

Veera Riikonen

**THE USE OF AFRICAN AMERICAN ENGLISH
NEGATION PATTERNS IN CORPUS OF
GLOBAL WEB-BASED ENGLISH**

Faculty of Information Technology and Communication Sciences
Bachelor's Thesis
October 2022

ABSTRAKTI

Veera Riikonen: The Use of African American English Negation Patterns in Corpus of Global Web-Based English

Kandidaatin tutkielma

Tampereen yliopisto

Kielten kandidaattiohjelma, englannin opintosuunta

Lokakuu 2022

Tämän tutkielman tarkoituksena on kartoittaa afrikkalaisamerikkalaiselle englannille tyypillisten kieltomuotojen käyttöä internetissä tapahtuvassa kielenkäytössä. Afrikkalaisamerikkalainen englanti on alun perin ollut syrjäytyn ryhmän puhuma englannin variaatio, sillä se on kehittynyt Afrikasta tuotujen orjien ja heidän jälkeläistensä keskuudessa. Myöhemmin tämä variaatio on levinnyt hip hop ja rap -musiikin sekä sosiaalisen median välityksellä myös muiden ryhmien käyttöön.

Tutkimuksen taustoittamiseksi tutkielmassa esitellään afrikkalaisamerikkalaiseen englantiin liittyviä erityiskysymyksiä sekä sen leviämistä maailmalle. Teoriaosuudessa käydään läpi myös afrikkalaisamerikkalaisen englannin kieltomuotoja sekä internetkielelle tyypillisiä ominaisuuksia.

Tämä tutkielma edustaa korpuslingvistiikkaa, ja tutkimuksen aineistona käytetään Corpus of Global Web-Based English -korpusta, joka on koottu teksteistä internetissä. Korpus on jaoteltu eri maantieteellisten alueiden mukaan, minkä vuoksi se sopii afrikkalaisamerikkalaisen englannin leviämisen tutkimiseen. Tutkielmassa tutkitaan kolmea kieltomuotoa, joista kaksi ovat kaksoisnegaatioita ja yksi on käänteinen kaksoisnegaatio, jossa kieltosana tulee ennen tekijää. Kaksoisnegaatioiden katsotaan yleensä amerikkalaisessa englannissa tuottavan positiivisen merkityksen lauseelle. Afrikkalaisamerikkalaisessa englannissa niillä kuitenkin on negatiivinen merkitys. Yksi tutkielman tutkimuskysymyksistä on, millaisissa konteksteissa kieltomuodot esiintyvät. Toinen kysymys koskee eroja eri kieltomuotojen yleisyydessä ja kolmas kieltomuotojen yleisyyttä eri alueilla. Tutkittavia alueellisia variaatioita ovat Filippiinien englanti, Singaporen englanti sekä australianenglanti.

Tutkielmassa kävi ilmi, että tässä korpuksessa afrikkalaisamerikkalaisten kieltomuotojen käyttö on äärimmäisen harvinaista, ja kaikilla tutkituilla alueilla kieltomuotoja käytetään hyvin vähän. Ne esiintyvät niin blogiteksteissä tyylikeinoina, kuin kommenttipalstoille kirjoitetuissa viesteissäkin. Tutkituista kieltomuodoista harvinaisin on muoto, jossa kaksoisnegaatio on käänteisessä järjestyksessä. Alueittain ei esiintynyt suuria eroja siinä, kuinka yleistä kieltomuotojen käyttö on, mutta Filippiineillä se on hieman harvinaisempaa kuin Singaporessa tai Australiassa.

Ongelmaksi tutkielmaa tehdessä muodostui korpuksen ikä, sillä se on koottu vuosina 2012–2013. Tämän takia monia verkkosivustoja, joissa kieltomuotoja esiintyi, ei enää ollut olemassa. Lisäksi aiemmissa tutkimuksissa on käynyt ilmi, että jotkin korpuksen paikkatiedoista eivät aina pidä paikkaansa. Kieltomuotojen harvinaisuuden vuoksi myös johtopäätösten tekeminen on hankalaa. Jatkossa vastaavan tutkimuksen tekemiseen olisi parempi käyttää toisenlaista aineistoa.

Keywords: African American English, internet language, corpus linguistics

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

Table of Contents

1 Introduction	1
2 Theoretical background	2
2.1 African American English and Its Spread.....	2
2.2 Negation Patterns of African American English.....	4
2.3 Internet language	5
3 Methods and Data	6
3.1 Corpus Linguistics	6
3.2 Data	7
3.3 Problems of GloWbE	9
4 Results	10
4.1 <i>Don't + V + no</i>	10
4.2 <i>Don't + V + nothing</i>	14
4.3 <i>Can't nobody + V</i>	17
4.4 Problems.....	18
5 Conclusions	19
6 Bibliography	21
6.1 Primary Sources	21
6.2 Secondary Sources	21

1 Introduction

It is widely known that English as a language is a lingua franca and spoken all around the world. Moreover, it is not only standard English that has spread among the language learners and non-native speakers. African American English (AAE), the variety originally spoken by the slaves and their descendants, has also become used by other speakers, both those whose first language is not English and those who would traditionally speak another variety of English. The spread and globalization of AAE has been seen as a result of hip hop and rap music and culture becoming more popular globally, and different countries and regions creating their own hip hop scene (Pennycook, 2007).

AAE has been stigmatized and an extra attention has been paid to its status as a variety of English (Mufwene, 2001 and McWhorter, 1997). Historically, AAE has also been regarded as “broken” and “slang” (Hickey and McQuitty, 2016). Even though this is not the view of linguists anymore, it is interesting that the variety that has its roots in the speech of a segregated group of people has gained popularity among all kinds of speakers of English.

This study will examine the use of African American English negation patterns in internet discourse, as represented by the Corpus of Global Web-Based English. There have been previous studies about the use of AAE in social media, especially Twitter, and it has for example been used to create a map of the different regional variations of AAE (Jones, 2014). The use of negations specifically has not yet been studied, which is why it was chosen as the topic of this thesis. Also, according to Labov (1968, 288), the negation patterns of AAE are non-standard. This makes them a good example of AAE, and thus studying them can be used to make some conclusions about how widely AAE has spread.

The first chapter offers theoretical background about African American English and the grammar of the negation patterns. The data and methods of the study are provided in the second chapter. After

that, results and discussion about them follows as well as discussion about some problems that this study faced.

2 Theoretical background

In this section some background on African American English will be provided, along with the grammar of the negation patterns of African American English. It will also briefly look into hip hop music as a vehicle that has helped AAE reach more global speakers. Finally, some theory on internet language will also be provided.

2.1 African American English and Its Spread

There has been a conversation whether African American English is its own language or a dialect of English. According to Mufwene (2001), many speakers of AAE do not consider it to be a language that should be distinguished from American English. McWhorter (1997, 2) has argued that AAE is not different enough to be considered its own language, and thus it is simply a dialect. He also notes that AAE is as different from standard English as the dialect of lower income Southern white speakers (1997, 2). However, that is not to say that it is completely like those dialects in terms of for example syntax, phonology, or lexicon, just that its status should not be different.

Mufwene (2001, 26-27) discusses some of the views according to which AAE should be considered a language of its own but suggests that they have their issues. For example, saying that AAE is a relexification of Niger-Congo languages would be too large of an abstraction, even though AAE does have similarities with those languages (Mufwene 2001, 27). Niger-Congo is one of the referential language groups in Africa, and languages belonging to that group are spoken across sub-Saharan Africa (Good, 2020). According to Good (2020), Niger-Congo languages include over 1,500 languages. This is why Mufwene (2001, 27) also notes that it would be difficult for AAE to be a relexification of such a various group of languages. Some scholars see AAE to be a part of a larger continuum of varieties known as Ebonics, along with West-African and Caribbean Englishes as they

have similar African origins (Mufwene, 2001, 27). However, Mufwene (2001, 27) also argues that African American speakers themselves do not consider AAE to be more similar to African or Caribbean Englishes than to American English.

African American English is often referred to as African American Vernacular English (AAVE), but these two terms differ from each other. AAVE is the variant spoken by the working-class African Americans, while AAE is more generally spoken by all African Americans (Thomas, 2007). According to Thomas (2007), AAE and AAVE pronunciations are mostly similar, but some morphosyntactic variables that appear in AAVE are missing from AAE. In this thesis the term AAE will be used, since it includes all varieties of African American speech.

It has previously been thought that African American English is a variety that is widely similar between all African American speakers (Wolfram, 2007, 295-6). That view, however, is not right and Wolfram (2007, 296) states that “In reality, regionality has played a significant role in the earlier development of varieties of AAE and it continues to play a significant sociolinguistic function in its development”. For example, the rhoticity of AAE has been shown to express some variation regionally (Wolfram, 2007, 296-8).

AAE is often simply defined as the language spoken by African Americans (Mufwene, 2001). Even though that is still where its origins are, it has spread around the world as a result of hip hop and rap music becoming more popular globally (Pennycook, 2003). Pennycook (2007, 105), states that there are rappers, such as Malaysian Too Phat, that have created their own local version of hip hop English, that has taken influence from AAE. This shows the spread of not only hip hop as a music genre but also AAE.

Nowadays, social media and especially “Black Twitter” has possibly accelerated its spread even more. According to Florini (2013), “Black Twitter” is the large presence of African American users in the microblogging site Twitter. Her study shows that the African American users perform Black

American cultural traditions through Twitter using language and modes of interactions and discuss topics that relate to African American popular culture, politics, and media (Florini, 2013). Since Twitter is used widely around the world, it is possible that internet users outside of the US are interacting with “Black Twitter” and adopting their way of expression and language use. Especially as it was mentioned before the African American popular culture, including rap and hip hop, is popular outside the community.

2.2 Negation Patterns of African American English

Negative concord, or double negation, has been seen to convey a positive meaning since the 18th century in standard English (Seright, 1966). However, it still appears with a negative meaning in many non-standard varieties of English, for example the so-called Traditional Dialects of British English that are spoken mainly in the rural areas of England, Scotland and Northern Ireland (Tubau, 2016, 144). One of the varieties in which this kind of double negation appears is African American English, the negative patterns of which we will be surveying in more detail, but it is worth noting that it is not the only one.

According to Green (2002, 77), African American English uses multiple negation to express a negative meaning, unlike general American English where a double negative would produce a positive meaning. It is also possible for a grammatically correct sentence of AAE to have more than two negators. These additional, so called pleonastic, negative elements do not add to the negative meaning, they simply agree with the first negator that conveys the negative meaning of the sentence (Green 2002, 78). The negation can appear in auxiliaries and indefinite pronouns (Green 2002, 78). Labov (1972, 226) states that African American English speakers use *nothing* the same way as general American English speakers use *anything* in negative sentences. So, instead of saying “I don’t know anything” they say, “I don’t know nothing”.

Another AAE negation pattern that is described by Green (2002, 78) is negative inversion in which “the initial negated auxiliary is followed by a negative indefinite noun phrase”. One of the examples she uses to explain this phenomenon is “Can’t nobody tell you it wasn’t meant for you”, where *can’t* is the negated auxiliary and *nobody* is the indefinite noun phrase. In negative inversion sentences the auxiliary cannot be positive, but sometimes the following noun can (Green, 2002, 79). These cases are rarer than the ones where both are negatives, but for example, “Don’t another person move” is possible (Green 2002, 79).

2.3 Internet language

Internet is nowadays an important medium for communication and that it is why internet language, or computer-mediated communication, has been actively studied. According to Squires (2010), many scholars who have studied it consider it to be a variety that is somewhere between speaking and writing. This is because it can be “informal, synchronous, and ephemeral— ‘like speech’—and/or editable, text-based, and asynchronous— ‘like writing’” (Squires, 2010). Internet language cannot be completely like speech as it happens via a computer and is usually written and as Crystal (2006, 26) says, the hardware that allows us to connect to the internet offers some restrictions on what can be said and how. For example, paralinguistic elements, such as the tone and the volume of one’s voice, must be expressed with spelling (Crystal, 2006, 36-37).

Furthermore, online interactions do not happen in real time the way speech does. Even in chatrooms and instant messaging the messages appear as full messages and not simultaneously as they are typed, therefore we cannot see how the recipient of our messages is responding to them before they type out their answer and there is always at least a small wait between interactions (Crystal, 2006, 33). This also means that the messages are usually somewhat preplanned and checked and possibly edited before sending, unlike speech that is uttered sound by sound and thus internet language is not as spontaneous as speech (Crystal, 2006, 43).

Squires (2010) argues, however, that the language on the internet is not only determined by the hardware and the restrictions of the technology. According to her, also the social characteristics and relationships of users of a platform are important factors that affect the language that is used online (Squires, 2010). Moreover, it matters where on the internet the language appears. Websites are usually much like standard written language as the writers of webpages do not necessarily know what kind of audience visits their page and want to keep it neutral (Crystal, 2006, 31). E-mails and blog pages are usually slightly more like speech than web pages, and chatgroups and instant messaging are even more similar to speech (Crystal, 2006, 31 & 44).

Social media has further developed internet language, and Hoffmann (2017, 2-3) says that it has made internet language more specialized, and therefore it is now harder to determine what the general internet language is like. Different digital platforms provide different templates for how the text and communication is organized, and often offer other ways such as “retweeting” to communicate and share ideas. (Hoffmann, 2017, 4). The different specialized text genres emerge as these templates offered by the digital platforms can be used for various purposes and topics (Hoffmann, 2017, 4).

3 Methods and Data

This section will focus on the methods and data used in this thesis. There will also be some general notions about corpus linguistics and discussion about the problems that arose using the chosen corpus.

3.1 Corpus Linguistics

Corpus linguistics is a field of linguistics that uses a corpus or multiple corpora as its source of data collection. A corpus is “a collection of spoken or written texts to be used for linguistic analysis and based on a specific set of design criteria influenced by its purpose and scope” (Weisser, 2016, 23). McEnery (2012, 1) defines corpus linguistics as “dealing with some set of machine-readable texts which is deemed an appropriate basis on which to study a specific set of research questions”. Corpus linguistics offers a good way for checking if linguistic theories appear in certain types of data (Meyer,

2002, 2). Corpora are used in many fields of linguistic research, for example discourse analysis, language learning, and sociolinguistics (McEnery, 2012, 26). According to Meyer (2002, 100) corpus linguistics has in the past been criticized for just counting frequencies and not really answering questions. He says that for a good linguistic study based on corpora to be conducted, the study should be quantitative and qualitative (2002, 102). Corpus linguistics usually uses tools to make it possible to search big sets of data quickly (McEnery & Hardie, 2012, 2). McEnery and Hardie say that these tools include concordances and often some form of frequency data. Concordances show the context in which the words appear, and frequencies show how often the words appear in the corpus, therefore they are useful in qualitative and quantitative research (McEnery & Hardie, 2012, 2).

Corpus linguistics has often been divided into corpus-based and corpus-driven linguistics (McEnery & Hardie, 2012, 3). Corpus-based linguistics simply uses a corpus or corpora as a data for studying a hypothesis or testing a theory, but in corpus-driven linguistics the corpus itself is the source of the hypotheses (McEnery & Hardie, 2012, 6). However, McEnery and Hardie (2012, 147) argue that corpus-driven linguistics does not exist since “observations, while important as necessary support for explanations, clearly do not themselves constitute explanations” – observations meaning the observations from corpus data and explanations the linguistic theory. This thesis represents corpus-based linguistics.

3.2 Data

The data was mainly collected from the Corpus of Global Web-based English (GloWbE). Also, some comparing searches were made in the iWeb-corpus. GloWbE is a collection of 1.9 billion words of text from the internet. Its special feature is that the words are organized according to the country where they are written and there are 20 of these regions represented in the corpus (Davies, 2013). GloWbE was chosen as the main source of data for this thesis because of this feature, since it shows where and to which extent the use of AAE negation patterns have spread. iWeb is significantly larger than GloWbE with 14 billion words, and therefore it was used for the purpose of comparison, in few

instances. It also consists of text from the internet, but is not sorted according to countries, like GloWbE.

Three negation patterns were chosen to be examined: “*don’t + V + no*” and “*don’t + V + nothing*”, with a verb in between as in “don’t say no” or “don’t say nothing” and “*can’t nobody + V*” with a verb in the end as in “can’t nobody tell me nothin”. They represent the typical negation patterns of AAE, as the first two are examples of double negatives and the third one also includes a negative inversion. The corpus search strings used for these patterns were “do n’t VERB no”, “do n’t VERB nothing” and “ca n’t nobody VERB”.

The research questions of the study are:

- 1) In which kind of contexts do AAE negations appear?
- 2) What kind of differences are there in the frequencies of use between the different types of negations?
- 3) What kind of differences are there in the frequencies of use in different countries?

The hypotheses are that the negation patterns of AAE have made their way into the internet language of non-native speakers and that there are some differences in their frequencies. Double negatives are more general type of negation pattern than negative inversions, so the hypothesis is that there are more results of “*don’t + V + no*” and “*don’t + V + nothing*”, than there are of “*can’t nobody + V*”. The contexts in which the negations are expected to appear are informal and probably more likely to be from a discussion forum than a blog text. Blog texts are more thought-out language and thus may not include such speech-like elements.

The analysis focuses on the different countries and how frequently the negation patterns appear in the texts collected from them. The United States is not examined in detail as that is where it originates and thus does not provide information on the spread of the AAE negation patterns. Also, since in

United Kingdom some of the results may be from speakers of Caribbean varieties that are similar to AAE, it is likewise excluded from the study. Moreover, for the same reasons the African varieties and Jamaican English are not included in the study. The countries which were studied were the Philippines, Singapore, and Australia.

Excluded from the analysis are also metalinguistic references, such as texts that teach about double negatives or present attitudes towards them. There were also some instances in “*don’t + V + no*” that are not double negatives. For example, in the clause “I generally don’t say no to alcohol”, the *no* is not in the place of the indefinite pronoun *any* and does not make it a double negative. Similar to this is the phrase “don’t take no for an answer”. Other examples where the construction is not actually a double negative, are places where a comma has been left out.

1) ... then by all means *don’t watch no body* is forcing you

(“... then by all means, don’t watch, nobody is forcing you”)

Example 1) is an instance from the data that simply misses a comma or a period, with the expected version presented in the parentheses. As internet speech is often quite informal, these kind of “shortcuts” are not surprising. These are not included in the study, either.

3.3 Problems of GloWbE

As mentioned before the texts of the GloWbE corpus are organized under the countries of their origin. Callies (2018, 158), however, notes that there are some problems with the identification of the country where the text has been produced. Firstly, after verifying that the texts in the corpus are marked in the right countries, he found that not all of them were sorted correctly and thus did not provide accurate information about the appearance of a result in a certain variety of English. Secondly, confirming some of the results was impossible, since the websites where the texts originally appeared, were deleted and not accessible anymore. This is something that was also apparent in the collection of the data for this thesis. Hoffmann (2018, 179) provides another notion on the accuracy of the

locations on GloWbE. He mentions that many of the results are comments on the websites and not written by the author, so their country of origin may be incorrectly marked. These things are important to take into account in analyzing the data, as they can skew the results slightly.

GloWbE has also been collected in 2012-2013, so it does not necessarily represent the most recent forms of internet discussion and writing, but it still offers information about language use in different countries in an accessible way. Furthermore, even though both Callies and Hoffmann provided criticism on the corpus, they nevertheless concluded that it is a useful resource as they used it in their studies. The user simply must be aware of the problems that the corpus has and take them into account when using it as a source.

4 Results

In this part the results of the corpus search will be discussed. The different negation patterns will be examined in order from the most popular to the least popular. In the end some problems regarding this thesis are presented.

4.1 *Don't + V + no*

Contrary to the hypothesis that the negation patterns of AAE are popular in internet speech, the “*don't + V + no*” pattern does not appear that often in the corpus data. It is the most common out of the three negation patterns studied, but still not very prominent. Its total frequency in the data is 1169 and the normalized frequency is 0.61 per million words (pmw), meanwhile the pattern “*don't + V + nothing*” has the frequency of 322 (pmw) and “*can't nobody + V*” that of 27 (pmw). This shows that altogether AAE negation patterns are not that typical in internet speech, at least according to the GloWbE corpus.

It is, not surprisingly, most common in the United States with 405 results out of all 1169. As mentioned in the “Methods and Data” section, however, the focus of this study is on the results from the countries where AAE is used mostly by non-native speakers of it. In the Philippines there seems to be a somewhat high frequency per million words compared to the other varieties, even slightly

higher than in the US, as the number for the Philippines is 1.18 and for the US 1.08. However, as will be observed in greater detail later when we take a closer look at the results, it can be seen that this frequency does not consist of instances of authentic internet writing. As we can see from the Table 1, out of the total of 50 results, only 2 are examples of it. The frequency per million words for these is 0.05. This shows that according to this data, AAE negation patterns have not made it largely into the internet writing of the Filipinos.

Region	Number of hits of genuine internet writing	Frequency per million words
The Philippines	2	0.05
Singapore	9	0.20
Australia	29	0.20

Table 1. The frequencies of the “*don’t + V + no*” pattern in the GloWbE.

The instances of negation patterns that could be verified as genuine cases where the pattern was used by a Filipino author seem to be used as some kind of stylistic devices in the middle of text that is standard English.

1) The Mazda 6 *don’t get no* love (GloWbE, PH, www.topdesingcompany.com)

2) But I *don’t need no* preaching when it comes to how to quit my 9-to-5 job...

(GloWbE, PH, www.linguist-in-waiting.com)

Example 1) is from a comment about cars and the rest of the text is written in more standard English. Example 2) is from a blogpost that talks about the issues with travel blogs. Like in example 1) the rest of the text sample does not appear to represent AAE. It could also be indication of mixing AAE with other types of English, instead of adopting it completely and using it more systematically. Obviously, very thorough conclusions about the context and way in which the negation patterns of AAE are used, cannot be made with this small number of examples.

Most of the 50 hits of the pattern in Philippine English were in fact song lyrics with 32 of the results being them. Other results that were not relevant to this study included quotes from interviews and excerpts from literature. Some of these are lyrics from hip hop songs, for example:

3) *We don't need no* water just let that mother burn! (GloWbE, PH, www.lyricskeeper.com)

According to the website from where this text sample originally is, these are lyrics from a song by the American rapper Nelly. Rap lyrics appearing in the data of study about AAE is logical considering that it has had an important role in popularizing AAE globally, as was mentioned before in the theory part. Rap music is not the only genre of music that appears in the results. The following example 4) is from an American gospel duo Mary Mary.

4) If God's in me, why I gotta think small /

Now you know that *don't make no* sense at all (GloWbE, PH, www.christian-lyrics.net)

Similarly to rap and hip hop, gospel music has employed AAE forms. Like, Green (2002, 79) notes AAE negation patterns appear in African American religious songs. Both the rap song and gospel song thus represent very typical examples of AAE, and if they are really posted on the websites from the Philippines, it can show that even though the music is spreading the negation patterns for people to hear, they are not necessarily using them themselves in internet language.

Singapore has a lower frequency compared to the Philippines and a lower frequency per million words (0.77), but there are more relevant examples. Out of the 25 results that are left after counting out the instances that are not double negatives, nine (pmw = 0.20) are internet writing. Most of the other kind of instances are again lyrics, and there are also quotes of spoken language from interviews. Unlike, both examples from the Philippines example 5) below is not a lone sentence of AAE in the middle of standard English, but the text uses other non-standard forms like “*imma*” and “*cuz*” as well.

5) *I don't got no* social life (GloWbE, SG, www.dramabeans.com)

This could be a result of the context as the examples from the Philippines were from blog posts and represent a more planned and refined form of writing, this example is from the comment section of a website, so the communication is more spontaneous and more speechlike. There are, however, also hits that are from blogposts and that have a very similar context to those from the Philippines. An example of this is 6) below.

6) *don't work no* more (GloWbE, SG, feedmetothefish.blogspot.com)

This text expresses an opinion on politics, and it is written in standard English. It, however, does employ some colorful expressions such as “your cronies and testicles carriers”. Again, the double negative is used as a stylistic choice.

According to this data, even though the percentage of internet writing among the results was a bit higher than in the Philippines, the “*don't + V + no*” -construction still is not common.

Australia has the frequency of 0.52 pmw, so it is again smaller than in the Philippines or Singapore, but a similar pattern can be observed. Out of the 77 instances, 29 seemed to represent internet language and the rest were, again, lyrics, quotes, and excerpts from books. If we look at the frequency per million words for those 29 examples it is roughly the same as that of Singaporean English, 0.20. This is also presented in the Table 1.

In the case of Australia in the corpus there were some missing contexts and websites, so confirming the examples to be either genuine internet language or something else, was not possible. The instances of actual internet language or discussion were both found on comment sections and as stylistic choices in blogposts or articles.

7) *we don't score no* tries much these days (GloWbE, AU, www.theroar.com.au)

Example 7) is from a message in the comment section of an article about rugby. While the rest of the message does not use quite as informal forms as in example 5) above, it does use expressions like “*bang on about*” and “*damn nuisance*”, so the writing is somewhat informal. The following example 8) is from a blog post that is stylistically very colorful, and it is used for a humorous effect, since the writer says that she sometimes yells it out in her “wannabe gangster voice”. It is even repeated in the very end of the text as a closing statement. As observed before, the same pattern or way of using AAE negations repeats again.

8) I'M A FREE INDEPENDENT WOMAN; I *DON'T NEED NO* MAN (GloWbE, AU, lipmag.com)

4.2 *Don't + V + nothing*

Region	Number of hits of genuine internet writing	Frequency per million words
The Philippines	0	0
Singapore	2	0.05
Australia	7	0.05

Table 2. The frequencies of the “*don't + V + nothing*” pattern in the GloWbE.

As mentioned already in the section 4.1, “*don't + V + nothing*” appears in the corpus a total of 332 times and has the frequency of 0.17 pmw. This double negative is also the most used in the US, with 92 out of the 332 hits being from there (pmw = 0.24). Similarly to “*don't + V + no*”, the Philippines has a relatively high frequency of 0.49 pmw of the “*don't + V + nothing*” construction compared to other countries. After analyzing the results of the previous double negative, it comes as a no surprise that in the case of this construction, too, many of the hits are not genuine examples of the internet writing of Philippine English users. In fact, as can be seen in Table 2, none of the hits represent it. Again, most of the results are from lyric websites with 13 of the results being them. However, in this section there are some metalinguistic references. These are not instances of genuine internet writing, but they show that the construction is used in spoken language but is not considered proper by these

writers online. This could show us that the use of these negative constructions is part of the spoken language of Filipino speakers but does not appear in this data for it consists of more refined written text. The example 9) below represents this kind of hit.

9) But no matter if a thousand and one American TV sitcoms use " *I don't know nothing*" as a staple expression, there's no way for it to be accepted as good English. (GloWbE, PH, josecarilloforum.com)

As can be seen from this, the writer sees the double negative as an improper use of English. Furthermore, the influence of American English in the use of this construction is visible in this example, although it is not necessarily mentioned to be AAE and instead of music, it is from TV.

Singaporean English does not have many results of "*don't + V + nothing*", only four of them or the frequency of 0.09 pmw. But unlike any of the Philippine English results, two or the frequency of 0.05 pmw of these hits are instances of genuine internet language. The example 10) is from a comment section of a blog text that discusses "dramas", a genre of Asian tv-series.

10) if you can't say something nice then *don't say nothing* at all (GloWbE, SG, koalaspplayground.com)

The rest of the comment seems to be informal but not completely speechlike much like the earlier examples. For instance, the expression "*old hag*" is used in the comment, which gives it a colloquial tone.

The other example of genuine internet language is

11) you *don't lose nothing*, she did! (GloWbE, SG, bongqiuqiu.blogspot.com).

It is not clear whether this is a comment or a blog text as the site does not exist anymore but from the context provided in the corpus, we can see that this text also uses informal language. "Go bonkers about shit" is one of these instances of informal language in the text. In these examples the double

negative is not as clearly a stylistic choice as they are quite informal in style. However, the grammar seems to follow the rules of standard English.

The number of hits for Australia is 30 and the frequency per million words 0.20. That is slightly higher than that of Singapore. Furthermore, there are seven examples that were identified as genuine internet writing, and that is also more than the number of them for Singapore. However, the frequencies per million words are again the same, 0.05 for both of Singapore and Australia as is shown in Table 2. Again, the hits that were not genuine internet language were song lyrics and quotes. This time only three of them were lyrics and the rest were quotes from other texts or from speech.

Example 12) is from a forum where people can ask questions and others can help.

12) I argue on the internet (not intentionally) but *don't say nothing* in School (GloWbE, AU, au.answers.yahoo.com)

Unlike the previous examples in this section, this comment does not contain any profanities or other kind of striking language. The double negative here does not seem to be used as device to make the text more interesting, as the text does not use any other tactics for that. Moreover, this text is not necessarily written to attract readers or entertain them as it is simply a comment that answers to other's comments. This site has, too, been deleted but from the context in the corpus we can see that it starts with "Thanks for the answers!" so it was presumably written by the writer of the original question. This makes the use of the double negative seem quite spontaneous, and not a stylistic device.

Example 13) below is more like the earlier examples in this section than example 12) as the surrounding text uses some crude language such as "*fucker*" being used to refer to people.

13) but they don't talk nothing when taj's (and everybody!) doing reo after reo...
(GloWbE, AU, stabmag.com)

This seems to be a comment on an article, and it is quite informal in style as it, for example, does not use capitalization. Therefore, the case is similar to the examples 10) and 11).

According to the data “*don't + V + nothing*” is not regularly used in the internet language of these areas. There were not many examples of the construction used in any of these varieties in genuine internet writing. As was mentioned earlier in this section it is not surprising because this was the case with the previous negative pattern also. Most of these hits seem to be from comments rather than blogposts which would support the hypothesis that the informal and more speechlike nature of comment sections makes the use of double negatives more common. However, once again the low number of hits complicates making conclusions, and in section 4.2 the double negative was used in blogposts also.

4.3 *Can't nobody + V*

It was mentioned before in section 4.1, that “*can't nobody + V*” is the least common out of the negations studied in this thesis. It appears only 24 times in the whole corpus, and most of the countries have no results at all. Even in the United States, where these negation patterns would be expected to be more common, it appears only 14 times. Furthermore, all of these results are either song lyrics or quotes of spoken language. So, it seems that, at least during the time the corpus was collected, the negative inversion “*can't nobody + V*” has not found its way into the internet language of even Americans. This is in line with the hypothesis that non-inverted double negatives would be more common than negative inversion, that was made in the section 3.2.

Out of the countries where there were instances of the negation pattern, only in Australia it represents actual internet communication. All of the others are either song lyrics or song titles. The only example from Australia is:

- 14) I feel great!!!! *Can't nobody tell* me nothin (GloWbE, AU,
www.goldcoastbootcampdeals.com)

It is from a blog, and the text talks about weight loss and how it has affected one's life. However, it is not written by an Australian, but an American who has made a visiting appearance on the blog.

As a comparison, a search was made in the iWeb-corpus, as it includes a larger number of words, and the same phenomenon can be seen there. iWeb is not sorted by countries, so it is not possible to examine its spread outside of the US, but out of the 14 billion words in the corpus, "*can't nobody + I*" appears only 90 times and has the frequency of 0.006 pmw. Many of the results are also from lyric pages, precisely like in the GloWbE corpus.

In conclusion, this negative inversion does not seem to have a significant place in the internet vocabulary of people, according to these corpora. It is possible that it is more frequent in social media that is not included in these corpora or that is newer, than 2013 when GloWbE was collected or 2017 when iWeb was. That, however, is a possible topic for a further study and will not be researched in this thesis.

4.4 Problems

The original aim of this study was to research the use of AAE negation patterns in the internet language of non-native speakers of AAE, but the problems that were mentioned in section 3.3 and the very small number of results appearing in the corpus searches made the conclusions that can be made somewhat weak. One big issue with especially the Australian section of the corpus is that many of the websites did not exist anymore and it was difficult to study the context of the examples in detail. Thus, the study ended up providing more information on whether the corpus is suitable for this kind of study and uncovered some issues that one might face when doing a similar study using the GloWbE corpus. The corpus is not best suited for studying this subject and for a further study another set of data, for example a specially constructed corpus, would be preferable.

5 Conclusions

In conclusion, this thesis studied the use of African American English negation patterns in the Corpus of Global Web-based English. This topic was chosen as AAE and African American culture have gained popularity around the world. Especially rap and hip hop music and BlackTwitter can be seen as driving forces of their popularization. The aspects which were studied were the frequency of the negative constructions, the frequency of these negation patterns in different countries and the contexts in which they appeared. The data used for the study was the Corpus of Global Web-Based English. Many regions that are included in the GloWbE were excluded from the study as they are similar to AAE because they are African varieties of English or spoken in the Caribbean where their history of being originally spoken by African slaves is similar to AAE. The varieties that were ultimately chosen were Singaporean English, Philippine English, and Australian English.

Contrary to the hypothesis of this thesis, all of the patterns studied were quite rare in the corpus. However, the other hypothesis was supported as the rarest of these negation patterns was the negative inversion “*can't nobody + V*”. The next was “*don't + V + nothing*”, and the one that appeared the most was “*don't + V + no*”. Many of the hits were song lyrics or other kinds of quotes and they had to be excluded from the study as they did not represent genuine internet language. The negatives appeared in comments that used informal language and even profanities, but also in some blog texts that were in other ways formal. In these blog posts the negative patterns seemed to be used as stylistic devices to make the text more interesting. The difference in the use of these negation patterns between the regions that were included in the study was not very significant. The authors of the texts from the Philippines used them slightly less than the authors from Singapore and Australia.

While doing this thesis I faced some practical problems. Some of them were related to the chosen corpus as it is already ten years old and some of the websites were not available anymore. Moreover, some of the locations in the GloWbE are placed incorrectly. The number of hits that did not represent genuine internet writing of these regions was really high and the number of actual hits that could be

included in the study was low. Therefore, making definite conclusions based on the results of this study was not feasible. This thesis did, however, provide some information on whether this corpus is suitable for such a study. The conclusion for that question is that another corpus would suit the aims of this kind of study better.

6 Bibliography

6.1 Primary Sources

Davies, Mark. (2013) *Corpus of Global Web-Based English*. Available online at <https://www.english-corpora.org/glowbe/>.

Davies, Mark. (2018) *The iWeb Corpus*. Available online at <https://www.english-corpora.org/iWeb/>.

6.2 Secondary Sources

Callies, Marcus. “Patterns of direct transitivization and differences between British and American English” In *Changing Structures: Studies in Constructions and Complementation*, edited by Kaunisto, Mark, Mikko Höglund, and Paul Rickman, 151-167. Vol. 195. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2018.

Crystal, David. *Language and the Internet*. Cambridge: Cambridge University Press, 2006

Florini, Sarah. “Tweets, Tweeps, and Signifyin: Communication and Cultural Performance on ‘Black Twitter.’” *Television & new media* 15, no. 3 (2014): 223–237.

Good, Jeff. “Niger-Congo, with a Special Focus on Benue-Congo.” In *The Oxford Handbook of African Languages*. 1st ed. Oxford University Press, 2020.

Green, Lisa J. *African American English a Linguistic Introduction* Cambridge, U.K. ;: Cambridge University Press, 2002.

Hickey, McQuitty. “Embracing African American English.” *English journal* 106, no. 2 (2016): 99–101.

Hoffmann, Sebastian. “I would like to request for your attention: On the diachrony of prepositional verbs in Singapore English” In *Changing Structures: Studies in Constructions and*

- Complementation*, edited by Kaunisto, Mark, Mikko Höglund, and Paul Rickman, 171-196. Vol. 195. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2018.
- Hoffmann, Christian R., and Wolfram Bublitz. *Pragmatics of Social Media* Berlin, [Germany] ;: De Gruyter, 2017.
- Jones, Taylor. "Toward a Description of African American Vernacular English Dialect Regions Using 'Black Twitter.'" *American speech* 90, no. 4 (2015): 403–440.
- Labov, William. *A Study of the Non-Standard English of Negro and Puerto Rican Speakers in New York City. Volume I: Phonological and Grammatical Analysis*, 1968.
- Labov, William. *Language in the Inner City : Studies in the Black English Vernacular*. Philadelphia: University of Pennsylvania Press, 1972.
- McEnery, Tony, and Andrew Hardie. *Corpus Linguistics : Method, Theory and Practice*. Cambridge University Press, 2012.
- McWhorter, John H. "Wasting Energy on an Illusion." *The Black scholar* 27, no. 1 (1997): 9–14.
- Meyer, Charles F. *English Corpus Linguistics an Introduction* Cambridge, UK ;: Cambridge University Press, 2002.
- Mufwene, Salikoko S. "What is African American English?" In *Sociocultural and Historical Contexts of African American English*, edited by Sonja L. Laenhardt, 21-51. Philadelphia, PA, USA: John Benjamins Publishing Company, 2001.
- Pennycook, Alastair. "Global Englishes, Rip Slyme, and Performativity." *Journal of sociolinguistics* 7, no. 4 (2003): 513–533.
- Pennycook, Alastair. "Language, Localization, and the Real: Hip-Hop and the Global Spread of Authenticity." *Journal of language, identity, and education* 6, no. 2 (2007): 101–115.

- Seright. "Double Negatives in Standard Modern English." *American speech*. 41, no. 2 (1966).
- Squires, Lauren. "Enregistering Internet Language." *Language in Society* 39, no. 4 (09, 2010): 457-492.
- Thomas, Erik R. "Phonological and Phonetic Characteristics of African American Vernacular English." *Language and linguistics compass* 1, no. 5 (2007): 450–475.
- Tubau, Susagna. "Lexical Variation and Negative Concord in Traditional Dialects of British English." *The journal of comparative Germanic linguistics* 19, no. 2 (2016): 143–177.
- Weisser, Martin. *Practical Corpus Linguistics: an Introduction to Corpus-Based Language Analysis* First edition. Chichester, England: Wiley Blackwell, 2016.
- Wolfram, Walt. "Sociolinguistic Folklore in the Study of African American English." *Language and linguistics compass* 1, no. 4 (2007): 292–313.