

A corpus approach to discursive constructions of a hip-hop identity

Kristy Beers Fägersten

Universität des Saarlandes, Germany

This chapter is an analysis of a 100,000-word corpus consisting of message-board postings on hip-hop websites. A discourse analysis of this corpus reveals three strategies employed by the posters to identify themselves as members of the hip-hop community in the otherwise anonymous setting of the internet: (1) defined openings and closings, (2) repeated use of slang and taboo terms, and (3) performance of verbal art. Each strategy is characterized by the codification of non-standard grammar and pronunciations characteristic of speech, as well as by the use of non-standard orthography. The purpose of the discourse is shown to be a performance of identity, whereby language is used and recognized as the discursive construction of one's hip-hop identity.

1. Introduction

When compared with face-to-face conversation, interaction via computer does not afford participants the same opportunities to actively gather information or even passively notice characteristics about each other through aural or visual cues. Despite the existence of web cameras and streaming video, typical computer-mediated communication (CMC) mainly consists of type-written text messages. Information about another's gender, race or age, for example, which one's appearance or style of speaking may reveal, can be undetectable – or even falsified – in CMC. As a result, there is both an element of anonymity and danger of deception associated with CMC, as well as with much other interaction or discourse on the internet. Some interlocutors take advantage of the anonymity to pass as someone or something they are not (Bechar-Israeli 1995; Cutler 1999; Nakamura 1995; Stone 1991; Turkle 1995; Van Gelder 1990; Wallace 1999); others take considerable measures to try to portray their true selves (Bechar-Israeli 1995; Cutler 1999; Warschauer 2000). Whatever their intentions may be, participants in CMC present themselves almost entirely through linguistic means, constructing their identities through their discourse. In cyberspace, you are what you type.

Corpora and Discourse

The challenges of different settings

Edited by Annelie Ädel
and Randi Reppen

Studies in Corpus Linguistics

青山学院大学図書館



00-08-12788

AMINS PUBLISHING COMPANY

The language particular to computer-mediated communication has been called the "third medium" (Crystal 2001:48; cf. Ong 1982), denoting a variety of language that is similar to spoken language, but in written form. Ferrara et al. (1991) refer to electronic discourse as 'Interactive Written Discourse' (IWD), Collet and Belmore (1996) use the term 'Electronic Language' (EL), while Takahashi (2003) uses the term 'Net-En' to refer specifically to computer-mediated communication in English, denoting a variety distinct from written or spoken forms of English. Considering the extent of comparisons between CMC and written and spoken varieties of language, both of which can vary greatly according to context, it is surprising that analyses of CMC have not focussed more on context-based variation. While CMC provides opportunity for enhancement of communication, different contexts of usage impose restrictions on it (Allwood 2000; Herring 2001). The adaptation of language to fit these different contexts thus raises the question of "how people tailor their use of written language to the conditions of medium and situation, as well as to their communicative wishes." (Hård af Segerstad 2002: 10) This acknowledgment of the effect of situation on communicative discourse reflects the "move from the 'language of CMC' to computer-mediated discourse" (Androutsopoulos 2006:421). This chapter will focus on identifying features of discourse styles in the context of hip-hop message boards. The language of hip-hop CMC will be viewed as an adaptation of written language to discursively construct identities within the context of internet message boards. Message board postings can be considered an example of the third medium (or IWD, EL, Net-En, etc.), as they reflect a tendency among contributors to write as they speak and can therefore be likened to spoken language, but nonetheless written by the 'speakers' themselves as they choose to represent their 'speech.' However, more central to the present analysis of hip-hop message board postings is the context of the discourse, in which language use contributes to a discursive construction of hip-hop identity.

In this chapter, I present an analysis of discursive constructions of hip-hop identities, based on a corpus of message board entries collected from hip-hop websites. Such a corpus study allows for both a quantitative and qualitative examination of the strategies employed by message board posters to establish their hip-hop identities in the otherwise anonymous setting of the internet. Traditionally, hip-hop communities consist of predominantly male, urban, African-American youths (George 1999). The practice of posting on a hip-hop message board serves to include posters in an online hip-hop community, where age, race, gender, and social background are not immediately known or obvious. The assertion of an online hip-hop identity is achieved discursively. The large amount of data in the hip-hop message board corpus (hereafter referred to as the HMBC) helps in determining the extent of linguistic systematicity in the discourse of the message board entries, which in turn encourages a micro-analysis of the most common strategies for discursively constructing a hip-hop identity, namely, (1) the use of

distinct openings and closings, (2) repetition of slang and taboo terms, and (3) performance of verbal art.

1.1 Hip-hop identity in message board discourse

Central to the culture of hip-hop is music, in particular, "a form of rhymed storytelling accompanied by highly rhythmic, electronically based music." (Rose 1994:2) Outside the internet realm, a hip-hop identity is furthermore defined and/or recognized in part by break-dancing, clothing, and graffiti (Berns & Schlobinski 2003; Cutler 1999; George 1999; McLeod 1999; Newman 2001; Rose 1994). In other words, in real as opposed to virtual settings, both aural and visual cues help to establish and identify a member of the hip-hop community. In a CMC environment and, central to this paper, in the message board environment, most visual and aural cues are minimized if not nullified. The anonymity associated with the medium of CMC and internet message boards does not allow for reliable identification or recognition of contributors, or even paralinguistic corroboration of their claimed identities. It is primarily through discursive means that contributors can identify themselves and each other, and thus it is a challenge to members of the hip-hop community to assert and maintain their hip-hop identities during computer-mediated communication. The discursive construction of a hip-hop identity is as much a function of *what* is posted as *how* it is composed. The content and form of hip-hop message board postings represent communal and cultural practices. Insofar as the members of the hip-hop community represent a subculture, their language can be considered a sociolect. Since hip-hop refers to a style of music as well as a cultural history of "dance, painting, fashion, video, crime and commerce [...]" (George 1999: viii), message board postings that do not concern a topic relevant to hip-hop culture risk being ignored or mocked, requiring posters to assert a hip-hop identity by way of showing a familiarity with cultural practices, events, and issues. In the example below, the poster dramatically laments the present state of hip-hop. Both the familiarity with and disapproval of the evolution of hip-hop help the poster to assert a hip-hop identity:¹

- (1) The question is what is rap going to turn into wildness and no realness? Where is the realness? Its all about clubs and partyin and fuckin and bling bling etc. everydamn thing is the same. Nothing creative, just the same ass thing.... It pisses me off. Fuck everyone for their opinions on rap when they don't even know what rap is! What hip hop is! The meaning of hip hop! And what it has revoultionized into! This Is SHIT!

1. Numbering of examples is for organization and reference only within this paper and does not indicate any order or chronology to the corpus postings.

In Example (2), the poster is more explicit about asserting a hip-hop identity, calling him/herself a 'hip hop person':

- (2) Peace, first I would like to say this has to be one of the most educating sites on Hip Hop, as a culture and as a lifestyle. As a hip hop person, I like to always educate myself on facts and to learn the history of something in my interest, and this website is doing that. Thanks.

Examples (1) and (2) reveal that the content alone of message board postings can establish or at least imply a hip-hop identity. But the examples also reveal two discursive practices of the hip-hop message board community, as well as one important characteristic of the language of message board postings. First, in Example (1), there are several lexical items which characterize the discourse of many hip-hop message board postings: 'fuckin', 'damn', 'ass', 'pisses', 'fuck', 'shit', and 'bling bling'. The repeated use of such taboo terms (e.g., 'fuckin') and hip-hop slang (e.g., 'bling bling', *flashy jewelry*) is a common discursive practice, discussed further below. Another common practice is the use of defined openings and closings, as can be seen in Example (2) in the form of the opening 'peace' (also an example of hip-hop slang), and the closing 'thanks'. There is, however, neither a defined opening or closing in (1), nor any use of taboo terms in (2), indicating a more fundamental stylistic difference between the two postings. The sentence structure, grammar and non-standard orthography used in Example (1) encourage a 'spoken' reading of the content. Example (2), on the other hand, with its opening, long sentences, standard grammar, spelling and punctuation, and closing is reminiscent of a written text such as a letter. Thus, among the message board entries can be found features of both written and spoken language: the language of CMC. It is the intervariation between (1) and (2) which underlines the necessity to move from the language of CMC to the context of CMC, calling attention to different discursive goals.

A large amount of data in the form of a corpus is conducive to the identification of patterns and tendencies in the discursive construction of a hip-hop identity. The emergent discourse strategies may not be evident in every posting, but a quantitative and qualitative discourse analysis reveals the message board forum as an intersection of written and spoken language in which posters discursively construct their hip-hop identity. Example (3) illustrates how aspects of both spoken and written language can be deftly combined with the content and form of message board entries to assert a hip-hop identity:

- (3) ey foo.....show meeh sum pic of yall nikka's breakin iight homie.....do dat fo ur boii b joe frum Under Rated Breakaz.

To the uninitiated or out-group members, this posting (and many others in the corpus) may be a challenge to comprehend. A not-too-divergent gloss would be:

Hey fool. Show me some pictures of you niggers break-dancing, alright, homeboy? Do that for your boy. I am Joe from Under Rated Breakers. In addition to two straight-forward assertions of a hip-hop identity – the poster refers to himself as ‘boi’, a hip-hop slang term, and ‘joe frum Under Rated Breakas’, implying he is the member of a break-dance group – there is also the use of slang and taboo terms (‘foo’, *fool*; ‘nikka’, *nigger*; ‘iight’, *alright*; ‘homie’, *homeboy*, *friend*; ‘boi’, *boy*, *breakdancer*). Therefore, in terms of content, a hip-hop identity is already well-established. Nevertheless, it is the form of the posting including non-standard orthography, grammar, and punctuation which dominates as an expression of the poster’s hip-hop identity. Throughout the corpus of hip-hop message board postings, there are similar examples of creative form coupled with appropriate hip-hop content. A corpus analysis of message board postings allows for certain features of discourse to be made salient, revealing how conventions of posting structure, form and content in turn allow posters explicitly or implicitly to assert their hip-hop identities.

In order both to acknowledge the distinct variety of language found on internet message boards, and to avoid unintentional alignment with written or spoken varieties of English, throughout this paper, the terms ‘to post’ (to write and submit a message on a message board), ‘posting’ (an individual message on a message board), and ‘poster’ (the submitter of a message) are used.

2. Methodology

A corpus-linguistic investigation of the discursive construction of a hip-hop identity is enabled by the quantity and quality of data available. Not only is there an abundance of websites dedicated to hip-hop culture, many with free access to message boards, but, most importantly, the message board postings themselves represent unique, raw data produced by the members of a socio-cultural community to which linguist-observers and/or out-group members might not have access. The postings furthermore capture the language of a specific medium as its users would have it represented. There are no questions of editing, nor is there, in contrast to speech, any need for transcription, which eliminates mediation and guesswork (Yates 1996).

The reliable identification of linguistic conventions in hip-hop discourse requires both a quantitative and qualitative approach to data and data collection. Determining the extent of linguistic systematicity throughout the discourse demands a large amount of data, while an examination of variations within the identified system also requires a micro-analysis. In an effort to collect enough data to discover potential discursive patterns warranting careful investigation, but at the same time reduce the chances of these patterns being specific to one on-

line community, the HMBC was composed from five different hip-hop websites; a description of the corpus and web addresses are available in Appendix A.² There are many websites devoted to hip-hop culture, but not all sites feature message boards and, among those that do, only some are heavily posted. The five sites chosen for this study featured message boards with many postings, facilitating data collection. After dates, e-mail addresses and URLs were edited out, the corpus totaled 102,343 words (tokens) with 9,822 distinct types. WordSmith Tools (Scott 2004a) was used to analyze the corpus in terms of word frequency, sorted lists, keywords and clusters. Frequency, keyword and cluster lists were compiled in order to identify any frequently occurring lexical items or clusters significantly associated with hip-hop discourse. During an initial survey of the corpus, a number of words ending in 'a' and 'z' were noticed, and thus a reverse-sort list was compiled to determine the extent of these and other non-standard spelling practices. The aim of the investigation as a whole was to identify significant forms and/or lexical items specific to the discourse of hip-hop in order to determine their discursive function within individual postings in particular, and to investigate how they correspond to or signal general strategies for the discursive construction of a hip-hop identity.

3. Openings and closings

Openings and closings can be found both in spoken language, for example in conversations (Schegloff 1972), and in written language, such as in correspondence (Danet 2002; Eiler & Victor 1988). Crystal (2001), Taboada (2004), Yates (2000), and Yates and Orlikowski (1992) provide evidence and examples of openings and closings in computer-mediated communication as well, such as in e-mails, chat group conversations and message board postings, arguing that the opening-body-closing structure of such CMC is similar to a written letter. They also present evidence of instances of CMC where the opening, closing or both are absent, a common feature of CMC which often results in rendering the text informal (Taboada 2004).

Examples (2) and (3) differ considerably in terms of stylistics and form, with (2) representing more standard, written English and (3) non-standard, spoken English. Their structures, on the other hand, are similar in as far as each includes a distinct opening and closing. Example (2) opens with the word 'peace', used within the hip-hop community as an expression of greeting or leave-taking. Example (2) then closes with the term 'thanks'. Example (3) opens with 'ey foo', a term of address which can be compared to more traditional openings of let-

2. <http://odur.let.rug.nl/~vdbeek/perl/lecture2.html>

ters or other written correspondence, such as 'Dear/hello/hi <name>'. The same post then closes with 'i b joe...' which can be similarly likened to such traditional closings as 'Sincerely yours, <name>'.

Example (1) above illustrates that openings and closings are similarly not an absolute feature of hip-hop message board postings. In fact, a close analysis of the corpus reveals that a lack of opening and closing is common to most postings belonging to the same thread, or group of related postings. The HMBC contains 1,512 unique postings, representing 464 comment-response postings and 1,048 in-thread postings. A total of 901 in-thread postings (86%) are without opening or closing, as illustrated below in Example (4) containing the first three postings of a new thread. The example features at least two different and anonymous posters, and each turn is prefaced with a number:

- (4) 1. CBS doin a hip-hop special tonight at 8 about hip-hop and it's powers....
i think i might peep it!
2. Yeah I'm a be peepin' it see whut it's about... supposed to be how deep the industry really reach...
3. And wasn't it on CBC not CBS?

Erickson (1999) refers to communication such as these message board postings as *persistent conversation* because, "although it is conversational in many aspects, it is also preserved for future review" (Taboada 2004:61). Thus, as opposed to face-to-face conversation where contributions are ephemeral, message board postings linger. Website users have the ability to read through previous postings and contribute to the topic at any time, so that the conversation as well as the topic 'persist'. Much like a chat room, the communication of a message board as interactive written discourse means that "a conversation can not be interrupted. [...] The nature of the chat room is, therefore, non-dyadic, unlike traditional speech. Chat room communications are not merely conversations between two participants, but conversations within a larger community" (Balfour 2004:9). The following postings belong to the same thread as in Example (4) but occur later in the thread. There are at least two different posters, and each turn is prefaced by a number:

- (5) 1. its funny that the hip-hop culture.. which everyone sees as a "ghetto" and poor culture.. has one of the most expensive 'uniforms'.. ive seen mens phat farm pants for over 200\$'s... in that price range you could be wearing a pretty nice suit and tie!
2. I second this in every way.
3. kinda scary that cats spending that much on clothes... I would never spend that much on clothing Fuck that I rather be able to buy a house for my kids and shit..
4. I find it amazing that parents today spend that much money on kids. I was never bought name brand shit. And when I was old enough to buy

my own.. the brand name shit most of the times fell apart faster then the no name shit. My first and only nikes only lasted 3 months.. wait till you all get kidsmine come at me with this shit al the timelike im gonna buy them \$100 pants they must be fucking crazy

5. Yeah... kat sound like he might be on to something...
6. I love hip hop.
7. that was kinda random...
8. yeah it was lol.

Examples (4) and (5) illustrate that this particular thread indeed flows much like a conversation, reducing the need for an opening and closing at every turn. However, due to the lack of opening or closing in the posting "I love hip hop" (posting 6) in Example (5), it is attended to and commented on (postings 7 and 8) for its lack of relevance to the topic of the thread, much like an interruption might be.

Herring (1996) explains the lack of openings in many e-mails as a result of headers including "to" and "subject" fields, making traditional openings redundant or superfluous. Message boards are categorized by topic and tend to include fairly distinct threads. An opening for the purpose of stating a topic, therefore, is unnecessary. Similarly, as threads represent persistent conversation, there is little need for closings. The corpus data suggest, however, that openings and closings for the purpose of general greetings and leave-taking similar to conversational openings and closings are common to general, non-threaded posts. Of the 464 comment-response postings, 283 (61%) included both an opening, and a closing, 116 (25%) included an opening or a closing, and 65 (14%) had neither an opening or a closing; each of these 65 was a response-posting from the webmaster addressing the previous comment. Such response-postings were common to this particular website, reducing the need and lowering the expectations for an opening or closing to each one.

The data suggest that the openings and closings of hip-hop postings, while structurally similar, are lexically different from those of traditional letters or CMC texts. Hip-hop specific openings and closings were identified as features of hip-hop message board postings via frequency and keyword lists. In the following sections, two frequent lexical signals of both openings and closings are presented, 'yo' and 'peace'.

3.1 Yo

'Yo' is one of a number of words and expressions closely associated with hip-hop. Hip-hop songs, for example, are peppered with 'yo' in their lyrics, reflecting its frequent use in the hip-hop vernacular. Picking up on the frequent usage and attention-calling function of 'yo', from 1988–1995 MTV aired a program devoted

to rap and hip-hop music called *Yo! MTV Raps*. In terms of recognition and use, 'yo' is now one of the most popular lexical items in hip-hop culture.

The frequency list compiled from the HMBC revealed 'yo' ranking at 91, occurring 166 times, or 0.16% of the 100,000 word corpus. In order to determine if this frequency was noteworthy, a keyword list was compiled on the corpus. Keywords in corpora are those words "whose frequency is unusually high in comparison with some norm" and therefore "characterize the text" (Scott 2004b: 19) under investigation. The 'norm' used for the comparison is known as a reference corpus; for this study, the reference corpus used was Text G from the FROWN files.³ This particular corpus was chosen due to both its size and content. At a total of 216,104 running words, it is approximately twice the size of the HMBC, providing a good basis for comparison in terms of quantity. Consisting of written American English, the FROWN corpus represents a potentially different kind of language than that of the HMBC, and thus significant keywords are more likely to be identified.

A keyword list compiled for the HMBC identified a total of 477 keywords, corresponding to approximately 5% of the total types in the corpus.⁴ Thus, a relatively high number of words in the corpus are key in that they occur unusually frequently when compared to the reference corpus. Since keywords may not be among the most frequent words in the corpus, a keyword list makes them salient in ways that a frequency list might not.

'Yo' ranked 28th in the keyword list, as shown in Table 1. The majority of the words in Table 1 represent function words and discourse markers characteristic of speech, and thus figuring as 'key' when compared to written texts. The keyword list suggests that 'yo' is a significant member of the hip-hop lexicon, which contains slang and taboo terms, as well. Unlike slang and taboo terms, however, 'yo' functions primarily as an opening to a posting. Of the 166 instances of 'yo', a total of 21 (13%) are used as alternative forms of 'you' or 'your', as in Example (6), while a further nine function as interjections, as in Example (7):

- (6) Go down to Fat Beats on Melrose and buy some b-boy videos and practice yo moves, don't forget to be original too.
- (7) Yo wuzzup this is Juan a former O towner but now i live down south in the Mia yo your site is phat! The breakers and the writers and even the design of the site is hot but yo the freestyles are kind weak and a lil wack but everything is phat! If u lookin for some flows email me and i'll hook it up but everything else on the site is phat! peace out.

3. <http://khnt.hit.uib.no/icame/manuals/frown/>

4. Keywords are types as well, so this figure is inflated in the same way that the number of total types is.

Table 1. Keyword list

N	Key word	Freq.	%
1	i	2 590	2.5
2	you	1 411	1.36
3	u	936	0.9
4	shit	474	0.46
5	ur	432	0.42
6	hip	374	0.36
7	hop	357	0.34
8	your	475	0.46
9	like	705	0.68
10	i'm	363	0.35
11	n	303	0.29
12	get	404	0.39
13	ya	266	0.26
14	dont	255	0.25
15	just	476	0.46
16	me	571	0.55
17	got	326	0.31
18	think	336	0.32
19	is	1 395	1.35
20	rap	221	0.21
21	i'll	215	0.21
22	fuck	203	0.2
23	if	552	0.53
24	im	195	0.19
25	can	419	0.4
26	b	206	0.2
27	know	327	0.32
28	yo	166	0.16

The remaining 136 instances of 'yo' function as openings, as in the opening of Example (7), above, and as the following examples illustrate:

- (8) yo dawg, always love the site, been a dope resource for me and other heads. always tell cats about it when they are looking for some good culture.
- (9) Yo wutup? Im just a beginner learnin how to pop. Im a gymnast so i can already do a lot of the power moves. I was just wonderin what crew Mr. Wiggles is from and what city its based out of. Thanks Yo.

In (9), 'yo' is even used in the closing. The poster self-identifies as a 'beginner', and thus may be over-using 'yo' as a known hip-hop lexeme. Nevertheless, the closing use of 'yo' reveals its possible dual function as both an opening and closing of a post, a feature which also characterizes 'peace'.

3.2 Peace

As in the case of 'yo', the word 'peace' was also identified as warranting a close analysis due to its association with hip-hop culture, as well as its inclusion in the keyword list. Within the hip-hop culture, 'peace' has specific semantics, meaning salutation or farewell. According to the website <http://www.rapdict.org>, the term 'peace' has been appropriated by the hip-hop community, which "updated the term to 'peace out'". The evolution of 'peace' into 'peace out' represents the preferred use of 'peace' as a closing. In the HMBC, 'peace' ranked at position 158 of the keyword list, with an occurrence of 63 times in the corpus (0.06%). While certainly not very frequent, the occurrences and usage of the term 'peace' in the HMBC nevertheless clearly reflect its dual function as an opening or closing. Of the 63 occurrences, only eight are used as an opening, as in Example (10):

- (10) Peace, checked out your site after hearing it on miami hip hop, the 2 of us were spot lited on the show so i wanted to see what you guys are all about, [...]

Example (11) shows how 'peace' is used as both an opening and closing in one and the same post, while Examples (12) and (13) illustrate the use of 'peace' and 'peace out' strictly as closings:

- (11) Peace, This is Mecca, a chicago based MC. [...] Peace, Mecca.
 (12) [...] If you have any questions, check out the forum message boards on the event. Peace and see you all in LA in Feb.
 (13) [...] If u lookin for some flows email me and i'll hook it up but everything else on the site is phat! peace out.

4. Slang and taboo terms

The data from the corpus suggest that the content of the message board postings is quite limited in scope, in terms of lexical density. While the type-token ratio 9.60 suggests a rather diverse vocabulary, the total number of types (9,822) is somewhat misleading. Many types are actually variants of one word, for example, 'please', 'pleez' and 'plizz', and thus when conflated, the total decreases. The total number of types could therefore be understood as much lower, suggesting a low lexical density for this corpus, which in turn indicates that "very few types occur very often". In fact, a content analysis of the postings of the HMBC reveals that one main focus is listening to or performing hip-hop music or texts, as illustrated in the following posting:

- (14) is it me, or are niggas on RB writing verses when they battle? =/... cause when im thinking of spittin, im thinking you actually sayign the shit out loud, mathcing sylables, having a flow to it.. not sitting there writing a diss essay lol.. maybe its just me and i need to change my style, but when i spit i spit so have people head movin wit mine na mean? i should prolly stick to cyphers haha

The content of this particular posting furthermore features a number of lexical items worth close examination. First, the terms 'battle' (a freestyle lyrical challenge with another contributor), 'spit' (to rap) and 'flow' (lyrical rhythm) are examples of hip-hop slang, the use of which functions as an in-group marker. Second, 'nigga' (nigger) and 'shit' number among a group of taboo words used particularly frequently in hip-hop discourse. Each of these lexical items are keywords in the HMBC.

4.1 Slang

The words 'battle', 'spit' and 'flow' together with their inflected variants each constitute under 1% of the corpus, but as slang, and thus words included in the hip-hop register (Berns & Schlobinski 2003), their frequency is high enough (or low enough in the reference corpus) to render them key. The use of 'battle', 'spit', and 'flow' in hip-hop discourse is illustrated in the following postings:

- (15) 1. yeah i might start soon cuz ive been busy accually MAKING this site and now that its finally settling down and its pretty active i can now relax and maybe battle.
 2. yeh g u shud ive never seen any of ur battles i look 4ward to seein ur battles n maybe if im feelin lucky i myt battle u.
 3. Yo fo real ? dat wud be kool G tu see u battlin u know een one or two of ure drops wen I first started on this site and it was real good so du ure thang G and show ery1 who runs thangs lol.
- (16) after reading through some of the nonsense posted on this site i had to join to try to bring common sense to the discussion. what rhyme has anyone ever heard 50 spit on that was hotter than most of the nigg@a he beefin with worst track. nas, jada, j, game and everyone else he beefin with are by far way superior lyricist. all 50 can spit is shoot this, 9 rounds that, and bonin groupies. ne one who says he is anywhere near the best probably gay and jsut want to be in his next candyshop video.
- (17) and i think game is the best rapper in the usa he is as hot as a motherfucker. i went and seen him live in glasgow and he was outstanding. i have seen 50 3 times live in glasgow and he was no were as good as game his flow is the best.

The examples of (15) illustrate the function of 'battle' as both a noun and verb, used to refer to the practice of challenging others to competitions in rapping. The use of 'spit' in Example (16) refers to the act of rapping, while 'flow' refers to the overall delivery of a rap. The use of these specific terms serves to establish the contributors of the postings as legitimate members of the hip-hop community.

Other keywords with specific hip-hop semantics include 'ill', 'tight', and 'sick' (positive, valued); 'peace' and 'safe' (salutation and/or farewell; see Example (21)); and 'holla' (recognize, acknowledge, communicate with):

- (18) i need a ill name or a name that fits me. i love graffiti.seeing my nigga who writes nerds influenced me. also many other graff heads i no like "coma.iw", "win.cas" and my nigga "nerds.NB". I try and try to get ill, but i don't know what the deal is. i think its because i cant find a name that fits me.
- (19) I'm not a huge snoop fan, he's kind of whack, but nate and warren g are tight.
- (20) Mic Club has a few weak beats on it (C Section, Drama A/T), but for the most part, Bis comes through with some absolutely sick rhymes over some equally sick beats (Master Thesis, Curriculum 101, Behind Enemy Rhymes, Allied Meta Forces....G Rap rips shit!!!!).
- (21) **SAFE**
U Peeps need to listen to lxxtra and then make ur judgments and trus me on this l!
- (22) somebody tell me how to find that song and album holla at yo boy tru miami soulja fan.

The use of slang in the message board postings reflects a familiarity with both linguistic and non-linguistic or cultural hip-hop practices, helping to identify each contributor as an in-group or community member. The use of these slang terms and other hip-hop jargon is a linguistic practice which, in the online context, enables the discursive construction of identity (Bucholtz 1999). As alternatives to mainstream language, hip-hop slang terms may even function as anti-language (Halliday 1976); that is, they contribute to the development of a hip-hop language which is incomprehensible to out-group members, thereby preventing access by out-group members to the hip-hop community. Nevertheless, the usage of both slang and taboo terms carries covert prestige (Trudgill 1972) and fulfils the same in-group member marking function.

4.2 Taboo terms

The high frequency of occurrence of 'nigga' establishes it as a keyword that characterizes at least this particular corpus if not hip-hop discourse in general. Along with 'shit' (0.5%), 'fuck' (0.45%), 'ass' (0.16%) and 'bitch' (0.13%), 'nigga'

(0.14%) ranks among the most frequent lexical (as opposed to functional) keywords in the HMBC. Such words are easily recognizable as belonging to a group of words collectively and commonly referred to as swear words, profanity or taboo terms (Beers Fägersten 2000, 2007). The simple identification of such taboo terms as key words in the HMBC is not enough to identify them as characteristic of hip-hop discourse. In fact, “[w]hile corpus data allows us to describe swearing in English, for example, it does not begin to provide an explanation for anything that we see within the corpus” (McEnery 2006:4). Based on examples of taboo terms in context, it is argued that, unlike the slang terms above, the use of swear words functions to marginalize hip-hop culture and the hip-hop community by virtue of their recognizability as taboo terms:

- (23) all u niggaz lost yall mind sayin tha black album is wack. if ya think tha black is wack den ya aint really listen to it. dat shit is da hottest album of tha year. yall need to sit back and listen to dat shit cuz dat shit is hot.
- (24) now its just a bunch of ignant cats playin into the stereo types of hiphop adn the black race but i wont comment anymore cause its not my place too comment on black issues. But if i was black fuck id have enough shit too say about what these stupid ass crunk rappers are sayin.
- (25) i just dont like it when i see a white kid or any race actin all hard and shit cuz their wearing fubu and rocawear and BX and shit that when i think its time to give an ass beatin to any race....actin all hard..BITCH PLEASE!!!

The most noteworthy feature of Examples (23)–(25) is the recurrent use of swear words within single postings. The practice of using taboo terms is thereby made much more salient, suggesting that such linguistic behavior is in fact characteristic of the members of the hip-hop community.

5. Verbal art

Verbal art, according to Joel Sherzer (1987:296), is:

discourse which creates, recreates, modifies, and fine tunes both culture and language and their intersection, and it is especially in verbally artistic discourse such as poetry, magic, verbal duelling, and political rhetoric that the potentials and resources provided by grammar, as well as cultural meanings and symbols, are exploited to the fullest and the essence of language-culture relations becomes salient.

In her discussion of verbal art and performance, Johnstone (2002:220) addresses the aesthetic aspects of discourse, claiming that “humans attend to how discourse sounds and looks as well as to what it refers to and what it is meant to accomplish”. She (2002:42) also says that “[a]s people construct discourse, they draw on

the resources provided by language and on the resources provided by culture". The postings of the HMBC can be considered verbal art due to the creativity resulting from an exploitation of available resources of CMC, namely the keyboard. In this section, verbal art in the form of alternative orthography including the use of numbers and keyboard symbols is presented as a common practice in computer-mediated hip-hop discourse.

Wilkins (1991) points out that asynchronous computer-mediated communication and spoken discourse have in common both the occurrence of second person pronouns as well as linguistic creativity. The HMBC supports this finding, as it includes many postings in which contributors seek to make contact with other members of the community, discuss opinions or ask to be acknowledged (Beers Fägersten 2006). There is also a high frequency of performance of verbal art. Although the HMBC is composed of written English, it can be argued that the content reveals features of spoken, conversational English. A noticeable characteristic of the content of the hip-hop message board postings is the explicit acknowledgement of interaction with others via the use of second-person pronouns. The pronoun 'you' ranks as the second most frequent keyword of the corpus and sixth overall with a frequency of 1,411 or 1.36%. To demonstrate that this frequency is particularly high within a written-language corpus, a frequency analysis of the HMBC has been compared to frequency analyses of other corpora, including the Lancaster-Oslo/Bergen Corpus (LOB; 1 million words; British English) Brown Corpus (1 million words; American English) and the British National Corpus (BNC; 100 million words; British English). Table 2 shows the ten most

Table 2. Ten most frequent words (types) by corpus

	HMBC	LOB ⁵	Brown ⁶	BNC ⁷
1.	the	the	the	the
2.	i/I	of	of	of
3.	and	and	and	and
4.	to	to	to	a
5.	a	a	a	in
6.	you	in	in	to
7.	is	that	that	it
8.	of	is	is	is
9.	that	was	was	to
10.	in	it	he	was

5. <http://alt-usage-english.org/excerpts/fxcommon.html>

6. <http://www.edict.com.hk/lexiconindex/frequencylists/words2000.htm>

7. <ftp://ftp.itri.bton.ac.uk/bnc/all.num.o5>

frequent words across each corpus. The types in Table 2 have not been conflated, thus alternative forms of 'you' have not been included.

It is clear that even in the genre- and register-specific HMBC, function words occur most frequently. Where the HMBC diverges is in the frequency of the first- and second-person pronouns 'I' and 'you'. The difference should be considered in terms of a low frequency of these pronouns in the comparison corpora versus a high frequency in the HMBC; 90% of the BNC and the entire LOB and Brown corpora are composed of written texts, where first- and second-person pronouns are less frequent (Biber 1988; Yates 1996).

A keyword link analysis of the corpus reveals 'you' to be the most frequently linked word, with 444 links to other keywords. A keyword link analysis reveals "which keywords are most closely related to a given keyword" (Scott 2004b). In other words, 'you' is the keyword which most often occurs in clusters or collocations with other keywords. When different forms of 'you' are accounted for, the number of keyword clusters increases. Although 'you' occurs 1,411 times in the corpus, corresponding to a frequency of 1.36%, when other forms such as 'ya', 'y'all', 'your', 'you're', 'u', and 'ur' are included in a frequency count, the total jumps to 3,715, or 3.67%, making the superordinate second-person pronoun the most frequent corpus-wide type. In other words, there is a clear tendency among contributors to explicitly acknowledge and appeal to interlocutors through the use of the second-person pronoun.

In the HMBC, the most frequent cluster is, not surprisingly, 'hip hop', occurring 274 times. The second most frequent cluster, however, is 'if you', with 190 occurrences. This cluster usually occurs in contexts where the contributor is seeking contact in order to appeal for assistance, as in the following examples:

- (26) also my friend told me that it's better to just buy a couple of break dancing tapes instead of going off to a school to learn. I'd like to know if you agree or if you have another opinion.
- (27) Whats killing me is some of these wack as south rappers. Ryhming like niggas were ryhmin in the eighties back in the bronx, And slingin that shit like they just created a new style. Holla if you here me.
- (28) Word up! Nigga from New York claim there hoods are the hardest same thing with cali niggaz. Dog, theses are all huge cities where you can move somewhere else in the same city and be straight. What the hell do y'all know bout them little muder towns like Gary, Indiana Flint, Michian Little Rock, amongst others. Don't praise the dirt that goes on in the hood. If u have another opinion we can discuss it in a civil manner.
- (29) YOU'RE FUCKED UP ON ECSTACY. U MUST BE DRUGGIN' IF U FUCKIN' THINK U CAN MESS WITH ME.

Postings (28) and (29) are considerably more aggressive than postings (26) and (27), partly due to the 'if u' clusters functioning as challenges to the addressee. The increased aggressiveness corresponds to the shift from the use of standard 'you' to non-standard 'u'. Much like threats, such challenges imply that the poster has the ability, social power or social status to question the beliefs or practices of another. Part of hip-hop culture is the practice of asserting your identity in terms of knowledge of hip-hop or talent in battling or rapping (George 1999; McLeod 1999; Newman 2001). Just as it is a lack of knowledge which motivated the postings in Examples (26) and (27), it is an assertion of knowledge which characterizes Examples (28) and (29). The polarization of status and power as well as their corresponding assertions of hip-hop identity are encoded in the alternate forms of 'you/u'. The use of both standard and non-standard orthography allows the posters to discursively construct their identities through the form-based strategy of performance of verbal art.

5.1 Non-standard orthography

The example postings thus far included give an indication of the extent of linguistic manipulation involved in performing a hip-hop identity. In Example (3), for instance, almost every word of the posting is either specific to the hip-hop genre (e.g., 'nikkas breakin iight') or written in an alternative manner (e.g., 'do dat fo ur boi'). The corpus data further suggest that the verbal art of hip-hop discourse is reflected mainly through a community-wide and systematic use of alternative, non-standard varieties of orthography which permeate nearly all word types. Of the ten most frequent words listed in Table 2, alternative spellings for seven of them were found in the corpus. Both sets of words, along with the corresponding frequency percentages, are presented in Table 3. Only the most frequent alternative for each standard form is presented in Table 3.

Table 3. The ten most frequent words and their alternative forms (if any)

	Standard	% of corpus	Alternative	% of corpus
1.	the	3.36	da	0.13
2.	i/I	2.53	–	–
3.	and	2.28	n	0.29
4.	to	2.04	2	0.17
5.	a	2.02	–	–
6.	you	1.38	u	1.33
7.	is	1.36	iz	0.02
8.	of	1.34	a	0.05
9.	that	1.14	dat	0.07
10.	in	1.11	–	–

The HMBC consists of written message board entries, and thus neither directly represents oral language nor does it feature prosodic mark-up. Nevertheless, the use of non-standard orthography often encourages an interpretation of the postings as oral language, giving the impression that the contributions are written in such a way as to indicate a particular pronunciation. The alternative forms 'da', 'n', 'a', and 'dat' are orthographic representations of the standard forms as they would be phonetically realized in speech. In word-initial position, the dental [ð], represented orthographically as 'th', is phonetically realized as the voiced, alveolar, plosive [d]; this is even a feature of African American Vernacular English (AAVE), which is associated with hip-hop culture (Feldman 2002; Rickford 1999; Rickford 2004). 'And' and 'of' are often reduced in informal speech such that 'and' becomes a syllabified [n] and 'of' becomes the lax vowel [ə].

In contrast to these four forms, the other three alternative forms '2', 'u' and 'iz' do not represent non-standard pronunciations. They do, however, further illustrate verbal art in computer-mediated discourse since each alternative spelling calls attention to the form of the message. The use of small 'i' for the first-person pronoun could also be considered an alternative to the standard; however, a general absence of capitalization has in fact become the standard in computer-mediated communication (Crystal 2001). The percentages in Table 3 clearly indicate that the standard forms are more common, but considering the corpus size, the frequency of the alternative forms is remarkable, particularly when one also considers that the use of each alternative may entail a deliberate effort to avoid writing the standard form. Herring (1996) points out the economical use of special characters and acronyms in computer-mediated communication. It is true that each alternative form is shorter than the standard, but one could also argue that their use is potentially more time-consuming due to the potential effort involved in *not* typing the standard form. Admittedly, this evaluation reflects an out-group member bias. For a community member or a seasoned contributor to hip-hop message boards, saying, thinking, and even typing, for example, 'dat' instead of 'that' may be effortless; after all, this form is a feature of one's hip-hop identity. The earlier discussion of the 'if you' cluster supports precisely this argument, as the alternative 'if u' clusters are shown to occur primarily in postings by contributors asserting their hip-hop identities based on their knowledge and expertise in – and thus familiarity with – hip-hop culture.

5.2 Use of numbers

Although the HMBC had been edited to remove dates and websites, the first frequency list compiled revealed a conspicuously high frequency of numbers, particularly single digit numbers, i.e., 0–9. A manual, qualitative investigation of the corpus revealed the frequent use of numbers as alternatives to letters, phonolog-

ical strings, and morphemes in the hip-hop message board postings. Because of the type-written form of message board postings, the physical similarity of some keyboard numbers and letters can be exploited, as in the following examples of substitution:

'9' for 'g', especially in 'ni99a'

(30) this cat wasn't even a street ni99a's he just new street ni99a's and told they stories, wich is cool but dont front like that's yo live.

(31) Ni99a's i had to be on the block I much rather be in the board room Bitches, word!

'0' for 'o'

(32) i love the page so pr0ps to u.

'5' for 's', '4' for 'A'

(33) 54F£

Unlike the alternative orthography of several of the most frequent words, these number substitutions do not represent non-standard pronunciations. Furthermore, as the numbers only replace single letters, the alternative forms are legible and quite easily comprehensible.

In the following examples, however, numbers are used to replace phonological strings and entire morphemes, which encourages and sometimes requires pronunciation, as the standard form is not always immediately recognizable from the altered, type-written form.

'1' for 'one'

(34) i was readin a boys source magazine n da black ppl wer sayin stuf like 'how can we call eminem a racist wen we disrespect ourselves by calin each other niggas'. iaint even 2 sure bout dis so if ne1 can xplain.

(35) i got some1 makin our sig its gonna have female gangstaz in front den in da background its gonna have 50 cent chingy and other people iight tell me wat yall think.

'2' for 'to-', 'to' or 'too'

(36) i run da streets 2DAY- get on ur knees 2 pray- u tryin' 2 be hard, but it doesn't increase da rage

(37) u aint no rapper plz all your lyrics is kept in a guitar-case fakest thinker huh? 2bad the world is mine sorry scar-face

'4' for free morpheme 'for' or phonological string [fɔr]

- (38) yeh g u shud ive never seen any of ur battles i look 4ward to seein ur battles n maybe if im feelin lucky i myt battle u.
 (39) just u wait kuz the U4iK is gonna tackle his bitch ass.

'8' for phonological string [eIt]

- (40) get yah daym facts str8 noob before you correct me again! i know my sheit.
 (41) soul tld u y blks h8 whites, racism dnt change and neva will, its still out dere no matter how da government trys 2 hide it.

In many of the above postings, further examples of non-standard orthography can be identified. The ability to determine the extent of usage and systematicity of alternative forms is a distinct advantage of a corpus study. A frequency list, for example, has revealed the most common words of the HMBC, as well as the systematic usage of numbers to replace letters, phonological strings and morphemes. Using WordSmith, it is also possible to view a corpus as alphabetical or reverse-sort lists. Each list facilitates further identification of recurrent, systematic uses of non-standard orthography in that similar forms are grouped together, for example, 'u', 'u'll', and 'ur'. The reverse-sort revealed a curiously large amount of words ending in the letters 'a' and 'z', encouraging further investigation and revealing a systematic usage of non-standard orthography for specific (morpho-)phonology.

5.3 Word-final 'a'

Excluding proper names and other words that end in -a in standard orthography, the total number of 'a'-final tokens (not including plurals) in the HMBC is 1,307, distributed over 139 types, corresponding to 1.28% and 1.37% of the total tokens and types, respectively, in the corpus. The non-standard orthography featuring final 'a' can be categorized according to the word it substitutes for (Examples (42)–(45)), or the sound string it is meant to represent in speech (Examples (46)–(48)). In general, final 'a' reflects (morpho-)phonemic reduction:

–a for 'have'

- (42) he shoulda neva gotten control of TS cuz now all thats left is armageddon and tony sunshine

–a for 'of'

- (43) i kno a buncha y'all faggots, ur so hungry, "u act BIGGA"

–a for ‘to’

- (44) if u wanna speak yo raise ya hand
cuz u dont wanna see the buckin if u disrupt the man

–a for [o]

- (45) LOL..NOW FELLAS, children tend to have wild imaginations. maybe he doesnt get enough positive attention at home..!

‘ma’ for [maI] in ‘my’

- (46) i write to improve ma skillz

In the discussion of the most frequent words of the corpus, ‘u’ was identified as an alternative form of ‘you’. Furthermore, in postings (3), (15), (21), (36) and (38), examples of an alternation between ‘your/you’re/ur’ can also be seen. The reverse-sort list revealed another non-standard variant used by the message board contributors, namely ‘ya’:

- (47) Ill never stop spitin, till ya run outa the shit that was prewritten
(48) stab u in ya bladder, and drown u in Piss Puddles

The use of both the ‘u’ and ‘ya’ forms in Example (48) is particularly illustrative of the different phonetics they are each intended to represent. The overall phonology of both postings (47) and (48) is particularly important to the contributor, since these postings are actually part of rap lyrics posted on the message board as part of a battle.

Many of the ‘a’-final tokens are the forms ‘da’ and ‘tha’, an alternative spelling for ‘the’. Unlike ‘da’, there is no obvious correspondence in pronunciation of the non-standard spelling of ‘tha’. However, Example (49) suggests that this variant may, in fact, be phonologically motivated:

- (49) i aint from tha D but i live here, in grosse point, where da rich mufuckas at, lol

This posting shows that the contributor indeed has both variants in his/her repertoire. The use of ‘tha D’ may be to avoid the alliteration which would result from ‘da D’ (Detroit), but still achieve poetic discourse with non-standard orthography. There is a switch to ‘da’ later in the posting, which further suggests that the earlier use of ‘tha’ is due to its phonological environment.

Additional final ‘a’ tokens include three different examples of elision – ‘hella’, ‘ima’ and ‘ma’ – where syllables or, in the case of multi-word expressions, entire words are omitted. In posting (50) it can be seen that, through conversion, the form ‘hella’ functions as an intensifying adverb, much like ‘really’ or ‘very’:

'hella' for 'hell of a'

- (50) Yeah he seems hella hungry on that joint.

Many examples of AAVE can be found in the language of hip-hop (Cutler 1999; Feldman 2002; Rickford 1999; Rickford 2004). Rickford (2004) identifies 'ama' as a feature of AAVE; the corpus includes the variants 'im'a', 'imma', and 'i'mma'.

'Ima' for 'I am going to'

- (51) Thinkin u a thug, why dont u bust slugs and Humor Me Cuz ima think u a poser till bullets rip thru my Computer screen

It has been claimed that the word 'motherfucker' (or 'mother fucker') is most frequently used by African Americans, especially males (Berger 1970; Hughes 1998). Although there is evidence of a wider social distribution of its use (Beers Fägersten 2000), the association with African Americans entails an association with AAVE, which is, in turn, associated with hip-hop. Nevertheless, the variants of 'motherfucker' are relatively infrequent and the anonymous nature of message boards makes it difficult to ascertain if its use is associated with a particular race or gender.

'ma' for 'mother' or 'motherfucker'

- (52) stick to the music ma fuckas
 (53) yeh ure journal u got sum deep stuff up in that ma ya know

The balance of the 'a'-final tokens reveals a systematic use of non-standard orthography to represent the sound [əɹ]. In the majority of cases, the final 'a' is a direct substitution for the letters 'er', corresponding to a nominal marker as in posting (54) or the comparative adjectival marker, as in posting (55). Final 'a' can also be found as a replacement for non-morphemic word-final 'er' in analogous environments, such as in Examples (56) and (57):

- (54) i thought it wuz hilarious dat a rappa would write about sum gawd dayum no tooth bitchez...
 (55) and shit still does happen 2 us. maybe not in america but in england it does. u got da national front and shit and my dad used 2 get jumped by white people wen he was yunga.
 (56) n another thing,black dudes dont wanna holla at white girls when black girls are around,but as soon as the black girls leave,they try hollerin..
 (57) I waz neva really feelin Obie lyrically but if itz anythin like got some teeth itz cool to thro on when you goin out n shit

Other examples of final 'a' do not constitute a direct substitution for final 'er', but rather for the word final phones [əɹ] or [øɹ]:

'fire'

- (58) king of the spit pit, i spit the fiya shit

'for'

- (59) oh i cant wait fa someone to hear this or see it. Ya get a hype feelin and from there ya feel at ya peak and just hold on to it

'sure'

- (60) 4sha!!!! peace to B.I.G

5.4 Word-final 'z'

A reverse sort revealed a total number of 533 words ending in -z, distributed over a total of 150 types. The final 'a' tokens effect a non-standard orthography intended to reflect pronunciation, encouraging readers to receive and process the text as if it were spoken. In contrast, final 'z' does not seem to elicit an alternative pronunciation. Like the morpheme /s/, 'z' is used to mark plurals, third person singular inflections, and possessives. The data suggest that final 'z' is used as a non-standard orthographic feature to reflect standard phonology, that is, when the phonological environment of the morpheme /s/ results in voicing, yielding word-final [z]. Posting (61) includes examples of plural, third person singular and possessive morpheme /s/ realized as 'z':

- (61) i ain't feelin obie'z shyt.....his songz waz whack n i wasn't feelin hiz cd...dat nicca iz jus an "AD"....jus gettin ppl'z attention 4 a sec.

The following examples illustrate that final 'z' is also used to substitute for the inflectional morpheme /s/, even when the phonological environment would not cause voicing:

Plural

- (62) I love Pac and Dear Mama is on his greatest hitz cuz it waz one of his greatest hitz ..Classic

Third person singular verb agreement

- (63) if you think pac'z overrated thatz your opinion

Final 'z' also appears in words ending in [z] or 's', regardless of the phonology:

'cuz', 'coz', 'kuz', 'becuz', 'becoz' for 'because'

- (64) Cuz he has a few slower beat songs which manz can kick back to.

'plz', 'plizz', 'pliz' for 'please'

(65) u aint no rapper plz all your lyrics is kept in a guitar-case

'asz', 'azz' for 'ass'

(66) I can still "Pop" my azz off at this age... I was Known as "Mr.Tic" becuz I could strobe my whole body like 3-D.

Furthermore, the use of final 'z' has been extended to words with no motivating phonology, indicating a trend towards word-final usage:

(67) Crooked I is heavy but know I wants to know, Ras Kass is on another level. If they all move in unison! the west can be what it can be "GFunkedcrazymutha-fuckers"

ANYWAYZ

LATERZ

The alternative spellings for both frequent function words as well as less frequent lexical items suggests that a general feature of the postings of the HMBC is the inclusion of verbal art, where verbal art is the creative exploitation of computer-mediated resources to achieve poetic and aesthetic content and form. The anonymity of the postings due to aliases or even lack thereof does not allow for a reliable survey of the extent of inter-poster variation, but the number of examples encourages the claim that the variation is not due to repeated entries from a single poster, but rather individual entries of many posters, which in turn offers support for the claim that performance of verbal art is common to the discursive construction of identity as a member of the hip-hop community.

5.5 Use of special characters

In Example (33), the use of numbers to substitute for letters is illustrated: 54F£. The use of the British monetary symbol '£' for capital 'E' suggests that the postings can be attributed to predominantly British contributors, and indeed the fuller contexts of postings such as (33) corroborate this conclusion, although the corpus data cannot confirm poster identity. Regardless of the origin, such postings also illustrate how contributors exploit the interface of computer-mediated communication, the keyboard, to accomplish yet another kind of non-standard orthography by substituting symbols or special characters for letters. This substitution occurs in words that are potentially offensive, suggesting self-censorship, possibly to avoid filtering (Crystal 2001; Dingwell 2004) which could result in non-publication of the posting:

- (68) i don't know about that. u put too many dumb n!gg@s in a room and some dumb n!gg@ shit is going to happen. somebody is about to get shot over this shit soon.
- (69) gotta give her sum credit, i mean she writez all her sh!t, not alot of rnb b!tchz do that
- (70) IF YOU PULL OUT YO WALLET YOU'LL GET SHOT UP AND F#CK IN THE A\$\$ WITH A PLUNGER BY THE PIGS

Other examples of non-standard orthography specific to the keyboard interface include the usage of capital letters. The conspicuous use of capital letters for some or all parts of chats, e-mails or other forms of computer-mediated communication has been conventionalized to be considered as shouting (Danet et al. 1997), and is thus cautioned against as an inflammatory practice (Crystal 2001). However, a total of 51 postings (3.4%) in the HMBC are written entirely in capital letters, as in posting (70). Such postings may not be intended as or, more importantly, even considered shouting, since the persistent use of capitalization throughout an entire posting neutralizes the shouting effect:

- (71) I AGREE, 50 IS HOT BUT NAS AND JADKISS ARE BETTER. NAS GAVE JAY-Z A HELL. SOME PEOPLE THINK HE WON. WHATS MAKES 50 THINK THAT HE WANTS SOME OF NAS.AND JADKISS IS WAITING FOR SOMEONE TO SAY HIS NAME ON A TRACK. JADKISS CAN DEFINTY GET WITH HIM. FAT JOE IS A SESON VETERN. 50 NEEDS TO MAKE FRIENDS AND NOT EMEMYS BEFORE SOMEONE ENDS HIS LIFE OR HIS CAREER.

Only when the use of lower-case letters is established as the norm can intermittent capitalization be attributed paralinguistic meaning:

- (72) Read what i said you DUMB OLD FUCK, i was pointing out how stupid you are.

Another variety of non-standard orthography in the hip-hop message board postings involves the use of alternating lower-case and upper-case letters. The data suggest that this practice is not indicative of any vernacular pronunciation or paralinguistic effect, but rather fulfils a purely poetic function, such as the nickname of the poster of the following example:

- (73) AbSoLuTeBbOy

6. Conclusion

Many studies of hip-hop culture and hip-hop identity have focussed on the content of hip-hop music and discourses (Berns & Schlobinski 2003; George 1999; McLeod 1999; Newman 2001; Rose 1994), attending to the content but failing to address the form. The examples presented in this chapter provide ample argument for considering the *how* as well as the *what* of hip-hop discourse. Keyword and word frequency analyses of the HMBC have revealed the discursive construction of identity to be a function of the lexical content of postings; alphabetical and reverse-sort lists have also helped to identify the systematic use of alternative forms or verbal art. Slang terms were identified as keywords used by posters to mark their in-group membership, and to mark the hip-hop community as a counter-culture. Each status is further established by the frequent use of taboo words, which, because of the general, society-wide recognizability of such terms as informal, non-standard and potentially offensive (Beers Fägersten 2000), marginalizes the community through less encrypted means than hip-hop jargon.

Throughout the corpus and in many of the above examples, however, content is dominated by form, drawing attention from what the posting is about to how the posting looks. The juxtaposition of the following extracts from postings illustrates an increasing saliency of form over content:

- a. As a hip hop person, I like to always educate myself on facts and to learn the history of something in my interest, and this website is doing that.
- b. i am white and porto rican but yu cant tell i look striat up with but i act hood not black not fuckn wiggash i at hood cuz that is where i am from
- c. ey foo.....show meeh sum pic of yall nikka's breakin iight homie.....do dat fo ur boii b joe frum Under Rated Breakaz

Implicit or explicit membership in the hip-hop community characterizes the content of each posting, but the forms of (b) and, to a greater extent, (c) corroborate the content by showing a familiarity with the verbal art practices of hip-hop. I have argued here that, in the context of online hip-hop discourse, 'as a hip hop person' does not serve to identify the poster as a member of the hip-hop community (despite the claimed affiliation with a breakdancing group) nearly as much as 'i b joe frum Under Rated Breakaz' does. The postings reveal that both the content and the form of discourse contribute to the discursive construction of identity. It is these two elements together which create a hip-hop persona.

The corpus data further suggest that such verbal art is frequent and, to a great extent, conventionalized in terms of wide-spread practice based on frequency of examples. Non-standard orthography including usage and integration of numbers and symbols would seem to be the theme resulting in variation of form among the

postings. The anonymity associated with the message board forum allows for the possibility that only a few contributors, or in the extreme even one contributor, is responsible for the variation between standard and non-standard orthography or between formal and informal styles. It is argued, however, that the number of postings in this corpus created from five different websites suggests a greater dispersal. It is furthermore argued that only through familiarity with and use of such verbal art practices can members of the hip-hop community discursively construct and recognize a valid hip-hop identity.

The message board medium affords interlocutors the time to produce and process language, and thus awards them the opportunity to exploit fully their linguistic resources, the resources of computer-mediated communication. Posters are not only able, but are expected to showcase their linguistic talents, rhythmic abilities, and familiarity with the practices of hip-hop and the conventions of computer-mediated communication to be identified and accepted as hip-hop. Consequently, the hip-hop message board postings reveal a deliberate exploitation of the language qualities of this medium. The corpus shows that members of the hip-hop community have adapted to the internet medium, indeed are embracing it and taking full advantage of the interface systematically to construct and assert their individual identities, and to establish community practices. In light of the oral and artistic traditions of hip-hop culture, internet message boards allow for creative codification of speech practices and therefore represent an ideal forum for members of the hip-hop community to discursively construct their identities.

References

- Allwood, J. 2000. An activity based approach to pragmatics. In *Abduction, Belief and Context in Dialogue: Studies in computational pragmatics*, H. Bunt & William Black (eds), 47–80. Amsterdam: John Benjamins.
- Androutsopoulos, J. 2006. Introduction: Sociolinguistics and computer-mediated communication. *Journal of Sociolinguistics* 10(4): 419–438.
- Balfour, J. 2004. *The third medium: Sociolinguistics and chat rooms*. <http://a.parsons.edu/~julia/thesis/oct13.pdf>
- Bechar-Israeli, H. 1995. From 'Bonehead' to 'cLoNehEAd': Nicknames, play and identity on internet relay chat. *Journal of Computer-Mediated Communication* 1(2).
- Beers Fägersten, K. 2000. *A Descriptive Analysis of the Social Functions of Swearing in American English*. Ms.
- Beers Fägersten, K. 2006. The discursive construction of identity in an Internet hip-hop community. *Revista Alicantina de Estudios Ingleses: Special Issue on Linguistics and Media Discourse* 19.
- Beers Fägersten, K. 2007. A sociolinguistic analysis of swear word offensiveness. *Saarland Working Papers in Linguistics*. <http://scidok.sulb.uni-saarland.de/sulb/portal/swpl/>, retrieved 15 March 2006.

- Berger, A. 1970. Swearing and society. *ETC: A Review of General Semantics* 30: 283–286.
- Berns, J. & Schlobinski, P. 2003. Constructions of identity in German hip-hop culture. In *Discourse Constructions of Youth Identities*, J. Androutsopoulos & A. Georgakopoulou (eds), 197–219. Amsterdam: John Benjamins.
- Biber, D. 1988. *Variation Across Speech and Writing*. Cambridge: CUP.
- Bucholtz, M. 1999. 'Why be normal?' Language and identity practices in a community of nerd girls. *Language in Society* 28: 203–223.
- Collopy, M. & Belmore, N. 1996. Electronic language: A new variety of English. In *Computer-Mediated Communication: Linguistic, social and cross-cultural perspectives*, S. C. Herring (ed.), 13–28. Amsterdam: John Benjamins.
- Crystal, D. 2001. *Language and the Internet*. Cambridge: CUP.
- Cutler, C. 1999. Yorkville crossing: White teens, hip hop and African American English. *Journal of Sociolinguistics* 3(4): 428–442.
- Danet, B., Ruedenberg-Wright, L. & Rosenbaum-Tamari, Y. 1997. Hmmm... where's that smoke coming from? Writing, play and performance on Internet relay chat. *Journal of Computer-Mediated Communication* 2(4) <http://jcmc.huji.ac.il/jcmc/vol2/issue4/danet.html>, retrieved 15 March 2006.
- Danet, B. 2002. *The language of E-mail*. European Union Summer School, University of Rome. <http://www.europhd.psi.uniroma1.it>, retrieved 15 March 2006.
- Dingwell, H. 2004. *Exploring web chat: Are Internet chat sites a form of community?* www.infinitecreations.net, retrieved 15 March 2006.
- Eiler, M. A., & Victor, D. 1988. Genre and function in the Italian and U.S. business letter. *Proceedings of the Sixth Annual Conference on Languages and Communications for World Business and the Professions*. Ann Arbor MI.
- Erickson, T. 1999. Persistent conversation: An introduction. *Journal of Computer-Mediated Communication* 4(4): <http://www.ascusc.org/jcmc/vol4/issue4/erickson.html>.
- Feldman, M. 2002. *African American Vernacular English in the lyrics of African American popular music*. <http://www.swarthmore.edu/SocSci/Linguistics/papers/2002/mattfeldman.pdf>
- Ferrera, K., H. Brunner & G. Whittemore. 1991. Interactive written discourse as an emergent register. *Written Communication* 8(1): 8–34.
- George, N. 1999. *Hip Hop America*. New York NY: Penguin.
- Halliday, M. 1976. Anti-languages. *American Anthropologist* 78(3): 570–584.
- Hård af Segerstad, Y. 2002. *Use and Adaptation of Written Language to the Conditions of Computer-Mediated Communication*. Göteborg: Göteborg University.
- Herring, S. 1996. *Computer-mediated Communication: Linguistics, social and cross-cultural perspectives*. Amsterdam: John Benjamins.
- Herring, S. 2001. Computer-mediated Discourse. In *The Handbook of Discourse Analysis*, D. Schiffrin, D. Tannen & H. Hamilton (eds), 612–34. Malden MA: Blackwell.
- Hughes, G. 1998. *Swearing: A social history of foul language, oaths and profanity in English*. London: Penguin.
- Johnstone, B. 2002. *Discourse analysis*. Malden MA: Blackwell.
- McEnery, T. 2006. *Swearing in English*. London: Routledge.
- McLeod, K. 1999. Authenticity within hip-hop and other cultures threatened with assimilation. *Journal of Communication* 39(4): 134–150.
- Nakamura, L. 1995. Race in/for cyberspace: Identity tourism and racial passing on the internet. In *Cyberreader*, V. J. Vitanza (ed.), 442–453. Needham Heights MA.

- Newman, M. 2001. 'I represent me': Identity construction in a teenage rap crew. *Texas Linguistic Forum, Proceedings from the Ninth Annual Symposium about Language and Society* 44(2): 388–400.
- Ong, W. J. 1982. *Orality and Literacy: Technologizing the word*. London: Routledge.
- Rickford, J. R. 1999. *African American Vernacular English: Features, evolution, educational implications*. Malden MA: Blackwell.
- Rickford, J. R. 2004. What is Ebonics? (African American Vernacular English). *Linguistic Society of America* on-line, https://lsadc.org/info/pdf_files/Ebonics.pdf
- Rose, T. 1994. *Black Noise: Rap music and black culture in contemporary America*. Hanover NH: Wesleyan University Press.
- Schegloff, E. 1972. Sequencing in conversational openings. In *Directions in Sociolinguistics: the Ethnomethodology of Communication*, J. Gumperz & D. Hymes (eds), 346–380. New York NY: Holt Rinehart and Winston.
- Scott, M. 2004a. *WordSmith Tools* (Version 4). Oxford: OUP.
- Scott, M., 2004b. *WordSmith Tools Help*. Oxford: OUP.
- Sherzer, J. 1987. A discourse-centered approach to language and culture. *American Anthropologist* 89: 295–305.
- Stone, A. R. 1991. Will the real body please stand up? Boundary stories about virtual cultures. In *Cyberspace: First steps*, M. Benedikt (ed.), 81–118. Cambridge MA: The MIT Press.
- Taboada, M. 2004. The genre structure of bulletin board messages. *TEXT Technology* 13(2): 55–82.
- Takahashi, J. 2003. *Do we talk (or write?) differently over the Net? A lexical enquiry into 'a' Net-EN*. <http://korpus.dsl.dk/cl2003/cdrom/papers/takahashi.pdf>.
- Trudgill, P. 1972. Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society* 1:179–95.
- Turkle, S. 1995. *Life on the Screen: Identity in the age of the internet*. New York NY: Simon & Schuster.
- Van Gelder, L. 1990. The strange case of the electronic lover. In *Talking to Strangers: Mediated therapeutic communication*, G. Gumpert & S. Fish (eds), 128–142. Norwood NJ: Ablex.
- Wallace, P. 1999. *The Psychology of the Internet*. Cambridge: CUP.
- Warschauer, M. 2000. Language, identity, and the internet. In *Race in Cyberspace*, B. Kolko, L. Nakamura & G. Rodman (eds), 151–170. London: Routledge.
- Wilkins, H. 1991. Computer talk. Long-distance conversation by computer. *Written Communication* 8(1): 56–78.
- Yates, S. 2000. Computer mediated communication: The future of the letter? In *Letter Writing as a Social Practice*, D. Barton & N. Hall (eds), 233–51. Amsterdam: John Benjamins.
- Yates, S. J. 1996. Oral and written linguistic aspects of computer conferencing. In *Computer-Mediated Communication: Linguistic, social and cross-cultural perspectives*, S. C. Herring (ed.), 29–46. Amsterdam: John Benjamins.
- Yates, J. & Orlikowski, W. 1992. Genres of organisational communication: An approach to studying communication and media. *The Academy of Management Review* 17(2): 299–326.

Appendix A

The corpus consists of message board entries collected from the following websites:

<http://p081.ezboard.com>
<http://www.hiphopsite.com>
<http://www.jam2dis.com/hiphopboard>
<http://www.thugz-network-board.tk>
<http://www.underworldhiphop.com>

Each website was last accessed for data collection on 22 November 2004.