

DELIMITING CULTURAL BORDERS:  
THE USE OF *WORDSMITH TOOLS* TO IDENTIFY  
CULTURAL DIFFERENCES IN LANGUAGE USES  
BY WHITE OR BLACK RAPPERS

PEDRO ÁLVAREZ-MOSQUERA  
*University of Salamanca*  
*pedro@usal.es*

ABSTRACT

In the field of linguistics, music has become an increasingly important field of study. Authors such as Alim (2006), Beal (2009), Cutler (2010) and Fought (2006), among many others, have based some of their studies on music in order to describe linguistic patterns, explore cultural phenomena, or study the use of a particular language feature. While conducting a sociolinguistic research project on processes of *language crossing* in rap music, we came up with two independent corpora corresponding to more than 30 rap songs by each ethnic group (European American and African American) studied. Our purpose is to explore how this data can potentially be used as a valuable source of information about these two interacting groups. This article presents some significant results from processing our corpora through *Wordsmith Tools*, which both delimits and reinforces cultural differences in rap language usage. The use of the term *nigga*, the presence of different terminology to refer to females in both corpora, and the explicit skin color references made by both rap groups creates an ethnic line that delimits their language uses and has particular relevance in the hip-hop context.

KEYWORDS: Delimiting language uses; ethnic corpora; *Wordsmith Tools*; rap music.

1. Introduction and theoretical background\*

In order to explore the ethno-cultural borders among individuals of the two ethnicities analyzed in this article, it is necessary to understand our approach in studying the processes of *language crossing* in rap music. *Language crossing* is defined by Rampton as “the use of language varieties associated

---

\* This article is funded by the Junta de Castilla and Leon project SA012A09.

with social or ethnic groups that the speaker does not normally ‘belong’ to” (Rampton 1995: 14). Other influential studies, such as Hewitt (1986) or Gilroy (1987), have contributed to building the foundation and describe some of the mechanisms of this complex linguistic process that takes place at the intergroup level. According to these studies, processes of language crossing have strong ties to a person’s identity, and they can help explain how an individual views him or herself in society. Lanehart affirms that “people create their linguistic systems (and we all have more than one) so as to resemble those of the groups with which from time to time they wish to identify” (Lanehart 1996: 322–323; see also Beal 2009: 226–227). However, when an individual appropriates a linguistic variety which is not his or her own, the existence of strong ties between language, identity and power (among other variables) makes the willingness to identify himself/herself with other social groups as a potential source of conflict. Rampton, while studying the linguistic interactions among speakers of different ethnic groups, observed that there were contextual situations where, for instance, the use of Creole by adolescents of Panjabi and Anglo descent “risked being seen as making a serious claim about really being black” (Rampton 1995: 217; see also Rampton 1995: 61). This observation can be also applied to the rap context, where the oral language plays a central role.

Although rap has become a worldwide phenomenon, its international origins are rooted in African American culture, and this connection remains strong (Smitherman 2000: 275). Some of the most important aspects that keep rap solidly connected to African American culture are the use of *African American Vernacular English* (AAVE) features, the persistence of African American oral patterns, and the strong concept of authenticity. The main reasons that explain the interconnection among these elements stem from this ethnic group’s own history. From the time of African slavery in America through the abolishment of slavery in 1865, Africans maintained a strong oral tradition that was borne out of the experiences and conditions that African slaves endured throughout this period. Smitherman describes this clearly when she asserts:

THE EBONICS SPOKEN in the US is rooted in the Black American Oral Tradition, reflecting the combination of African languages (Niger-Congo) and Euro American English. It is a language forged in the crucible of enslavement. US-style apartheid, and the struggle to survive and thrive in the face of domination. [...] Using elements of the white man’s speech in combination with their own linguistic patterns and practices, enslaved Africans developed an oppositional

way of speaking, a kind of counterlanguage, that allowed for the communication of simultaneous double meanings  
(Smitherman 2000: 19; see also Coupland 2007: 104.)

Based on Smitherman's words, one can infer that because AAVE is intrinsically woven into the history of the African American community, there are ethnically marked words, expressions and oral patterns present in their speech. Therefore, the appropriation of AAVE features by non-African American speakers might present, as Rampton anticipated, important identity implications that can potentially become a source of conflict between *crossers*<sup>1</sup> and AAVE speakers.

When these appropriations are transferred to the rap scene, the prominence of the oral component not only underlines the ethnic origins of this music genre, but also reinforces them because it incorporates other cultural patterns. To understand this, we must briefly refer to some relevant aspects of the *Black Oral Tradition*. More specifically, it is important to emphasize the African concept of *Nommo*, which is, "the magic power of the Word" that slaves carried from Africa to the United States (Smitherman 2000: 203; see also Alim 2006: 106). Due to its prominence in rap, Smitherman defines rappers as "a post-modern African griot, the verbally gifted storyteller and cultural historian in traditional African society" (Smitherman 2000: 269; see also Baker 1993: 91; Costello and Wallace 1990: 96–98). This ethno-cultural significance is also reflected in other daily oral practices, such as *signifying* and *playing the dozens*, or *boasting* and *bragging* (among others), which are associated with African American speakers and are also prevalent in rap music (Fought 2006: 172–173; Kochman 1981: 66–67; Morgan 2001: 56–57; Smitherman 2000: 218–222). The former two broadly refer to a verbal game that Mitchell-Kernan defines as "the recognition and attribution of some implicit content or function which is obscured by the surface content or function" (apud Morgan 2002: 56), although *playing the dozens* can be described as a specific type of *signifying* that includes the famous *yo mama jokes*. *Boasting* and *bragging* constitute a vocal exercise of self-praise or claims to superiority over others (Kochman 1981: 63). *Bragging*, and its more exaggerated and comic version, *boasting*, are normally misperceived by non-African Americans as being arrogant. *Bragging* and *boasting* also have an

---

<sup>1</sup> *Crosser* defines any individual performing a process of language crossing into any language variety associated with a social group that this individual does not normally associate with (Rampton 1995: 14).

important social function because “superior performance within a community context has a spiritually unifying uplifting effect, one which [...] allows the group and performing individuals to achieve a sense of fulfillment” (Kochman 1981: 67). In fact, Fought underlines how these practices serve “as an explicit indicator of how apparently spontaneous and insignificant joking rituals can be crucial windows into the construction of ethnic identity” (2006: 166).

Taking into account the prominent linguistic and socio-cultural load present in rap music, it is easy to imagine how language can be used to delimit the borders between African American rappers and non-African American rappers. In fact, since individuals from different ethnic groups have become rappers (and are actively involved in hip-hop culture), it draws attention to the importance of the language that they use because it carries with it implications of *authenticity*, especially within the context of the United States (Cutler 2010: 249).

Rap, as we know it today, originated and developed in the United States, where racial tensions have always run deep. In fact, these tensions can be explored from the linguistic point of view by analyzing AAVE. AAVE is spoken by a minority group (African Americans) and it connotes low social status that is often associated with any non-standard language variety. However, AAVE has been a central component of rap ever since its beginnings in the South Bronx of New York City (Baker 1993: 88), and it functions as a clear indicator of *authenticity*. Baker considers rap “not only a style most readily accessible to black urban youngsters, but also a representational black urban *authenticity* of performance” (Baker 1993: 62; see also Bennett 1999: 3; Best and Keller 1999). This should not be underestimated, since AAVE features and other African American oral patterns dominate the rap music scene. There have been cases, such as the European American rapper *Vanilla Ice*, Robert Van Winkle’s artistic name, who tried to connect his origins and past to the ghetto (another site of *authentication*). This reinforces the importance of authenticity: “Vanilla Ice’s desire to be a ‘white negro’ [...] to ‘be black’ in order to validate his status as a rapper hints strongly at the degree to which ghetto-blackness is critical code in rap music” (Rose 1994: 11–12). White rappers clearly understand the significance and importance of authenticity in terms of language uses or socio-cultural origins (Álvarez-Mosquera 2012: 109–123).

In this article, based on the intercultural and interethnic nature of rap music, we maintain that the linguistic component is at the core of this genre, due to its capacity to lend the speaker (the rapper) authenticity. Since it has

been proven that language elements and linguistic patterns of a particular language variety can be appropriated by speakers of other varieties (Rampton 1995), we hypothesize that although white rappers may try to gain authenticity by reproducing AAVE features, significant linguistic differences still exist. In order to test our hypothesis, we will analyze two ethnic corpora, created from over 70 songs by African American and European American rappers and evaluated using *WordSmith Tools*.

## 2. Materials and method

This study represents the continuation of a previous research article presented at the University of Coruña (Álvarez-Mosquera 2010), where we focused on a few linguistic features that partially confirmed the existence of an ethnic borderline with regards to language use. In Álvarez-Mosquera (2010), *WordSmith Tools* was also used to observe quantitative and qualitative differences (and even pinpoint some similarities) by focusing on a specific set of culturally loaded linguistic uses that were relevant in the selected rappers' corpora (see Section 3). Methodologically, we will use the same corpora and apply similar procedures. Regarding the corpora, the following sociolinguistic parameters were taken into consideration for gathering the corpora subject of analysis:

- (i) Origin:<sup>2</sup> All rappers were born in the New York area in order to avoid regional linguistic variation as much as possible.
- (ii) Age: All rappers were of similar age (20–30 years old) at the date of publication of the songs that we analyzed in order to avoid generational differences in their language uses.
- (iii) Gender: All rappers analyzed in this study were men in order to avoid gender differences in language.

Taking these selected parameters into consideration isolates the *ethnic variable* by avoiding other social factors that may alter our concluding results. In

---

<sup>2</sup> This parameter might be affected by rappers' mobility. Chris Palko (Cage) was born in the city of Würzburg (Germany), but his parents, who were American, moved back to New York when he was four years old. In addition, the rapper Erik Schrody (Everlast) was born in Long Island, but he also lived in Los Angeles.

that respect, focusing on rappers from New York was also very relevant for this study. Contemporary rap music started in New York City, and the city's size makes it possible to identify six well-established rappers who were born there. Also, from a linguistic point of view, there is less overlap among local language varieties associated with European Americans and AAVE (Wolfram and Christian 1989: 18–19) than in many other areas of the United States (e.g. Southern varieties). We also believe that it was necessary to analyze albums that were published in approximately the same year to restrict the possible variations resulting from *trendy* language uses.

Our analysis does not only take into account sociolinguistic factors, but it also examines the structure and content of rap lyrics in order to obtain results that are as accurate as possible. Therefore, we only accounted for the words in the choruses in rap songs once because of their repetitive nature (by removing their consecutive appearances), which could have tainted our final results by significantly increasing the frequency of a particular term. To illustrate this point, the presence of any word (e.g. crime) would be multiplied in a chorus by the number of lines and its repetition within the chorus of a song. As a result, this specific word would have been recorded as one of the most frequent terms in our global *Wordsmith Tools* list, in spite of possibly only appearing in one particular song, by one rapper. In addition, format words such as *title*, *album*, *chorus*, etc. contained in the lyrics were also removed so that only the actual words spoken by the rappers were the subject of analysis in this study. The next step consisted of listening to every song and using the official lyrics and information provided in the original CDs, as well as the website Ohhla.com<sup>3</sup> (*Original hip-hop lyrics archive*), to remove the sections sung by other artists whose social characteristics (ethnicity, age, origin) did not meet the research needs of our study.

In order to observe ethnic differences from a corpus-based approach, we selected around 12 songs from an album by the Beastie Boys and Public Enemy (1980's), Everlast and 2Pac (1990's) and finally, Cage and 50 Cent (2000's)<sup>4</sup>. Due to the nature of our data (songs) and the editing that took place in order to fulfill the parameters that we outlined, our corpora were

---

<sup>3</sup> We used this website to obtain the lyrics that were not included in the original CDs. The website was also used by Morgan (2002).

<sup>4</sup> The selection of decades, ranging from the 1980's through the 2000's, practically covers the entire period during which rap music has been available to the mainstream market. In addition, including more than one decade prevents our analysis from overemphasizing particular language uses that respond to contextual factors that would have the potential to distort more conclusive results. For example, the term *9/11* might appear in many rap songs after 2001, but for

largely reduced, resulting in an European American corpus of 14,865 words and 17,221 words for the African American corpus. Nevertheless, we intend to increase the size of each corpus for further study.

Finally, the last step of our methodological approach consists of processing the corpora through *WordSmith Tools*, in order to identify cultural differences in language uses by the three European American rappers and the three African American rappers. More specifically, we used *Wordlist*, which is a tool that allows you to create word lists ordered alphabetically or by frequency, along with generating statistical information in order to investigate the use of the term *nigga* and the different terminology used to refer to females. Additionally, the *WordSmith Tools* option *Concord*, which allows you to search for a word/phrase and display it as concordance lines, was used to study the explicit skin color references.

### 3. Corpus analysis

This article represents a continuation and a more in-depth analysis of an earlier research work, which presented preliminary results exploring these ethnic corpora by using *Wordsmith Tools* from a sociolinguistic perspective. In that study, we used this software to detect quantitative and qualitative differences and search for similarities. Due to its relevance for the present article, it is important to shortly rehash some of the most relevant data that led us to detect some ethnic bias (Álvarez-Mosquera 2010).

First, both ethnic groups presented similar patterns in their use of the personal pronoun *I* (and other references to the first person, see Table 1). This finding reinforced the rapper's connection to the *Black Oral Tradition* and the role of the African *griot*.

However, there was a significant divergence in the way white and black rappers conveyed the sense of community. The less frequent use of the pronoun *we* (Table 1), and the scarce references to the *hood* (see Table 2), set white rappers apart from black rappers, underlining their inability to relate to both the central idea of community, and their major site of *authentication*.

Similarly, although *gun*, *kill*, *shot*, *fight* and *police* were widely used by both ethnic groups, the term *police* drew another ethnic boundary between both groups as a result of their different experiences with law enforcement (see Table 3).

---

obvious reasons, earlier rappers did not use it at all. Thus, we scrubbed the data to ensure that these types of patterns did not impact the data.

Table 1. Use of *I/We* by European American and African American rappers.

White rappers					Black rappers				
N	Word	Freq.	% Texts	%	N	Word	Freq.	% Texts	%
1	I	489	3.3135	3 100	1	I	550	3.2031	3 100
2	MY	264	1.7889	3 100	2	YOU	429	2.4984	3 100
3	YOU	196	1.3281	3 100	3	MY	286	1.6656	3 100
4	I'M	187	1.2671	3 100	4	ME	252	1.4676	3 100
5	ME	151	1.0232	3 100	5	I'M	236	1.3744	3 100
6	LIKE	119	0.8063	3 100	6	GOT	152	0.8852	3 100
7	GOT	117	0.7928	3 100	7	GET	148	0.8619	3 100
8	HER	101	0.6844	3 100	8	YOUR	138	0.8037	3 100
9	ALL	84	0.5692	3 100	9	WE	121	0.7047	3 100
10	GET	82	0.5556	3 100	10	THEY	117	0.6814	3 100

Table 2. Use of *hood* by European American and African American rappers.

N	Concordance (White rappers)
1	like I'm from the hood
2	not everyone can relate to the hood
N	Concordance (Black rappers)
1	it's all good I ain't fresh out the <b>hood</b> , I'm still in the hood Black rims
2	I'm still in the <b>hood</b> Black rims, black hemi
3	That's why we never ever ever see you in the <b>hood</b> with it
4	I have ya outlined in chalk (I-I Get It) In the <b>hood</b> if ya ask about me
5	I come creepin through the <b>hood</b> wearin teflon Hit the corners
6	From <b>hood</b> to the 'burbs, everyone of you niggas
7	cheap tricks from gettin on her... life in tha <b>hood</b> ...



Table 3. Use of violent terminology by European American and African American rappers.

White rappers			Black rappers		
	Frequency	Position		Frequency	Position
GUN	27 times	43	GUN	17 times	97
KILL	14 times	110	KILL	14 times	127
SHOT	17 times	89	SHOT	15 times	119
FIGHT	7 times	229	FIGHT	4 times	485
POLICE	1 time	2.431	POLICE	7 times	312

In other words, white rappers incorporated and reproduced terms within the scope of social experiences of a white individual (e.g. owning a *gun* – 27 repetitions), but they could not appropriate all that is encompassed by the *black experience*<sup>5</sup> (e.g. being a suspect by the color of your skin; perceiving *police* as an oppressive force – 7 repetitions).<sup>6</sup>

In order to reaffirm these preliminary results and corroborate our hypothesis, we considered it essential to increase the range of features and linguistic patterns that we analyzed. The results we are presenting here not only reinforce, but also expand the conclusions derived from the previous study.

<sup>5</sup> In general terms, the expression *Black Experience* refers to the collection of historical, social and economic conditions that this ethnic group has endured throughout its history on the North American continent.

<sup>6</sup> By using *WordSmith Tools* in our analysis, we learned that white rappers' usage of explicit violent terminology was quantitatively higher than that of the black rappers. However, due to the difference in the violent terms most repeated in the white corpus (e.g. *gun*, *fight*, *shot*, etc., but not *police*), combined with the fact that black rappers usually locate their raps in the ghetto (place of authentication with many violent connotations), led us to analyze one rapper of each ethnic group from a qualitative point of view, embracing the entire semantic frame of violence (Álvarez-Mosquera 2013). The results indicated that 2Pac's (African American rapper) wider variety of terminology not only undermined the quantitative significance of many violent words (less repetitions of the same term), but they also could be denotative of the ghetto context, a fact that adds credibility or authenticity to the rapper. For example, while Everlast (European American rapper) used the word *gun* (8 times) and *chrome .45* (type of gun) on one occasion, 2Pac referred to firearms as the *nine* (9mm gun), *my steel*, *my pistol*, *my glock* (type of gun), *semi-automatic Mack 11* (type of gun) or *my R-15* (type of gun) on different occasions. With regards to the use of the term *police*, we must note that, unlike 2Pac, there were not any references to the police as a source of violence or conflict in the white corpus. This unequal use of the term *police* might again undermine the authenticity of white rappers, since it suggests that they do not experience the same social reality as black rappers.

In this study, we again used the *Wordlist* option to obtain the list of words contained in every corpus, organized by the number of times a particular word was been used by every ethnic group while rapping (frequency). Notably, we added the same *Stopword list* containing around fifty grammatical words (e.g. *the, a, to, of,* etc.) to both corpora. Their high index of frequency and their low relevance for our study allow us to suppress them in order to obtain a clearer picture of the results and significant differences.

*Nigga* is the first word to analyze. This controversial term stands out in the African American corpus (position 16 in the ranking, Table 4).

Table 4. *Nigga* stands out in the African American corpus.

N	Word	Freq.	%	Texts	%
1	I	550	3.2031	3	100
2	YOU	429	2.4984	3	100
3	MY	286	1.6656	3	100
4	ME	252	1.4676	3	100
5	I'M	236	1.3744	3	100
6	GOT	152	0.8852	3	100
7	GET	148	0.8619	3	100
8	YOUR	138	0.8037	3	100
9	WE	121	0.7047	3	100
10	THEY	117	0.6814	3	100
11	NOW	106	0.6173	3	100
12	IS	102	0.594	3	100
13	KNOW	102	0.594	3	100
14	LIKE	102	0.594	3	100
15	ALL	100	0.5824	3	100
16	NIGGA	96	0.5591	2	66.7
17	BE	92	0.5358	3	100
18	BUT	92	0.5358	3	100
19	WHEN	85	0.495	3	100
20	DON'T	84	0.4892	3	100

In fact, once we added all of its *lemmas*, or lexical units presenting variations of the word *nigga* (e.g. *Niggaz*, *niggas*, *niggaz'll*, etc.), this ethnic reference (and way of addressing another African American) registered 188 repetitions, moving up to the sixth position on the *Wordlist*. While *nigga* for many years constituted a derogatory way to refer to African Americans, nowadays *nigger* (or *nigga*, depending on the spelling) has been strongly stigmatized as an insult by the public (Kennedy 2002: 172). Despite this fact, the implications of this word have been renewed, especially among African Americans themselves. As Kennedy (2002: 174) states, “[b]lack use the term with novel ease to refer to other blacks, even in the presence of those who are not African American”. However, cases of using *nigga* (more than *nigger*) among individuals of other racial groups as a sign of affection can be also heard in the streets of major American cities, although the social reactions are very different (Spears 2006). The controversy around this word makes its study even more relevant from an ethnic point of view. Returning to our corpus-based approach, our results show that despite the high rate of appearance, only two of the three African American rappers studied here use this word (see Table 4); everyone apart from Public Enemy.

Although further research has to be done, we propose two possible reasons to justify these results. First, Public Enemy’s distinct language usage could demonstrate the sentiment in the 1980’s, when rap music was struggling to gain acceptance in the mainstream market and it was being strongly criticized by the media, politicians, parents, and conservative groups (Rose 1994: 4) for its violent language and explicit content. To underline the effect of this pushback, Public Enemy also seemed to reduce its usage of AAVE, in order to increase its mainstream appeal. This idea is supported by the fact that this rapper scored lower than the other two black rappers in the use of AAVE features, such as *ain’t*, *copula deletion*, the absence of the third singular present tense *-s*, and the realization of the final *ng* (velar) as *n* (alveolar) in gerunds (Álvarez-Mosquera 2012: 73–88). Additionally, *flippin the script* or *semantic inversion*, a practice associated with AAVE speakers, may help explain these results. AAVE speakers can use *Standard American English*<sup>7</sup>

---

<sup>7</sup> It is also known as MUSE (Mainstream U.S. English), among other names. For the purpose of this paper, we would like to point out that we are referring to this variety of English in the broad sense, including standard and non-standard uses. It is the English variety spoken by most of the Caucasian population of United States and many speakers of different ethnic groups. *Standard American English* is also the institutional language variety, the vehicle of communication in education and an indispensable tool for professional and social progress.

words to invert their meaning, or provide them with a new one. In Smitherman's words:

Semantic inversion/flippin the script was an act of linguistic empowerment as Africans in America took an alien tongue and made it theirs, simultaneously they created a communication system that became linguistically unintelligible to the oppressor, ever though it was his language. [...] When an AAL term crosses over and gains linguistic currency in the EAL, AAL speakers generate a new term to take its place. [...] It symbolizes linguistic resistance to the dominant culture's lexicon. (Smitherman 2000: 280.)

The fact that Public Enemy does not use the term *nigga*, might also indicate that the *semantic inversion* took place after their album's launch.

As for the European American rappers, the word *nigga* presents zero instances in their corpus. This drastic imbalance represents a profound ethnic barrier between both groups, but its implications are greater in rap terms than in social terms. In the hip-hop scene, white rappers cannot use the term *nigga*, due to the fact that their cultural background does not share the *black experience*. Therefore, while appropriating other AAVE features might be relatively easy to a certain extent (e.g. violent terminology), they cannot cross this ethnic line without being offensive, due to the racist connotations and historical underpinnings of the word. This undermines their chances of sounding more *authentic*, since African American rappers frequently use this term.

A closer look at our corpora demonstrates that the use of the word *nigga* is not the only source of divergence between the two groups. Although references to women in general are quite numerous in both corpora, the presence of the word *bitch* (39 repetitions, position 45) stands out in the African American corpus. The topic of sex and women has been broadly covered in many studies and it could even be considered a type of rap itself: "while some raps convey social and cultural information, others are used for conquering foes and women" (Smitherman 2000: 207). Smitherman also remarks on the presence of women in rap when she explains that "while boastful raps are used to devastate enemies, love raps help in gittin ovuh with women" (2000: 208). In our study, we took into account five references to females that are common in social terms, but also relevant in terms of frequency in both groups' corpora<sup>8</sup> (see Table 5).

---

<sup>8</sup> This data contains the lemmas of every word.

Table 5. WordSmith Tools (female references).

White rappers			Black rappers		
	Frequency	Position		Frequency	Position
SHE	96 times	9	SHE	61 times	28
GIRL	63 times	17	BITCH	59 times	29
BABY	12 times	123	BABY	34 times	55
Other references			Other references		
WOMAN	3 times	715	GIRL	22 times	76
BITCH	1 time	1.341	WOMAN	3 times	777

An in-depth analysis of this topic could warrant a separate study, especially if one were to examine the specific relationship between gender and its cultural implications (Smitherman 2000: 208–209). However, our results begin to reflect some of the most important cultural differences between both groups. In general, there is a balance in the number of references to women. The African American corpus presents 179 instances, while there are 175 references to women in the European American group. Although both groups agree on the most common way to refer to females (*she*), they vastly differ when it comes to using lexical words to refer to them. The fact that *bitch* is the second most used female reference (the first lexical one) in the African American corpus (59 times), while *girl* is the second most used female reference in the European American corpus (63 times) may lead us to dismiss the ethno-cultural differences between both groups of speakers and conclude that African American rappers are more disrespectful towards women. This sentiment can be taken further if we observe the third most common reference to women: *baby*. *Baby* presents even more repetitions in the African American corpus than the word *girl* itself.

However, despite being a pejorative word in Standard American English, we should try to understand the use of this terminology in light of African American rap culture. Smitherman, an African American author herself, expresses it in the following way:

Contrary to popular stereotype, black men never really regarded black women as sex objects, pure and simple, for the love rap, based on the African view of the reconciliation of opposites, is a synthesis of emotional and intellectual appeal and has as its ultimate objective the conquest not simply of the woman's body, but her mind as well.

(Smitherman 2000: 208–209.)

Smitherman (2000: 209) insists that “[r]appin also accounts for what whites often label as ‘aggressive,’ ‘brash,’ ‘presumptive,’ or ‘disrespectful’ behavior by black men to black women”. However, she challenges us to look beyond the literal interpretation of this terminology in Standard English and reflect upon the cultural context of African Americans. By doing so, we can see how these references to women can be understood as playful (in most cases), rather than a harmful attack. In opposition to this cultural pattern, European American rappers resort to using the term *girl*, due to the fact that their own language variety (which would be closer to Standard American English) would not allow them to use this term without sounding inappropriate or misogynistic.<sup>9</sup> Again, the language used by both groups draws an ethnic borderline between them, which depicts the profound difference that context and culture can have.

The last topic of analysis in this paper is the use of explicit skin color references by both groups. We sought to understand if these types of references would be a mistake for European American rappers to make, as alluded to by Erik Parker, editor of *Vibe*<sup>10</sup> magazine, in the article “Rap’s white invasion” (2003). *Wordsmith Tools*, through the *Concord* option, finds all of the occurrences of a search word and displays the results together with a piece of co-text before and after the word. Since this option provides the textual context around the term, it is vital in identifying accurate references. Although our data yielded a low number of instances of explicit skin color references, the results point towards two clear, opposing tendencies.

In the European American corpus, we searched for the term *black* and found eight instances, but only one presented slight ethnic connotations:

- (1) “Black kids call me Whitey.” (Everlast).

Continuing with our research, we found that African American rappers’ use of the word *white* was present in four cases, with three of them explicitly referred to skin color, and carrying strong ethnic connotations:

- (2) “It’s the white man I should fear.” (2Pac)

---

<sup>9</sup> Further qualitative research is needed.

<sup>10</sup> *Vibe magazine*, which was founded in 1993, promotes urban culture, global music, lifestyle, etc. to more than eight million readers throughout the world. Website: [www.vibe.com](http://www.vibe.com).

- (3) “So just get you a white girl, don’t fuck with no black bitch.” (50Cent)
- (4) “The one who makes the money is white not black.” (Public Enemy)

Despite the few cases found in our corpora, the results support Parker’s statement, since skin color references about the other ethnic group have different connotations. While the European American group only uses these references rather anecdotally, the African American rappers associate white skin color with fear or wealth. Example three is particularly interesting, since African American rappers refer to women as *white girl* vs. *black bitch*, thus reinforcing the differences between how the two ethnic groups refer to women (see previous section).

In order to complete our approach, we also decided to look into the use of the same ethnic terms contained in each corpora in relation to their own racial groups. On this particular occasion, we included the complete list of results (with a larger number of cases beyond those only referring to skin color), in order to illustrate the different patterns in both groups. Consider the results in Tables 6 and 7.

Table 6. The use of *white* by European American rappers.

N	Concordance
1	the saints come marchin’ in . . . I messed the alpine <b>white</b> , classic rapper’s delight
2	Black kids call me Whitey Spanish kids Whito <b>White</b> kids call me king of
3	Nike’s Walked and talked like a rapstar But was <b>white</b> , and did it right Before girls,
4	what we’re all about We went to <b>White</b> Castle and we got thrown out
5	when I ring the bell I chill at <b>White</b> Castle ‘cause it’s the best But I’m fly
6	make them smile From <b>White</b> Castle to the Nile Back in the day
7	Got the ladies of the eighties from here to <b>White</b> Castle Hold it now
8	bass and highs <b>White</b> Castle fries only come in one size What you see

As depicted in Table 7, the African American corpus contains eleven repetitions of the term *black*. More importantly, out of the eleven instances, ten ca-

Table 7. The use of *black* by African American rappers.

N	Concordance
1	It's the hour to the minute – time to blow <b>BLACK</b> All right party people, bust a groove
2	The one who makes the money is white not <b>black</b> You might not believe it but it is like
3	She don't want a brother that's true and <b>black</b> If you're light, you're alright -
4	I'm still in the hood <b>Black</b> rims, black hemi, nigga see me when ya see me
5	out the hood, I'm still in the hood Black rims, <b>black</b> hemi, nigga see
6	St. Tropez Get a tan, I'm already <b>black</b> , rich, I'm already that Gangsta,
7	don't fuck with no <b>black</b> bitch I got two felonies, from sellin that crack shit
8	No more hesitation each and every <b>black</b> male's trapped And they wonder why
9	There's a ghetto up in Heaven and it's ours, <b>Black</b> Power is what we scream
10	Fresh out of jail, life's Hell for a <b>black</b> , celebrity So that's the reason why I call
11	glock So now I gotta throw away Floatin in the <b>black</b> Benz, tryin to do a show a day

cases specifically refer to skin color (e.g. Fresh out of jail, life's Hell for a *black*). However, the list of references obtained in the European American group offer fewer cases in total (eight instances of the word *white*), and only two present racial connotations (e.g. *Nike's Walked and talked like a rapstar [b]ut was white*) (see Table 6). Therefore, these results reinforce the pattern previously detected when analyzing the use of the term *nigga*, the words *police* and *hood*, and female references. That is, European American rappers present cultural and linguistic limitations that might undermine their authenticity.<sup>11</sup> In this case, our data confirms that skin color references are a relevant issue in rap music and culture, and they underline the strong connection between rap and the African American community.

<sup>11</sup> By focusing our study within the United States, where the presence of the African American ethnic group is relevant in the rap scene, especially in New York, where rap music started, we are underlining the importance of being authentic in rap. Despite that, we must acknowledge that hip hop has become global and therefore, there may be European American rappers who establish their own credibility by targeting their specific audiences and distancing themselves from African American linguistic features.



#### 4. Conclusion

This article has proven that rap music presents an opportunity to study the complex interaction between ethno-cultural factors and language. This study's relevance lies in the fact that rap is strongly rooted in African American culture, although there is an increasing number of non-African American rappers. As we hypothesized, from the linguistic point of view this situation is very significant, since it places the linguistic component at the core of this genre, due to its ability to lend the speaker authenticity, a key concept in the hip-hop scene. By maintaining a sociolinguistic approach and using *WordSmith Tools* in our analysis, we explored their language uses in order to track the potential presence of diverging, or converging, patterns. This software has been effective in delimiting cultural borders between both ethnic groups. In fact, the *WordSmith Tool* results in this study have confirmed the preliminary findings of our previous work (Álvarez-Mosquera 2010), contributing relevant information that depicts the existence of profound differences in rappers' language production.

A clear example of a word that alienates both groups is the use of the term *nigga* 'nigger'. By using *Wordlist*, we determined that this controversial term is present in zero instances in the European American corpus, while it is used quite frequently in the African American group. However, the statistical information shown by this software revealed that this word is absent in Public Enemy's songs. This fact weakens our conclusion that *nigga* is a clear delineating factor between the two groups of rappers. In order to further test our conclusion, Public Enemy should be the subject of greater research.

*Wordlist* was also the tool selected to analyze the use of female references. The data obtained emphasized the existence of an ethnic barrier in rap. In this case, cultural differences become more prominent when we specifically focus on the frequency of use of the word *bitch* by African American rappers and *girl* by the European American rappers. Both ethnic groups process the term *bitch* in different ways, due to their cultural background. European American rappers cannot use this term without sounding extremely offensive, opting instead for the term *girl*, which is a word that is more closely aligned with their own language variety and cultural expectations. However, AAVE speakers filter the negativity associated with using the term *bitch* by cognitively processing this language usage within their own socio-cultural context. In fact, the use of these cultural resources justifies why the word *bitch* is the second most used female reference.

Finally, skin color references illustrate two different meanings, according to the rappers' corpora. By using the *Concord* option in *WordSmith Tools* to observe the searched terms in their textual context, we verified that while European American rappers barely used explicit references to the skin color of their own group, or the African American group, African American rappers not only used them, but they also attached important social relevance to skin color. These results emphasize the important role that race still plays in rap music.

*WordSmith Tools* can be an advantageous instrument to locate and define cultural borders between these two groups of speakers, as well as to help enrich the sociolinguistic field. This study proves that although the linguistic component is placed at the core of rap music, due to its capacity to lend the rapper authenticity, and despite acknowledging that language elements and linguistic patterns of a particular language variety can be subject to appropriation by speakers of other varieties (Rampton 1995), the marked African American origins of this music genre constitute a serious obstacle for European American rappers in the context of the United States. Therefore, the data obtained in this study corroborates and expands on the findings in Álvarez-Mosquera (2010), addressing the linguistic limitations of European American rappers, which can undermine their authenticity. Although further research is still necessary to expand upon our conclusions, the results raise interesting questions regarding ethnic relations, language, and identity construction especially among younger generations. Using *WordSmith Tools* to compare ethnic corpora among other groups that interact can help detect converging, or diverging, patterns in intercultural communication, keeping in mind each particular social context. Analyzing such similarities and differences would help us understand how identities are linguistically negotiated, and how group members draw the linguistic limits of their own identity and its consequences for the out-group individuals.

## REFERENCES

- Alim, S.H. 2006. *Roc the mic right: The language of hip hop culture*. London: Routledge.
- Álvarez-Mosquera, P. 2010. "Exploring the use of WordSmith Tools for sociolinguistics purposes: A case study of cultural loaded language uses in White and Black rappers' corpora". In: Moskowich, I. et al. (eds.), *Language windowing through corpora*. Universidade da Coruña. 39–48.

- Álvarez-Mosquera, P. 2012. *Identidad y language crossing: El uso de inglés afroamericano por raperos blancos*. Bern: Peter Lang.
- Álvarez-Mosquera, P. 2013. "Exploring the semantic frame of violence: The need of integrating quantitative and qualitative analysis". In: Llanes, À. et al. (eds.), *Applied linguistics in the age of globalization*. Lleida: Edicions i Publicacions de la Universitat de Lleida. 439–446.
- Baker, H.A. 1993. *Black studies, rap, and the academy*. Chicago: University of Chicago Press.
- Beal, J.C. 2009. "'You're not from New York City, you're from Rotherham': Dialect and identity in British indie music". *Journal of English Linguistics* 37(3). 223–240.
- Bennett, A. 1999. "Rappin' on the Tyne: White hip hop culture in Northeast England – An ethnographic study". *The Sociological Review* 47(1). 1–24.
- Best, S. and D. Keller. 1999. "Rap, black rage and racial difference". *Enculturation* Spring 2.2.
- Cutler, C. 2010. "Hip-Hop, White immigrant youth, and African American Vernacular English: Accommodation as an identity choice". *Journal of English Linguistics* 38. 248–269.
- Costello, M. and D.F. Wallace. 1990. *Signifying rappers: Rap and race in the urban present*. New York: Ecco.
- Coupland, N. 2007. *Style: Language variation and identity*. Cambridge: Cambridge University Press.
- Fought, C. 2006. *Language and ethnicity: Key topics in sociolinguistics*. Cambridge: Cambridge University Press.
- Gilroy, P. 1987. *There ain't no black in the Union Jack: The cultural politics of race and nation*. London: Hutchinson.
- Hewitt, R. 1986. *White yalk Black talk: Inter-racial friendship and communication amongst adolescents*. Cambridge: Cambridge University Press.
- Kennedy, R. 2002. *Nigger: The strange career of a troublesome word*. New York: Pantheon.
- Kochman, T. 1981. *Black and White: Styles in conflict*. Chicago: University of Chicago Press.
- Lanehart, S.L. 1996. "The language of identity". *Journal of English Linguistics* 24. 323–331.
- Morgan, M. 2001. "'Nuthin' but a G thang': Grammar and language ideology in hip hop identity". In: Lanehart, S.L. (ed.), *Sociocultural and historical context of African American English*. Amsterdam: Benjamins. 185–207.
- Morgan, M. 2002. *Language, discourse and power in African American Culture*. Cambridge: Cambridge University Press.
- Rampton, B. 1995. *Crossing: Language and ethnicity among adolescents*. New York: Longman.
- Rap's white invasion. 2003. <<http://www.theage.com.au/articles/2003/04/23/1050777285605.html>>. (Last accessed 03 Sep 2011.)
- Rose, T. 1994. *Black noise: Rap music and Black culture in contemporary America*. New England: Wesleyan University Press.

- Smitherman, G. 2000. *Talkin that talk: Language, culture and education in African America*. New York: Routledge.
- Spears, A.K. 2006. "Perspectives: A view of the "N-word" from sociolinguistics". *Diverse Issues in Higher Education*.  
<<http://diverseeducation.com/article/6114/>>
- Wolfram, W.A. and D. Christian. 1989. *Dialects and education: Issues and answers*. Englewood Cliffs: Prentice Hall.

**Address correspondence to:**

Pedro Álvarez Mosquera  
Departamento de Filología Inglesa  
Universidad de Salamanca  
Placentinos, 18  
37008 Salamanca  
Spain  
[pedro@usal.es](mailto:pedro@usal.es)