackness"- what we describe as the shared cultural traditions or practices of African Americans- to certain experiences or tropes such as the rural, folk black experience.

We selected Langston Hughes and the Harlem Renaissance as the earliest historical point for our inquiry, and we used Margaret Walker, Gwendolyn Brooks, Maya Angelou, and Alice Walker, poets whom Hughes directly influenced, as comparisons. We created a text database of the collected poems of the five authors and assessed the frequency of words/phrases related to three main categories that recur in the scholarship of black poetry: memory, identity, and music.

METHODOLOGY

First, we developed a question that text mining could answer: Are there distinctive characteristics of our poets' works that can constitute the unique category of black poetry? Next, we chose our position: Black poetry can be uniquely categorized as its own literary genre. We developed three variables to test this hypothesis:

1) Memory- This category attests to the importance of remembering and documenting struggle in black art. Slavery, Jim Crow segregation, and the sociopolitical and economic marginalization of black Americans are some of the experiences that recur in black poetry.
2) Identity- The words in this list relate to black people's efforts to define themselves and to establish an accurate individual and collective cultural identity free from the overbearing influence of whites.
3) Music- Early African-American music reveals the search for black identity. Our poets recognize that music operates as a vital element of black art on the continent and diaspora. They include music and musical forms and acknowledge the performative aesthetic of black poetry.

Lastly, we collected all the works of our poets (a total of 94,527 words), digitizing the texts by scanning them. When our digitization was complete, we used the software program *Voyant Tools* to test the frequency of the words we hypothesized would occur based on our categories.
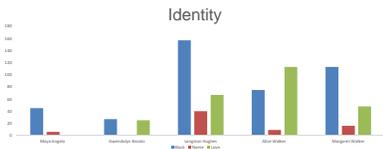
RESULTS

We selected these terms to represent the characteristics of African American poetry, as they best visualize the qualitative analysis of the overlapping themes of the selected works.
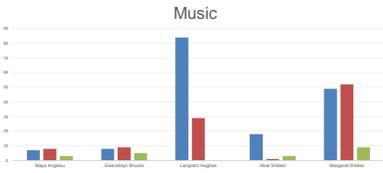
MEMORY

Memory
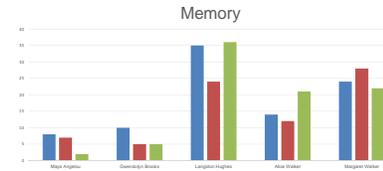
IDENTITY

Identity

MUSIC

Music

CONCLUSIONS

The data collected from the first category, memory, yields some significant conclusions. Firstly, all of the poets frequently use the term blood. The recurrence of blood purports the connection that black poetry has to ancestry and heritage. It also points to a collective-consciousness among African-Americans. In addition to blood, death is another word that our poets frequently use. This heavy usage may indicate a link between black poetry and the historical oppression blacks have experienced. The second data set, identity, reveals that race is an important component of identity for all of the writers. Furthermore, the results suggests that the older poets, Hughes and Margaret Walker, connect naming to identity. In contrast, the younger poets place significantly less emphasis on this correlation. The results yielded from the last category suggests that Hughes and Margaret Walker draw upon the trope of song more heavily than the other poets. This usage attests to their reliance on the vernacular and oral tradition.

REFERENCES

Angelou, Maya. *The Complete Collected Poems of Maya Angelou*. New York: Random House. 1994.Print.
B., Du Bois W. E., Henry Louis Gates, and Terri Hume. Oliver. *The Souls of Black Folk: Authoritative Text, Contexts, Criticism*. New York: W.W. Norton, 1999.Print.
Brooks, Gwendolyn. *The Blacks*. Chicago: Third World Press, 1994. Print.
Gates, Henry Louis. *Figures in Black: Words, Signs, and the "racial" Self*. New York: Oxford University Press, 1987. Print.
Hughes, Langston. *The Collected Poems of Langston Hughes*. New York: Vintage, 1995. Print.
Schulz, David A. *Coming up Black; Patterns of Ghetto Socialization*. Englewood Cliffs, New Jersey: Prentice-Hall, 1969. Print.
Thompson, Daniel C. *Sociology of the Black Experience*. Westport, Connecticut: Greenwood, 1974. Print.
Walker, Alice. *Collected Poems: Her Blue Body Everything We Know - Earthling Poems 1965-1990*. London, England: Phoenix, Orion, 2005. Print.
Walker, Margaret. *This is My Century, New and Collected Poem*s. Athens, GA: University of Georgia Press, 1989. Print.

ACKNOWLEDGEMENTS

Summer Scholars Program — Center for Undergraduate Research and Scholarship

Innovate · Investigate · Discover