

**A Qualitative and Quantitative Analysis of
Selected Ethnic and Racial Terminology
Present in Assorted Public English Corpora**

Natalia Pietraszewska, PhD Student
Philological School of Higher Education in Wrocław, Poland

Abstract: The central focus of the article is to analyse large collections of linguistic data in order to unveil language patterns pointing to the pervasive racist sentiments. Lexical choices made by speakers are assumed to reflect their attitudes to certain minority groups. Since languages reflect ideas present in a society, it is possible to find examples of racial and ethnic prejudices expressed either overtly with ethnophaulisms or more inconspicuously in certain lexical choices. This study is concerned with discriminatory practices in the form of racist and ethnic slurs as well as the most frequent collocates pointing to negative stereotypes of selected ethnic minorities. The search for biased language was conducted in large British and American English corpora, including Internet treated as corpus. The author attempts to present how the visual and cultural differences as well as the portrayal of particular minorities in media may influence language pertaining to ethnic groups. The results indicate that ethnic stereotypes are still reproduced in both British and American English. Prejudiced language appears to be associated with both visibility and socio-economic status of a particular group in a certain country.

Keywords: stereotyping; corpus analysis; ethnophaulisms; racist discourse

The knowledge of language enables people to categorize and classify by equipping them with useful generalizations that simplify what they perceive. Culture affects the mind, and because language is one of the tools by which it may be accomplished, language thus affects human cognition. Languages enable humans to interpret each experience, code it and connect it with their conceptual systems (Fitch and Sanders 2005: 6-7). It can be observed that lexical and grammatical patterns are partly conserved in certain social groups and are determined by the setting in which they were formed. They also reflect the general knowledge of a given community (Murphy 2002: 420).

The need to classify and organize knowledge may lead to the formation of cognitive schemata. This process was explained in the works of Eleanor Rosch (1978) and George Lakoff (1987), who pointed out that people have a tendency to group objects that quite often cannot be distinguished by defined traits. Such categorization is based on prototypes, which may be understood to be members of a category that fully fit the definition and to which the other category members are compared before being recognized as similar and categorized (Rosch 1978: 10-11). This may lead to the

emergence and preservation of stereotypes, which are understood as mental shortcuts that allow people to process what they perceive in relatively short time (Duranti 1997: 32). Cultural knowledge may take the form of cognitive schemata which are shared within a society, as Naomi Quinn (2005: 38, in Brown 2006: 100) believes. Similarly, Charles Stangor and Mark Schaller (1996: 10) note that stereotypes are not only individual but also cultural phenomena and that they are shared between individuals, learned and reproduced. Thus, in various cultures, different groups and categories of people are stereotyped. One of the examples mentioned by David Schneider (2004: 230) is the fact that in the United States, the most common stereotypes are associated with African Americans. This may imply that cultural and social factors influence the process of stereotyping. The number and kind of stereotypes present in a given language may therefore reflect the cognitive system of a certain speech community. The choice of linguistic features that a person uses when relating to a particular topic reflects their attitude toward it (Van Dijk 1987: 11). Linguistic expressions may also contribute to the creation and transmission of racism and other prejudicial attitudes (Ruscher 1998: 241-242). Teun van Dijk (2004: 351) noted that it is possible to divide racist language practices into two categories. The first is overtly hostile and derogatory racist language aimed at ethnically different groups. It encompasses derogatory terms, including slurs and impolite forms of address. The second form of racist discourse is used in interactions with other members of speaker's ethnic group and is about the stereotyped race or ethnicity (van Dijk 2004: 351). It is possible to note this form in the way other ethnic groups are portrayed. Such a biased representation of the stereotyped group is acquired due to various strategies that according to Teun van Dijk (2004: 352), "(...) may appear at all levels of text and talk (...)". The analysis of structures used in certain types of biased discourse helped Teun van Dijk (2000: 147) to identify some of the ways of conveying prejudiced ideas, including rhetorical devices such as the use of metaphor, metonymies, selection of lexicon, and even emphasizing certain topics in order to focus attention on negative information about the ethnic minority. In this day and age overt forms of verbal discrimination are generally considered to be politically incorrect, and the racist discourse directed at dominated ethnic group members is less blatant. It is more probable to come upon such negatively charged expressions in an informal conversation or on an Internet forum.

One of the main instances of racist discourse is racial epithets, also called ethnophaulisms, which are based on conceptual metonymies. These pejorative terms help to gain access to a certain conceptual category because they refer to some of the features representing its members (Kövecses and Radden 1998: 54). As Janet Ruscher (2001: 20–21) noted:

[e]pithets boast varying etymologies, which provide insight into how outgroups are — or historically have been — perceived. For example, research demonstrates that perceivers are sensitive especially to characteristics that minimize within-group variability and maximize between-group variability (...) Not surprisingly, then, a large number of epithets point

to intergroup differences with respect to alleged physical characteristics (e.g., darkie, redshanks, skirt), cultural and eating practices (e.g., porker, spaghetti eater, frog), and names (e.g., Paddy, Heiny, Guido (...))

The described properties are usually derogatory and refer to easily recognizable and visible attributes of certain group members. The prevalence of ethnophaulisms in relation to the size of the stereotyped group was researched by Mullen and Johnson (1993) and described by Charles Stangor and Mark Schaller (1996), who noted that in the case of smaller groups, the complexity and number of ethnic epithets was significantly lower. Therefore, the comparison of both the quantity and quality of slurs and the ethnic structures of particular countries may help to assess whether the degree of stereotyping of certain groups reflects their visibility in certain societies.

1. Corpus Linguistics in the Search for Racist Stereotyping

In order to establish the prominence of racist language, it is advisable to assess the extent to which prejudice markers are present in diverse English corpora and compare the findings for various societies. Professor Wolfgang Teubert (2005: 4) noted that corpus linguistics means “(...) an insistence on working only with real language data taken from the discourse in a principled way and compiled into a corpus”. The researched corpus may be any database or collection of authentic texts reflecting real language, which may be in the form of transcripts of everyday communication between members of one or various communities, newspaper articles, advertisements, letters, *etc.* Michaela Mahlberg (2005: 17) outlined the framework for the theoretical approach based on corpus linguistics and noted that “[m]eaning becomes observable through the repeated patterns of words, which are visible in corpora.” Similarly, John Sinclair (2004: 189) observed that such patterns are the reflection of language use in the social group. A certain corpus consisting of texts produced by members of a language community enables the researcher to find recurrent expressions and topics that are often discussed and thus draw conclusions regarding users’ attitudes and beliefs (Teubert 2005). Thus, corpus linguistics is a useful approach in studies of discourses that are visible in the corpus mainly in the form of patterns of words and phrases to which speakers are oblivious.

Large collections, such as the Corpus of Contemporary American English or the British National Corpus, include texts representing a language as a whole. Both corpora are balanced, which means that they include many genres of spoken and written language derived from various sources. The most viable source of data is, however, the Internet. Its contents comprise all the articles and publications available via the World Wide Web, making it the richest database of a wide variety of texts. What is more, since Internet users feel anonymous, they formulate their thoughts in a less careful way than they would in other media or communicational situations. Thus, the Internet may be a source of terms and phrases that reflects people’s thoughts in an

accurate way. The variety of texts present on the Internet is astonishing, and the freedom of speech and anonymity available ensure that racist discourse is much easier to note than it is within utterances from other collections. Search engines, namely Google, enable one to assess the frequencies of rarely used terms or collocations that would be impossible to find in other corpora, as noted by Irene Mittelberg (2007: 43). This suggests that search engines may be a vital source of linguistic data that reflects real language usage in the most accurate way.

One of the methods that is useful in elucidating the aforementioned patterns is establishing a concordance, which is understood to be all occurrences of the linguistic item in a given corpus within its linguistic context (Tribble, 2010: 169). The information about terms that co-occur with the search term more frequently than is possible merely due to random distribution is essential for studies of context. As John R. Firth (1957: 181, in Manning and Schütze 1999: 141) observed: “[c]ollolocations of a given word are statements of the habitual or customary places of that word.” They may, therefore, be a sign of some conceptual schemata reflecting the concepts linked with a certain term. Michael Stubbs (1996: 172) noted that: “(...) words occur in characteristic collocations which show the associations and connotations they have, and therefore the assumptions which they embody.” One of the examples of collocation that has become a fixed phrase and may support Stubbs’ view is a frequent co-occurrence of the terms *illegal* and *immigrant*, suggesting that immigration is perceived as something wrong and possibly leading to a priming effect (Hunston 2002: 119). The frequency of epithets or collocations of a term describing a minority, as well as some negative stereotypy terms, should differ in various countries because it reflects the views on the described group that are held by the general public. Moreover, the analysis of the typical adjectives collocated with a name of an ethnic group should also shed some light on the kind and extent of negative racial or ethnic stereotyping in each case.

The analysis of prejudice markers visible in the various corpora in this study involved three steps, each focused on a different type of data and corpus. First, the origins and denotations of the derogatory epithets listed on the Racial Slur Database website¹ were checked. The racial slurs listed on this website were collected with the aid of suggestions from Internet users who have ensured that the collection encompasses a wide variety of epithets and is possibly representative of derogatory slang. The possible etymologies of the collected slurs were also provided on the website. The epithets were first grouped according to the minority they were used to describe and subsequently organized on the bases of origin and the feature they referred to.

A study of the common collocations of terms connected with stereotyped populations was conducted with the aid of specialized tools available on the website of the Corpus

¹ <http://www.rsd.org/> (Retrieved 11.2011)

of Contemporary American English (COCA)¹. These tools enable searches in both COCA and the British National Corpus corpora. As the first step, the frequency of a term denoting a minority was established. Subsequently, a search for most common collocates was conducted. The results were grouped according to their frequency. However, in order to exclude high-frequency words, such as articles or prepositions, the result list was limited to collocations with mutual information scores over 3.0. The mutual information score provides information about the relevance of each collocation and allows for the disposal of high-frequency words, as was explained by Mark Davies (2008) on the website². Thus, the MI score makes it possible to examine collocates that are specific only to the search term and do not simply collocate with most of the terms in the corpus. The list obtained as the result of the search was further processed, and typical collocations with terms playing a role in the stereotyping process were selected in order to compile a list of negative collocates. Frequency in the corpus determined the order of collocates on the list. Collocates pointing to prejudice were compared for the American English corpus (COCA) and the British National Corpus (BNC).

The final step was based on the assumption that the Google search engine is the most widely used search tool in the world³. One of its features, Google Suggestions, may prove valuable in uncovering the extent of stereotyping in the portrayal of the selected minorities. After the user types a part of a term or sentence, Google Suggestions functionality hints at the way in which it may be finished. The suggestions are based on search query popularity. This means that if the user types “*why are (a name of the group)*” or “*(name of the group) are*” the suggestion will, in most cases, be the adjectives that are typed after this sequence by the majority of users. Most importantly, suggestions are based on constantly updated statistics regarding the popularity of queries, so they reflect the current state of computer-mediated discourse.

2. Ethnic Groups Studied

In order to examine the claim that the racial epithets and other forms of racial discourse pertaining to each minority group vary by country, the information on the statuses of various ethnic groups in the United States of America and Great Britain was investigated, and the data were compiled. The data for each country were obtained from the appropriate census survey. Therefore, it is vital to remember that the ethnic origin data in the census is a reflection of respondents’ perceptions of their ethnic backgrounds. Thus, ethnicity should be understood as a perceived status and not necessarily based on any objective markers of ethnicity or race. There are many ethnic groups in most English-speaking countries. However, the United States seems to be a country that has been shaped by immigration waves and is comprised of a considerable

¹ <http://www.americancorpus.org/> (Retrieved 11.2011)

² <http://www.americancorpus.org/> (Retrieved 11.2011)

³ <http://www.experian.com/hitwise/online-trends-search-engine.html> (Retrieved 11.2011)

number of ethnicities. Data concerning the number of members of minority groups and their socio-economic status in the US were obtained with the aid of the Fact Finder application available on the US Census Bureau's website¹. The most populous group was composed of citizens who categorized themselves as White – those originating from Europe, the Middle East, or North Africa. The largest minority group was comprised of people of Hispanic ethnicity. The second largest minority, Black or African American people, included those having ancestors in any of the Black racial groups of Africa. People categorized as Asian (Asian American) originated from the Far East, Southeast Asia, or the Indian subcontinent. The second country examined was the United Kingdom, which had an overall non-White population of almost 8% in its 2001 census². In 2001, the largest ethnic minority group in the United Kingdom was comprised of people of Indian origin, who accounted for nearly 2% of the total population and 23% of the ethnic minority population. This group was followed by Black minorities of various origins (mainly Caribbean and African). The third largest minority was composed of Pakistanis (16% of the total minority population), which comprised 1.8% of the total population. Unfortunately, the obtained data do not reflect the percentage of one of the most negatively stereotyped groups in many North American and European countries – the Muslim minority. This is due to the fact that this group consists of people of different ethnic origins who are connected by religion. However, as Muslim stereotyping seems to be associated with the Middle Eastern and Asian origins and cultural differences, it is advisable to consider it a manifestation of prejudiced discourse.

In order to study the attitudes towards ethnic and racial minorities, three minority groups were chosen. Firstly, Muslim stereotypes were researched because the members of this religious minority are culturally distinct from the White, predominantly Christian majority. Therefore, immigrants of Middle Eastern backgrounds are an obvious target for ethnic stereotyping. Notably, in the public opinion survey conducted in 2006 in the United States by the Council on American-Islamic Relations, approximately 25% of respondents described Islam as a religion of hatred and violence, and 26% had negative associations with the word *Muslim*, such as *violence, hatred, terrorists, war or guns*. The second selected group was the Black minority, which was stereotyped mainly due to differences in appearance and its presence in the media in the context of criminal activity. The survey conducted by Lawrence Bobo and James K. Kluegel (1997: 100-101) revealed that 31 per cent of White respondents gave Blacks a low rating in terms of intelligence, 47 per cent rated them as lazy, 54 per cent claimed that Black people are prone to violence, and as many as 59 per cent expressed the opinion that the Black community tends to live off welfare programs. Finally, examining the stereotyping of Hispanics,

¹ <http://factfinder.census.gov> (Retrieved 11.2011)

² <http://www.statistics.gov.uk/census2001> (Retrieved 10.2011)

who make up the most numerous minority group in the United States, may shed some light on attitudes towards immigration. Hispanic and Latino Americans originate from countries in Latin America or Spain and constitute the largest minority group in the United States (Dinnerstein and Reimers, 2009: 168). There are many factors that contribute to prejudice against this ethnic group, such as the issue of poverty, because the average income of a Hispanic American is considerably lower than the national average. Moreover, according to the 2000 census, education levels are much lower among Hispanic Americans than among non-Hispanic White Americans. The media also focuses on illegal immigration. In 2002, about 20 per cent of news stories related to Hispanics mentioned this topic. Also, 10 per cent of the news items studied mentioned Latino crime (Méndez-Méndez and Alverio 2003). Such portrayals also contribute to the formation of stereotypes.

3. Stereotyping of the Muslim Population

The majority of epithets enlisted in the Racial Slurs Database are references to easy-to-note features, such as appearance or religious customs. This finding is in accordance with the fact that the majority of those features are based on characteristics that are easy to notice (Ruscher 2001: 20–21). Thus, there are numerous epithets pertaining to culturally affected appearance, predominantly clothing items such as turbans or burquas. Moreover, religious practices, namely communal prayers, which may be conducted in places attended by the general public, may distinguish Muslims from other groups. The results reveal that many epithets pertain to the conflict between radical militant Muslim groups and Western civilization in terms of associating Arabs and Muslims with terrorism. Most of these terms refer to suicide bombings, which are unfamiliar to members of western societies. This finding suggests that stereotypes concerning Muslims may have been influenced by the perception of radical Islamic groups, whose activities have been constantly reported in the media since 9/11. However, whether the media coverage of the events connected with radical Islamists, such as acts of terror, could have influenced the attitudes and linguistic behaviour of society remains debatable.

In order to find the most commonly searched collocations on the Internet, a Google Suggestions search was used, as previously described. The results were provided in the form of a list of expressions beginning with certain phrases. The sample output of the data is shown in the Figure 1. The search is case-insensitive, thus all the names of racial and ethnic minorities were entered in lower-case letters.

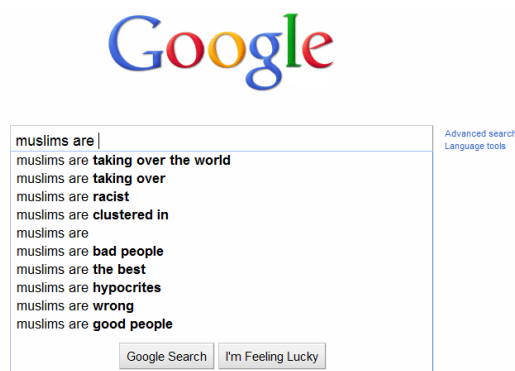


Figure 1. An example of the results of a Google Suggestions search – a list of the most commonly searched expressions beginning with a phrase “Muslims are” (www.google.com).

Other searched phrases and results were:

- 1) “Arabs are”, with the results, from the most to the least common: *racist, White, democracy’s new pioneers, Semites, lazy, revolting, Caucasian, hamites, Muslim*;
- 2) “why are Arabs”: *so rich, protesting, so rude, so racist, revolting, so backwards, so smart, in Africa, and Israelis fighting, White*;
- 3) “why are Muslims”: *so angry, circumcised, so hated, so sensitive, bad, not allowed to eat pork, so backward, so annoying*.

The outcome clearly indicates that this minority is regarded by many as angry and revolting. The concern with violent events reported in the media may be noted in search results, such as *taking over the world* or *and Israelis fighting*. The portrayal of Muslims as racist and oversensitive may be linked to the protests in the Arab world against the publication of drawings of Mohammed in newspapers. Many of the collected expressions point to the fact that perceptions of the Muslim minority are rather negative, with cultural and religious differences and terrorism being the main concerns of the English-speaking majority.

The search for collocates of the terms *Muslim* and *Islamic* in COCA and BNC helped to form a list of collocates in order of decreasing frequency. Selection of these collocates with their respective MI scores is shown in Tables 1 and 2.

Table 1. Assorted collocates of Muslim in the Corpus of Contemporary American English and British National Corpus

Collocate	MI score in COCA	MI score in BNC
fundamentalists	8.14	9.94
extremists	7.56	8.90
fundamentalist	6.97	8.29
militants	6.50	8.68
extremist	6.40	8.90
militant	5.98	6.49
radical	4.61	4.35
terrorists	4.45	5.03
terrorist	3.24	-
guerrillas	-	6.06

Table 2. Assorted collocates of Islamic in the Corpus of Contemporary American English and British National Corpus

Collocate	MI score in COCA	MI score in BNC
fundamentalism	10.05	10.78
jihad	9.93	12.47
extremism	9.78	6.71
fundamentalists	9.23	11.06
militants	8.77	9.38
fundamentalist	8.60	10.97
militancy	8.59	5.43
extremists	8.46	8.24
radicalism	8.41	5.28
militant	8.17	7.05
extremist	7.56	6.83
radicals	7.53	7.44
radical	6.68	4.80
terrorists	5.66	3.80
terrorism	4.87	3.87
terrorist	4.70	4.64
terror	3.62	-

In both researched corpora, the most frequent negative collocation with the term *Muslim* is *fundamentalists*. Other forms of this lexeme also frequently collocate with both search terms. All collocations of *Islamic* and *Muslim* and the notion of fundamentalism are characterized by very high MI scores, which mean that these collocations are fairly common in the English language. Most of the frequent collocates are associated with fighting and aggression (*militant*, *terrorists* or *guerillas*). Interestingly, a frequent collocate for *Islamic*, having high MI score in both corpora, is *Jihad*. It is the third most frequent negative collocate in the Corpus of Contemporary American English and the first most frequent negative collocate in the British National Corpus. This suggests that the notion of religious war influences perceptions of Islam. A noticeable difference is that while collocates *terror* and *terrorist* are fairly frequent in COCA, there was no such collocate found in BNC. This may stem from the fact that the perceptions of Muslims in the United States are affected by the threat of homeland terrorism, which is still one of the major fears in American society. It is possible that in the United Kingdom, perceptions of Muslims are more greatly affected by certain international events that emphasize the militancy displayed by Muslims in their countries of origin. On this basis, it may be concluded that different events could shape the image of the Muslim minorities in the United States and the United Kingdom. However, in both countries, perceptions are rather unfavourable and point to the possible threat that Muslims pose to society.

4. Stereotyping of the Black Population

It is worth noting that the number of epithets targeting Black society found in the Racial Slur Database was 182, which is considerably more than the 36 targeting Hispanics or the 24 targeting Muslims. This may be linked to the fact that the Black minority is still more visible, making up 12 per cent of the US population. Moreover, although they only make up 2 per cent of the population of Great Britain, they are constantly presented in the media, especially in connection with crime and poverty issues. This minority also has the longest history of being subjected to racial prejudice in the United States. As may be expected, there are still many sentiments connected to the slavery era, and epithets clearly alluding to the oppression of the Black minority are still present in the discourse. The analysis of epithets referring to the Black minority shows that most of the terms (over 50) are associated with appearance, which yet again supports Janet Ruscher's (2001: 20–21) theory. The physical characteristics used to create stereotypes include not only skin colour but also the hair and facial features, which suggests that appearance dissimilarities have a great influence on human thought and hence on linguistic choices. Another notable group of epithets is those mocking the pronunciation or syntax commonly used by African Americans. Interestingly, the long presence of the Black minority in American society gave rise to the significant number of epithets associated with either the depiction of Blacks in popular culture or their cultural contributions, such as music.

In the Google Suggestions search for collocations involving terms describing the Black minority, two words were chosen because members of this minority are commonly described as *Blacks* or *African Americans*. The results were as follows:

- 1) for the search phrase “African Americans are”: *not African, not Black, what percentage of the population, mixed, lazy, Hebrews, better athletes, obese, ignorant*;
- 2) for the search phrase “Blacks are”: *racist, lazy, fools, the most racist, loud, the real Jews, bad tippers, given citizenship*;
- 3) for the search phrase “why are African Americans”: *poor, Black, lighter, more athletic, at risk for hypertension, lighter than Africans, tall, fearful of water, so fast, so athletic*;
- 4) for the search phrase “why are Blacks”: *so lazy, so loud, so fast, so rude, hated, democrats, more athletic, so racist, so fat, poor*.

A slight difference between the search results for the politically correct term *African Americans* and the slightly less formal *Blacks* may be noted. The search for *African Americans* yields more neutral results, which may be accounted for by the assumption that racially prejudiced people would not utilize such a politically correct phrase. The search for *Blacks* reveals that the most persistent stereotypes are those of laziness, poverty and bad behaviour.

The most frequent collocates of the terms *African Americans* and *Blacks* were found and compared for COCA and BNC. The term *African Americans* collocates with neutral terms in the British corpus. However, in the American English corpus, some of the frequent collocates were concerned with the common stereotypical view of the African American minority. There were only two negative collocates found: *poverty*, with a mutual information score of 3.39, and *low-income*, with an MI score of 4.82. The term *African American* is, however, politically correct and does not seem to activate negative stereotypes of the Black minority. Therefore, a search for collocates of the term *Blacks* was conducted. The results are presented in Table 3.

Table 3. Assorted collocates of Blacks in the Corpus of Contemporary American English and British National Corpus

Collocate	MI score in COCA	MI score in BNC
uppity	8.26	9.74
unemployed	6.15	3.95
inferior	5.30	6.15
violence	5.06	3.44
slavery	4.54	5.83
impoverished	4.43	-
slaves	4.42	5.06
low-income	4.29	-
poor	3.76	2.76
killing	2.36	3.91
arrest	0.21	4.80

The most frequent negative collocation of this term in both analysed corpora is *uppity*, which may be considered an extremely racist term. While, in the United States, Black people are associated mostly with terms pointing to their economic status, namely high levels of poverty (displayed in such collocates as *poverty* and *low-income* and *impoverished*), in the United Kingdom, collocates show concern with the issues of crime and violent behaviour within the Black minority (common collocates such as *arrest* or *violence*). Moreover, in the United States, Black people are still discussed in the context of slavery, which may contribute to the portrayal of Black Americans as a passive and inferior group. However, it is significant that the number of these negative collocates is considerably lower than the number of negative collocates applied to Muslims. Thus, it may be inferred that Black citizens do not stir up negative emotions to the extent that Muslims do.

5. Stereotyping of the Hispanic Population

Hispanic and Latino Americans are people originating from countries in Latin America or Spain and constitute the largest minority group in the United States. The data obtained from the Racial Slur Database clearly points to the link between widely held stereotypes of Latino people and the epithets used in naming them. Not surprisingly, most of the slurs are related to the issue of immigration, *e.g.*, '*border-hopper*'. Although illegal immigrants comprise only a small minority of all Hispanics in the United States, images of Latinos attempting to illegally cross the border are pervasive in the media. Since there are no clear physical characteristics that distinguish Hispanics from non-Hispanic White Americans, there was only one epithet relating to such

differences. The slurs referring to Hispanics could not be based on cultural differences, as in the case of the Black and especially the Muslim minority, partly due to the fact that in the case of this minority, successful assimilation into the dominant society was observed (Dinnerstein and Reimers, 2009: 189). The epithets alluding to customs are focused on the diet of people of Hispanic origins and there are very few epithets referring to the occupations of poor immigrants. As the Hispanic wave of immigration is quite a recent event, only one epithet alluding to a famous Hispanic person was found. It is highly likely that with the progress of Latino people, more epithets of this kind may be used in the close future.

Google Suggestions results were studied in a similar fashion as with Muslim and Black stereotypes. The two terms for which the search was conducted were *Hispanics* and *Latino*. The results were as follows:

- 1) for the search phrase "Hispanics are": *not White, racist, from, surging in Arizona, what race, from where, taking over, White, not a race, short;*
- 2) for the phrase "Latinos are": *one, back in the game, White, lazy, taking over, Caucasian, lousy lovers, not White, not a race;*
- 3) for the phrase "Why are Hispanics": *at greater risk for ptsd, more prone to diabetes, called Latinos, good at soccer, Latino Americans, so rude, racist, called Latin, always late, poor;*
- 4) for the phrase "Why are Latinos" : *so attractive, called Latinos, so short, so loud, always late, racist, so racist, considered White, discriminated against, not a race.*

Due to the fact that the majority of people labelled as Hispanics are Mexicans, a search was also conducted with this term, leading to results such as "Mexicans are": *racist, lazy, White, Native Americans, hot, rude, a band of illiterate Indians, annoying, short, Indians;* "Why are Mexicans": *short, so fat, so short, so rude, called beaners, brown, so hot, so lazy, so racist.*

The outcome indicates that the questions of the racial or ethnic affiliation and the "Whiteness" of Hispanic people are of great concern for the general public. Common stereotypical beliefs may also be observed. Rudeness, poverty and laziness are attributed to Latinos. Interestingly, there are also more positive images, namely the physical attractiveness of members of this group.

In the case of the Hispanic minority, the searches were made for the terms *Hispanics* and *Latinos*, which are used interchangeably. A search for collocates of the term *Mexicans* was also conducted. The results of the search are presented in table 4.

Table 4. Assorted negative collocates of Hispanics, Latinos and Mexican in the Corpus of Contemporary American English

Collocate	MI score for <i>Hispanics</i>	MI score for <i>Latinos</i>	MI score for <i>Mexican</i>
immigrant	4.58	6.30	5.94
immigrants	4.25	4.17	4.47
low-income	4.53	5.88	-
immigration	3.49	3.68	4.89
illegally	3.11	-	5.92
poor	3.00	3.10	1.91
poverty	2.79	3.32	4.31
gang	2.47	4.32	3.97
illegal	2.17	4.11	3.38
lazy	-	3.50	4.60

Collocation lists were compiled only for the results from the Corpus of Contemporary American English since in the British National Corpus, no negative terms associated with the search terms were found. The lack of negative collocations for the Hispanic minority in British English is in accordance with the fact that the Hispanic minority in the United Kingdom is not very visible and thus not stereotyped by the dominant society. The collocations of terms associated with the Hispanic minority are predominantly based on the notion of immigration. Moreover, the economic status of this minority plays a vital role in the stereotyping process. There are, however, slight differences between the ways in which Hispanics are portrayed. In the case of the term *Hispanics*, collocates refer to the issues of poverty and immigration. The terms *Latinos* and *Mexicans* are additionally commonly associated with crime, namely the existence of gangs. In general, it seems that the Hispanic minority is still perceived as a society of immigrants, often illegal, who are poor and may be a moderate threat to the dominant society, hence the collocation *gangs*. However, there is a significantly lower variety as well as mutual information scores of negative collocates pertaining to this group than to Blacks or Muslims. This is quite surprising given the fact that the Hispanic minority is the largest in the United States. It may be a result of the aforementioned rapid and full assimilation of this minority or the fact that neither their appearance nor customs differ greatly from those of the dominant group.

6. Conclusions

In order to uncover the racial and ethnic prejudice present in English-speaking societies, a multi-faceted study was conducted. Ethnic slurs were categorised according to their origins. Most of the epithets found are based on negative characteristics attributed to the members of the groups or to their appearance and customs. The majority of racial slurs included in this study are based on easy-to-spot characteristics, such as skin colour in the case of Blacks or the turbans worn by some Muslims. A significant fraction of epithets may also be closely linked to the portrayal of a minority in the media. In the case of the Black and Hispanic minorities, there are many epithets pertaining to poverty, reliance on welfare programmes, violence and, consequently, crime.

The second part of the analysis, the search for the most popular collocates with neutral terms denoting minorities in national corpora, was aimed at finding patterns that were observable in the dominant discourse. The search based on corpus linguistics methods revealed that many frequent collocations pertain to the negative traits attributed to a minority. One must consider that such collocations are not only the result of negative stereotyping but also influence human thought and help to spread stereotypes. It is clear that racist ideologies are quite pervasive and are still being reproduced, regardless of the channel of communication and the introduction of the concept of political correctness into the official discourse. What is more, the representation of minorities in the media may contribute to the formation of stereotypes.

Undoubtedly, much more research on the language of prejudice should be done in order to identify more racist and ethnic stereotyping in corpora. It would also be advisable to compare the portrayals of minorities with the image of the White majority to determine how vast the differences between them are. Such research should be also extended to encompass the identification of other structures used in the prejudiced discourse, *e.g.*, those mentioned by van Dijk (2000: 147). Moreover, involving other national corpora in such extended research would allow researchers to define the link between the status of a minority and stereotyping in language more precisely.

7. References

Primary Sources

Bobo, Lawrence & Kluegel, James R. (1997). Status, Ideology, and Dimensions of Whites' Racial Beliefs and Attitudes: Progress and Stagnation" [In:] Steven A. Tuch and Jack K. Martin (eds.) *Racial Attitudes in the 1990s*. Westport: Praeger: 100–111.

Brown, Penelope (2006). Cognitive Anthropology. [In:] Christine Jourdan, Kevin Tuite (eds.) *Language, Culture, and Society: Key Topics in Linguistic Anthropology*. Cambridge: 96–114.

- Dinnerstein, Leonard & Reimers, David M. (2009). *Ethnic Americans: A History of Immigration*. New York: Columbia University Press.
- Duranti, Alessandro (1997). *Linguistic Anthropology*. Cambridge: Cambridge University Press.
- Fitch, Kristine L. & Sanders, Robert E. (2005). *Handbook of Language and Social Interaction*. Mahwah: Lawrence Erlbaum Associates.
- Hunston, Susan (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Kövecses, Zoltan & Radden, Gunther (1998). Metonymy: Developing a Cognitive Linguistic View. *Cognitive Linguistics* (9) 1: 37–77.
- Lakoff, George (1987.) *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. Chicago: University of Chicago.
- Mahlberg, Michaela (2005) *English General Nouns: A Corpus Theoretical Approach*. Amsterdam, Philadelphia: John Benjamins.
- Manning, Christopher D. & Schütze, Hinrich (1999). *Foundations of Statistical Natural Language Processing*. Cambridge: The MIT Press
- Mittelberg, Irene; Farmer, Tomas A. & Waugh, Linda R. (2007). They Actually Said That? An Introduction to Working with Usage Data through Discourse and Corpus Analysis. [In:] Monica Gonzalez-Marquez, Irene Mittelberg, Seana Coulson, Michael J. Spivey (eds.) *Methods in Cognitive Linguistics*. Amsterdam, Philadelphia: John Benjamins: 19–52.
- Murphy, Gregory L. (2002). *The Big Book of Concepts*. Massachusetts: MIT Press.
- Rosch, Eleanor (1978). Principles of Categorization. [In:] Eleanor Rosch, Barbara B. Lloyd (eds) *Cognition and Categorization*. Hillsdale: Lawrence Erlbaum Associates: 27–48.
- Ruscher, Janet B. (1998). Prejudice and Stereotyping in Everyday Communication. [In:] Mark P. Zanna (ed.) *Advances In Experimental Social Psychology* (vol. 30) San Diego: Academic Press: 241–307.
- Ruscher, Janet B. (2001). *Prejudiced Communication: A Social Psychological Perspective*. New York: Guilford Press.
- Schneider, David J. (2004). *The Psychology of Stereotyping*. New York: The Guilford Press.
- Sinclair, John (2004). *Trust the Text: Language, Corpus And Discourse*. London: Routledge.
- Stangor, Charles & Schaller, Mark (1996). Stereotypes as Individual and Collective Representations. [In:] Macrae, Neil, Miles Hewstone, Charles Stangor (eds.) *The Foundations of Stereotypes and Stereotyping*. New York: Guilford: 3–40.
- Stubbs, Michael (1996). *Texts and Corpus Analysis*. Oxford: Blackwell.
- Teubert, Wolfgang (2005). My Version of Corpus Linguistics. *International Journal of Corpus Linguistics* 10 (1): 1–13.
- Tribble, Christopher (2010). What Are Concordances and How Are They Used? [In:] Anne O'Keeffe, Michael McCarthy (eds.) *The Routledge Handbook of Corpus Linguistics*. London, New York: Routledge: 167–193.
- Van Dijk, Teun, A. (1987). *Communicating Racism: Ethnic Prejudice in Thought and Talk*. Newbury Park: Sage.
- Van Dijk, Teun A. (1998). *Ideology. A Multidisciplinary Introduction*. London: Sage.

Van Dijk, Teun A. (2004). Racist Discourse. [In:] Ellis Cashmore (ed.) *Routledge Encyclopedia of Race and Ethnic Studies*. London: Routledge: 351–355.

Secondary Sources

Firth, John R. (1957). *Papers in Linguistics, 1934-1951*. London: Oxford University Press;181.

Mullen, Brian & Johnson, Craig (1993). Cognitive Representation in Ethnophaulisms as a Function of Group Size: The Phenomenology of Being in a Group. *Personality and Social Psychology Bulletin*, 19: 296–304.

Quinn, Naomi (2005). How To Reconstruct Schemas People Share. [In:] Naomi Quinn (ed.) *Finding Culture in Talk: A Collection of Methods*. New York: Palgrave Miller: 33– 81.

Internet Sources

Council On American-Islamic Relations Research Center (2006). *American Public Opinion about Islam and Muslims*. Washington.

http://www.cair.com/Portals/0/pdf/american_public_opinion_on_muslims_islam_2006.pdf (Retrieved 03.2012).

Google Trends <http://www.google.com/trends> (Retrieved 04.2012).

Méndez-Méndez, Serafin & Diane Alverio (2003). *Network Brownout 2003: The Portrayal of Latinos in Network Television News*. <http://www.nahj12.com/NAHJbrownoutreport03.pdf> (Retrieved 04.2011)

Teun A. van Dijk (2000). *Ideology and Discourse: a Multidisciplinary Introduction*.. English version of an internet course for the Universitat Oberta de Catalunya.

<http://www.discourses.org/UnpublishedArticles/Ideology%20and%20discourse.pdf>. (Retrieved 12.2012).

The Racial Slur Database <http://www.rsd.org> (Retrieved 12.2012).

US Census Bureau, American FactFinder. <http://factfinder2.census.gov/main.html> (Retrieved: 11.2012).