

Thug Breaks Man's Jaw: A Corpus Analysis of Responses to Interpersonal Street Violence

Andrew Brindle

St. John's University,

Department of Applied English,

499 Tam King Road, Tamsui, New Taipei City, Taiwan

E-mail: andrewbr@mail.sju.edu

Accepted on October 19, 2015

Abstract: This paper examines a corpus of online responses to an article in an online edition of the British tabloid newspaper *The Sun* describing an act of interpersonal street violence between two men. The article produced 190 responses from readers, which were collected and compiled into a corpus that contained 6,606 tokens. Employing a corpus-driven approach, the data was investigated by undertaking concordance analyses of keywords and collocates of those words. The data was further analysed by taking into consideration multimodal information such as user names and avatar images in order to examine the significance of stating gender in correlation with views expressed. The findings indicate that regardless of the negative depiction of the aggressor in the article, the assailant and his actions were defended by certain posters, and at times admired and praised, while the victim was criticised for his lack of fighting skills, and not considered as innocent. However, the data also revealed that other respondents rejected such violence, demonstrating a continuum of reactions among the tabloid readership who responded to the article. The study found a marked difference of stances between those who stated that they were male to those who did not. The paper concludes by discussing the hypothesis that masculine identity and specifically hegemonic masculinity is constructed from multiple identities. Furthermore, the importance of investigating and analysing online peer groups is emphasised as an invaluable source in comprehending aspects of social behaviour within contemporary society.

Keywords: masculinities; interpersonal violence; corpus; discourse; online peer groups.

1. Introduction

The Internet and other Web-derived data have become a vast resource for corpus linguistics and natural language processing. In this study, texts of computer-mediated communication (CMC) taken from a message board of an online edition of a British tabloid newspaper, *The Sun*, were built into a corpus and analysed. The study researches the responses readers posted to an article in the newspaper which detailed an act of interpersonal street violence between two men in which one man was seriously injured. This research utilises a corpus-driven approach in order to discuss the attitudes articulated by the posters towards the act of violence, which it is argued, reflects upon the virtual identity of the posters.

As a result of the anonymity and freedoms of time and space, virtual identity is thought of as more unstable, performed and fluid than 'real' identity (Benwell and Stokoe, 2006), yet such a definition has similar qualities to postmodern identity which is described as both constructed and discursive (Bauman, 2007). Thus, this analysis aims to not only highlight the posters' attitudes towards such violence, but furthermore, demonstrate traits of identity through discursive accomplishment.

The theme of the discussion board is centred on an act of violence between two men. A great deal of what is bad in the world, from genocide to interpersonal violence, is the product of men and their masculinities (DeKeseredy and Schwartz, 2005). Work by criminologists such as Anderson (1990) have argued that instances of interpersonal violence originate from strongly held values in the construction and defence of personal street status and that violence is a tool for both the formation of and the protection of self-image. Furthermore, Messerschmidt (2004) writes that among certain men violence is a core component of masculinity and a means of proving one's manhood. However, Winlow (2001) considers that street and pub fights function as a means for working-class men to actualise a masculine identity due to the loss of traditional industrial job opportunities in a postmodern society. Clearly, violence is one means by which certain men live up to the ideals of hegemonic masculinity; such practices may be learned through interactions with particular peer groups, or virtual peer groups, and understood as a form of social constructionism (Hall, 1996).

2. Data

The article which produced the data for this study was published in the online version of *The Sun*, a British tabloid newspaper, on January 8th, 2013. The article, found in the News section, was titled, "*Thug breaks man's jaw outside takeaway in unprovoked attack...because he was ginger*" below which were two pictures taken from CCTV footage, the first showing a larger man punching another man. The second photograph shows the smaller individual falling to the ground in the street. After six sentences of the article, a CCTV video clip of the attack is embedded into the page. Further down, there is another picture which depicts the larger man exiting a store and confronting the smaller man and a fourth photograph which shows the moment in which the smaller man was hit.

The article states that a man was attacked and left seriously injured in what is described as *an unprovoked attack*. The story contains a large proportion of direct quotes as the injured man describes the incident and the long period of physical and psychological recovery. He relates how he went into a pizza takeaway restaurant with his girlfriend

and was verbally abused before being physically attacked upon leaving. He was left unconscious with a badly broken jaw and needed three months to recover. In the article he is clearly depicted as a blameless victim, whereas the other man is presented as the guilty aggressor. The article states that the attacker was still being sought by the police at the time of publication.

The article produced 190 responses from readers containing a total of 6,606 words. If readers wished to comment on the message board, they would first have to create an account by either using an existing Twitter or Facebook account, or by creating a new account with the newspaper. The reader would have to submit a user name; an avatar could also be added. Using these two sets of information, 60 posters used male names or provided pictures of males, whereas only 4 posters indicated that they were female. The other poster provided user names and avatars which did not indicate gender.

3. Methodology

Once the corpus of reader responses was compiled, the first stage of the analysis consisted of a study of frequency data. From the word list ordered by frequency, it was possible to gain an understanding of aspects of the corpus which occurred often and therefore had the potential to demonstrate the lexical choices that the tabloid readers who responded to the article made, which could relate to the presentation of particular discourses or attempts to construct identity.

After a word list organised by frequency had been analysed, a word list arranged by keyness was observed. For this study, a reference corpus of general English was constructed which consisted of newspaper articles from the British newspaper *The Guardian*. Wordsmith Tools was set to perform a log likelihood statistical test for each word, which gave a probability value (p value). This value designates the degree of confidence that a word is key due to chance alone, the smaller the p value, the more likely that the word's presence in one of the corpora is not due to chance but the result of the author's choice to use the word consciously or subconsciously. Wordsmith uses a default of $p < 0.000001$.

Once the keywords lists had been studied, to gain a more comprehensive insight into the data, the keywords were observed in context by undertaking a concordance analysis. A concordance-based study is able to disclose a range of discourses; therefore the notions of semantic preference and semantic prosody are important concepts. Semantic preference is defined by Stubbs (2001: 65) as "the relation between a lemma or word form and a set of semantically related words", thus it is related to the notion of collocation. Semantic preference is the meaning which arises from the common semantic features of collocates of a given node word (McEnery et al. 2006: 84).

Semantic preference is linked to the notion of semantic prosody (Louw, 1993) where patterns in discourse can be

established between a word and a set of related words that indicate a discourse.

To gain further insights into the usage in context of the keywords, a study of collocates of those words was undertaken in which mutual information (MI) was utilised to calculate the strength of collocation. The MI score measures the degree of non-randomness present when two words co-occur. An MI score of 3 or higher can be considered to be significant (Hunston 2002: 71). However, one problematic issue with MI is that high scores may be achieved by relatively low frequency words, therefore this must be taken into consideration as words with a low frequency ought not to be considered as significant in spite of a high MI score.

Once this procedure had been completed using the whole corpus, a second corpus was compiled of texts by posters who indicated that they were male. This was achieved by observing the usernames or the avatars. If in each case the posters indicated that they were male, the text was added to the second corpora.

4. Findings

Frequency is one of the most fundamental concepts in the analysis of a corpus (Baker, 2006), which is able to provide insights which illuminate a range of themes. The twenty most frequent words in the corpus are as follows: *the* (289), *a* (203), *to* (176), *and* (152), *he* (117), *of* (103), *is* (102), *this* (90), *I* (86), *him* (84), *for* (78), *that* (74), *in* (69), *his* (59), *be* (58), *it* (53), *not* (53), *with* (52), *on* (49), *out* (49). It is apparent that the most frequent words in the corpus are grammatical words (function words). Such words belong to a closed grammatical class consisting of high frequency words such as articles, pronouns, conjunctions and prepositions. These groups of words do not necessarily provide insight to the discourses found within the corpus as most forms of language contain a high proportion of functional words. However, by taking into consideration the most frequent lexical words such as nouns, verbs, adjectives and lexical adverbs, a clearer notion of the discourses within the corpus is attained: *like* (41), *get* (40), *guy* (27), *victim* (25), *hope* (24), *someone* (24), *people* (23), *thug* (22), *ginger* (19), *punch* (19), *know* (18), *got* (17), *think* (17), *fight* (15), *man* (15), *prison* (15), *catch* (14), *scum* (14), *caught* (13), *coward* (13).

The second list presents a clearer picture of what the corpus is about. There are words associated with acts of violence (*victim*, *punch*, *fight*). Another group of words are used to describe a person negatively (*thug*, *scum*, *coward*). An additional aspect of the list is that *catch* and *caught* are both present, therefore the lemma CATCH is significant. When this is taken into consideration with *prison*, it can be seen that another theme is prominent. Beside the lemma CATCH, other verbs are also present (*like*, *get*, *hope*, *know*, *got*, *think*). Further analysis of collocational data and concordance lines will be needed to understand the context of these verbs within the corpus, although when the COCA corpus of general English (Davies, 2012) is referenced, it

can be seen that such verbs are also of high frequency in a corpus of general English.

By considering frequency beyond the single word further insights can be gained. The most frequent 3-word clusters are: *is going to* (7), *he is a* (6), *lock him up* (6), *him up and* (5), *on the wrist* (5). By observing the concordance lines for the most frequent cluster *is going to*, certain themes can be seen: (1) *suspended, you just know He is going to get off lightly* (2) *moment of madness and this kid is going to pay the ultimate price* (3) *punching someone. This place is going to the dogs at a rapid rate* (4) *the lad who threw the punch is going to have to deal with the 'victim's friends* (5) *going to be mugged or someone is going to attack me, so im always ready* (6) *kick someones head in . Apparently UKIP is going to sort this out....!* (7) *just a stupid no brain thug who is going to jail.*

The writer of line 1 predicts the assailant is not going to be severely punished, whereas in lines 2 and 7 the opposite prediction is made. Line 4 remains within the theme of punishment, but states that the attacker is going to have to face the victim's friends. In line 3, the poster describes the decline of society and social behaviour, which is similar to the sentiment found in line 6, which states that a British political party claims to have a solution for such a situation, although the use of the word *apparently* appears to contest such a claim.

The poster of line 5 describes the actions that he or she would take if attacked in similar circumstances, thus claiming to be more prepared to act in the instance of street violence than the victim was. Such examples highlight the diverse responses to the act of violence depicted in the article.

Of the six instances of *he is a*, five refer to the attacker. He is described as: *a self centred thug, a threat to the society and a trained fighter*. Although further analysis is necessary, the data demonstrates certain themes within the corpus; the assailant is condemned for his action and his fighting ability is discussed. Two clusters are evident which combine to produce the phrase: *lock him up and* followed by a phrase such as *throw away the key* demonstrate the punishment the posters consider the aggressor deserves. Another frequent cluster is part of the phrase: (*slap or smack*) *on the wrist*, thereby predicting that he will be dealt with lightly by the law. Thus it can again be seen, there are various reactions and opinions to the violence: the aggressor is condemned for his actions, that society is described as having poor moral standards, and that he will not receive adequate punishment for his actions.

It can be observed that by analysing frequency lists, discourses within the corpus may be highlighted. In the following section, keywords of the corpus will be discussed.

The keywords with the highest levels of keyness are as follows: *him* (113.53), *this* (86.86), *I* (69.8), *guy* (66.26), *someone* (58.89), *hope* (58.89), *thug* (53.98), *victim* (53.56), *punch* (46.61), *ginger* (46.61), *get* (46.10), *he* (44.93), *like* (42.27), *why* (37.08), *your* (35.69), *catch* (34.34), *scum*

(34.34), *don't* (31.88), *coward* (31.88), *out* (28.41). These words could be divided into three separate groups. There are functional words: *him, this, I, he, why, your, don't, out*, verbs: *hope, get, like, catch*, and nouns or adjectives: *guy, someone, thug, victim, punch, ginger, scum, coward*. It is not possible to present all the findings in this paper due to space restrictions, therefore certain keywords from each will group be selected for further analysis.

If the functional words are taken into consideration, *him* has the highest level of keyness and is the second most frequent among the keywords on the list. As this word is also most likely to be referring to one of the two male actors in this instance of street violence, further analysis may provide further insights as to how the two men are constructed. Therefore, the keyword *him* was studied in context by considering the concordance lines.

Of the 84 instances of *him*, 73 are referring to the attacker, of which 36 depict him negatively, 10 reference the victim and one refers to a person in a hypothetical situation, thus the aggressor appears to be the primary focus of the posters. If the collocates of *him* with the strongest levels of MI scores are calculated, the following list is provided: *suspended, teach, catch, example, sentence, years, really, throw, someone, prison*. This appears to indicate that within the corpus there is a dominant discourse associated with the attacker being caught and punished for his actions. This is confirmed when the word *him* is seen in context. A principle discourse focuses on the attacker being caught: *Catch him and jail him ASAP*. Another concordance line within the same semantic field describes the same notion more strongly: *Scum of Britain!!!PLEASE catch him*. Another example using a phrase which was found in the most frequent clusters is as follows: *Find that punk, lock him up and throw away the key*. Therefore it can be seen how the posters react to such acts of violence. Another discourse within the corpus denigrates the attacker, as the last example demonstrates with the term *punk*. Other examples include: *UK is full of scum like him! / No other word for him COWARD.*, and referring to him as a *mug brained idiot*. However, not all of the posters refer to the aggressor in such negative terms, nor do all the people who responded to the article believe that he should receive a prison sentence for his actions as the following examples indict: *but the other kid does square up to him / I doubt he wudda smacked him like that completely unprovoked / What if the 'victim' offered him out to begin with?* This appears to indicate that there are some writers who do not accept the opinion of the article and are willing to consider alternative scenarios for the event which took place, thus indicating that the response to acts of violence among the message board posters is not homogenous.

Another keyword which provides insights into the corpus is *punch*. There are 19 instances of this word, all of which are in the form of a noun. When analysed in context, opposing discourses are evident; 9 of the lines either defend the attacker or are appreciative of his fighting skills, whereas only 5 instances denounce his actions. Another 5 instances of *punch* were classified as neutral, neither defending nor denouncing the attack. Examples of

instances which depict the attack positively are as follows: *Boom! What a punch! / great punch*, and *the kid knows how to throw a punch*. Such examples are in contrast to the article which clearly denounced his actions. It can be seen that the writers of these examples value the act of violence regardless of the fact that it left one man seriously injured. As previously stated, the corpus does not contain a single discourse; other examples of *punch* in context are more condemning: *The guy who threw the punch is a bully / it was a dangerous cowardly sucker punch*. This brief study of *punch* demonstrates that both qualitative and quantitative analysis is necessary not only to discover discourses within a corpus, but also to comprehend their statistical significance.

Another keyword of interest is *victim*. There are 25 instances of this word; they all refer to the man who was left unconscious with a broken jaw. However, when the concordance lines are studied, it can be seen that a number of writers are using this word ironically when labelling this man as a victim of crime or cast doubt on the interpretation of events depicted by the newspaper. Seven of the writers do not consider the injured man to be blameless as the following examples demonstrate: *the person who hit the deck was not a victim / Looks like the 'victim' called the other guy out of the joint then got punched / I doubt that he's the complete victim he's made himself out to be*. These writers do not appear to accept the opinions of the newspaper nor the evidence provided by the link to the CCTV footage which clearly illustrates the assault. The writer of the second example places the word within quotation marks to emphasise the fact that it is doubted whether the person is in fact blameless. Other writers demonstrate a different opinion, as the following examples illustrate: *it looks like he had 20lbs over the victim / That victim could have been a brain op patient*.

4.1 Men Only Corpus

Once the analysis had been completed using the corpus of all the postings, a second corpus was compiled using posts in which the writers had indicated that they were male either via a username or avatar (64 posts from 60 unique profiles). This procedure was undertaken in order to observe if the stating of gender had a significant impact on the findings. This corpus contained 2,088 tokens; the ten most frequent lexical words were as follows: *like* (14), *get* (12), *victim* (10), *punch* (9), *think* (9), *ginger* (8), *guy* (8), *looks* (8), *people* (8), *attack* (6). When this list is compared with the complete corpus, it can be seen that *punch* is much higher, as the posters discuss the blow which broke a man's jaw. Secondly, both *hope* and *thug* are no longer present. This appears to signify that those who indicated that they are men refer to the aggressor as a *thug* less frequently, nor do they use the word *hope* as often. When *hope* is studied in the first corpus, it is most often used in phrases whereby the writer expresses a wish that the aggressor is caught and sent to jail. The weakening of such a sentiment in the second corpus is of interest. This pattern of differences is reinforced when the keywords with the highest levels of keyness are observed: *him* (42.29), *punch* (38.79), *victim* (36.65), *this* (35.23), *ginger* (34.48), *guy* (34.48), *your*

(32.37), *looks* (28.45), *I* (28.04), *coward* (25.85). *I* is now present on the list, which appears to indicate the posters who state that they are male express their opinions more frequently than those who do not. However, it must be noted that *coward* has a high level of keyness, reinforcing the notion that the posters are not a homogenous group with shared set of values.

When *punch* is observed in context, only one post refers to the assailant as a *bully*, the other instances of the word are used in phrases which demonstrate an appreciation of his fighting technique or describe it in more detail. The posters use *victim* either ironically or by stating that he was not a victim, only a minority labelled the injured man as such. The word *ginger* also demonstrates the lack of support for the injured man in the corpus; of the eight instances of this word, six are used by posters when casting doubt on his degree of innocence in the act of violence. When *looks* is observed in context, again a significant difference from the first corpus is seen, as only one instance of this word is used in a context which criticises the attacker. Again the majority of posters cast doubts on the innocence of the second man as the following phrases demonstrate: *It looks to me like the ginger lad goes out for a fight. / This isn't as innocent as it looks. The ginger lad walks out first. / It looks like the ginger fella took offense*. Thus there appears to be less condemnation of the attacker in the second corpus than in the first.

The data demonstrates that the writers who posted on the newspaper message board in response to the article are clearly of differing opinions. There are those who accept the views presented by the newspaper which condemn the assailant and his actions, clearly articulating how he should be punished as a consequence of his actions. Others use the incident to express an opinion the England has and is still experiencing a decline in social standards and morality, and furthermore that law enforcement is too lenient to effectively respond to such a situation. However, there is another statistically significant discourse within the corpus which is contrary to those which condemn or criticise the violence. In this discourse, violence is seen as something which is appreciated and respected, where the aggressor is not depicted as the guilty party and where the victim is not seen as blameless. This second semantic field becomes much more evident when the data is grouped according to those who indicated that they were men.

5. Discussion

This corpus linguistic study, which is based on online responses to a newspaper article, contains sociological and cultural components. As the data was collected from a single newspaper, it is not possible to state that the findings reflect upon a larger social group other than those who posted on the website. Furthermore, the data does not provide insights into the level of influence the article had on the readers, as the stance taken by the newspaper journalist may, or may not have affected the responses found in the data. However, attempts were made not to decontextualise the data as the language of the article and images found both in the article and the avatars of the

posters may have influenced certain stances taken by the writers. As the study was sociolinguistic in nature, the corpus was not annotated with a grammatical tagger, and due to the small size of data, it was possible to consider and review each post before deciding if it met the required criterion to be added to the second corpus. The effect of irregular spelling found in the data was minimal, as again through the small size of the corpus, it was possible to return to the source of the data to observe the context in which it was located.

As previously stated, the data for this study is narrow in scope, and therefore does not shed light on stances held by individuals outside of the particular target group. However, the findings are of relevance in the fields of identity construction, masculinity and violence. The findings have demonstrated a continuum of opinions and stances on the online message board in response to a specific act of violence. Such expressions of opinions and stances may be considered to be a reflection of an aspect of identity the writer constructs for himself or herself. Early CMC scholars described how the Internet liberated people from social constraints through there being a supposedly unbiased and non-prejudiced environment. The Internet was also believed to provide a measure of anonymity; however this perception has now appeared to have lessened due to the rise of social media networks in which any form of Internet activity is traceable and where users are aware of a degree of accountability regardless of the spatial distance when interacting on the Internet (Thurlow et al. 2004).

Only a very small number of posters indicated that they were female, in contrast to a much larger number who claimed to be male, although by using the information provided on the message board, a significant proportion of the posters provided no information related to their gender. When the whole corpus was analysed, it appeared that the principle discourses within the texts were critical of the violence which took place. However, when the second corpus was built containing posts by writers who indicated that they were male, the discourse which accepted this violence and questioned the degree of innocence of the injured man, became much more prominent.

By manipulating the data in this manner, it has been possible to focus on the responses of individuals who stated masculinity to be part of their identity. Masculinity can be defined as the trait of behaving in ways that society considers to be typical and acceptable for males. Masculinity, like gender, is constructed and therefore is something that has to be worked at. Boys and men have to prove their masculinity constantly (Kimmel 2001: 269). One form of proving masculinity for certain individuals is to condone violence, as the data has shown. Hegemonic masculinities (Connell 1995) are characterised as the variety of masculinity capable of marginalising and dominating not only women, but also other men. It is dependent on subordinate masculinities, since it must contradict them. However, in the data presented in this paper, it can be seen that subordination is achieved through violence, where the weaker injured male has been subordinated and rejected by other men, although this form

of action was not uniform, again demonstrating the lack of a homogenous response. Thus, one trait of hegemonic masculinity is the use and acceptance of violence against other men as a means to subordinate others.

For certain researchers, such as Whitehead (2002: 93-94), discourse is focused upon as a means to comprehend how men practice hegemonic masculinity and perform identity work. Masculine identities can therefore be understood as effects of discursive practices; they are fashioned within institutions and are historically constituted. An online message board such as the one used in this research is one location where identity may be constructed, practiced and maintained. One way that the gender order is maintained is by linking notions of appropriate and inappropriate gendered performances to different types of identities, and as the data has shown, one form of behavior which is seen to be acceptable by certain individuals, is the use of, or appreciation of violence.

According to Whitehead (2002: 33-34), the notion that masculinity is a singular rather than multiple identities has been viewed as problematic, particularly where gender identities and power relations are contextualised practices. In order to comprehend the diversity of masculinities, it is necessary to study the relations, such as subordination and dominance, between the different forms of masculinity. These relationships are constructed through practices that may intimidate or exploit others (Kiesling 2006: 118). The data has shown how these complex positions regarding responses to violence interact. Masculinity is not a fixed trait, but a social process dependent upon restatement, and which, in various forms, involves language, thereby centrally situating linguistic issues in the theorising of gender. Men who heavily invest in a particular masculinity will attempt to communicate in a manner particular for that specific trait (Moita-Lopes 2006: 294). Masculinities are not displaced from a social context, but embedded and implicated in the lives of men.

Responses to acts of violence may be considered by certain individuals to be a tool for both the creation of and defence of self-image. Using corpus linguistic methods, the data has highlighted discourses which demonstrate that a wide range of stances exist, which in turn signifies the plurality of masculinity. Analysis of the data has shown that there exists a subculture of violence, whereby acts of aggression are respected and esteemed, and therefore a means to construct a masculine identity.

Researchers such as Messerschmidt (2004) argue that with the loss of traditional industrial job opportunities and the shift towards a service-based economy, certain working class men have found new means of establishing masculinity; violence and street fights are one means of doing so. Such a form of masculinity emphasises toughness and a willingness to fight and defend oneself in the face of perceived threats or challenges by other males. The findings have shown that all men do not respond to violence in the same way. This will reflect upon an understanding of what men are and the consequences of acts such as the one focused upon in this study. However, it has been shown that for certain men, violence is considered a means for

validifying masculinity through peer support which encourages and legitimises acts of aggression. Hegemonic masculine discourses and practices, such as violence, may be learned through interactions, both virtual and face to face, which justify the relevance of studying online communication.

Although researchers such as Winlow (2001) consider violence amongst males to be a consequence of the destabilising effects of postmodernism, others such as Pinker (2011) describe how violence has been a constant trait throughout human history. Therefore, it may be argued that violence and aggression by men is more closely linked to aspects of patriarchal and hegemonic masculinity (Connell, 1995) than it is with social responses to the effects of postmodernism.

6. Conclusion

The Internet is a location where individuals may construct identity by expressing stances and through interactions with other Internet users. Whenever an individual interacts in a social environment, an aspect of their identity is revealed, and as identity construction and maintenance is a continual process, further construction takes place; this also applies for an online environment. The identities that individuals construct and the interactions they make on locations such as message boards may not necessarily be totally reliable or accurate.

However, Wiszniewski and Coyne (2002) argue that regardless of the reliability of the interaction or identity construction, a reflection of the authentic identity is formed which will reveal an aspect of the user's identity.

This paper has demonstrated that by employing a corpus linguistic approach, multiple expressions of identity and identity construction on the Internet may be studied. Discourses of violence and masculinities have been discussed, and the continuum of responses to violence observed and analysed. The results indicate that for certain individuals, violence and aggression are esteemed character traits, while others rejected and condemned them, thus confirming the notion of multiple masculine identity traits rather than a singular stereotypical construction. In addition, it has been shown that posters who state that they are men are more likely to regard interpersonal violence as an acceptable trait of masculinity.

Furthermore, the study has demonstrated that Web-derived data may be collected and filtered in various ways using contextual information in order to shed light on sociolinguistic and identity traits of particular target groups.

References

1. Anderson, E. (1990). *Streetwise: Race, Class and Change in an Urban Community*. Chicago: University of Chicago Press.
2. Baker, P. (2006). *Using Corpora in Discourse Analysis*. London: Continuum.
3. Bauman, Z. (2007). *Liquid Times*. Cambridge: Polity.
4. Benwell, B. & Stokoe, E. (2006). *Discourse and Identity*. Edinburgh: Edinburgh University Press.
5. Connell, R. W. (1995). *Masculinities*. Berkeley; Los Angeles, University of California Press.
6. Davies, M. (2012). COCA: Corpus Of Contemporary American English. <http://corpus.byu.edu>
7. DeKeseredy, W.S. & Schwartz, M.D. (2005). Masculinities and Interpersonal Violence. In M.S. Kimmel, J. Hearn, R.W. Connell (eds.) *Handbook of Studies on Men & Masculinities*. Sage Publications: London (pp. 353-366).
8. Hall, J.K. (1996). Who needs "identity"? In S. Hall and P. du Gay (eds.) *Questions of cultural identity*. London: Sage, pp. 1-17.
9. Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
10. Kiesling, S. F. (2006). Playing The Straight Man. In D. Cameron & D. Kulick (eds.) *The Language and Sexuality Reader*. Oxford, Routledge. pp. 118- 131.
11. Kimmel, M. S. (2001). Masculinity as Homophobia: Fear, Shame and Silence in the construction of Gender Identity. In S. M. Whitehead & F. J. Barrett (eds.) *The Masculinities Reader*. Cambridge, Polity. pp. 266-287.
12. Louw, B. (1993). Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In Baker, M., Francis, G. & Tognini- Bonelli, E. (eds.) *Text and Technology: In honour of John Sinclair*. Philadelphia and Amsterdam: John Benjamins, 157-176.
13. McEnery, T., Xiao, R. and Tono, Y. (2006). *Corpus- Based Language Studies: An advanced resource book*. London and New York: Routledge.
14. Messerschmidt, J. (2004). *Flesh and Blood: Adolescent Gender Diversity and Violence*. Lanham: Rowan and Little Field.
15. Moita-Lopes, L. P. (2006). On being white, heterosexual and male in a Brazilian school: multiple positionings in oral narratives. In A. de Fina, D. Schiffrin & M. Bamberg (eds.) *Discourse and Identity*. Cambridge: Cambridge University Press.

16. Pinker, S. (2011). *The Better Angels of Our Nature: Why Violence Has Declined*. New York: Viking.
17. Stubbs, M. (2001). *Words and Phrases*. Oxford: Blackwell.
18. Thurlow, C., Lengel, L. & Tomic, A. (2004). *Computer Mediated Communication: Social Interaction and the Internet*. Thousand Oaks, CA: Sage Publications.
19. Wiszniewski, D. & Coyne, R. (2002). Mask and Identity: The Hermeneutics of Self-Construction in the Information Age. In K. A. Renninger & W. Shumar (eds.) *Building Virtual Communities* (pp. 191-214). New York, New York: Cambridge Press.
20. Whitehead, S. M. (2002). *Men and Masculinities: key themes and new directions*. Cambridge, Polity Press.
21. Winlow, S. (2001). *Badfellas: Crime, Tradition and New Masculinities*. Berg: Oxford.