SOCIAL MEDIA, HATE SPEECH AND CONFLICT: INTERPLAY OF INFLUENCES

CHINWE ELIZABETH UZOCHUKWU (PhD)
Department of Mass Communication
Nnamdi Azikwe University, Awka- Nigeria
ce.uzochukwu@unizik.edu.ng
+2348037917259

&

EKENE GODFREY OKAFOR
Department of Mass Communication
Nnamdi Azikiwe University, Awka- Nigeria
ekeneokar@gmail.com
+2348068683384; +2348185725259

ABSTRACT
Given the pronouncement of a new bill by the Nigerian senate to douse the increasing hate speech that results in conflict occasionally fuelled by the social media, this work aims at examining the interplay between hate speech, social media and conflict in the society. The design adopted in the study is the Critical Discourse Analyses (CDA). Fifty-three textual documents downloaded from the social media comprising speeches made by some Nigerian personalities were analyzed. Hinged on the spiral of silence theory, the study considers the effect of hate speech on the formation of public opinion by some people especially those who form their opinions from the dominant view. While the study found the existence of hate contents on the social media, the extant literature shows that measuring hate speech requires knowing the hate words or hate targets priori and that the description of hate speech tends to be wide, sometimes extending to embody words that are insulting of those in power or minority groups, or demeaning of individuals who are particularly visible in the society. As the study also revealed, while hate speech may be prone to manipulation at critical times such as during election campaigns, accusations of promoting hate speech may be traded among political opponents or used by those in power to curb opposition and criticism, suggesting the need for intermediaries to advance the fight against hate speech because of the tendency of negative opinion formation by those exposed to hate messages given that some efforts are motivated by the impulse to ban hate speech as it can provoke pain, distress, fear, embarrassment and isolation to individuals.

Keywords: Social Media; Hate Speech; Conflict; Violence

INTRODUCTION
The growing concern over the spate of violence in Nigeria has necessitated the enactment of a bill on hate speech. The International Convention on Civil and Political Rights, Article 20 (2) defines hate speech as any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence. A new bill on hate speech is being enacted by the Nigerian Senate. The bill proposes that any person found guilty of any form of hate speech that results in the death of another shall die by hanging upon conviction. The bill though yet to be passed into law, sponsored by, Senator Aliyu Sabi Abdullahi (APC, Niger) arguably seeks the establishment of an 'Independent National Commission for Hate Speeches', to enforce hate speech laws across the country. The bill states: "A person who uses, publishes, presents, produces, plays, provides, distributes and/or directs the performance of any material, written and/or visual, which is threatening, abusive or insulting or involves the use of threatening, abusive or insulting words or behaviour, commits an offence…''

In addition to the various avenues for the propagation of conflict and hate speech in the Nigerian society, the social media had in recent times taken the center stage by making conflict situations more personalized, reactive and ravaging unlike the conventional media. Social war which exists on the social media among users interpersonally had also become conspicuous; this instigates conflict in the larger society given the peculiarity of the social media at allowing the spread of very provocative information. Indeed, the prevalence of hate speech and foul language on social media bordering on political and national issues, and even social interaction in Nigeria, especially on Facebook, Twitter, YouTube and LinkedIn is becoming worrisome. Apart from undermining the ethics of journalism profession, it contributes in bringing disaffection among tribes, political class, and religion or even among friends in the society (Alakali, Faga & Mbursa, 2016).

With the increased diffusion and use of the social media among people, it is common these days to find hateful messages that attack a person or a group because of their nationality, race, religion or gender on the internet (Taylor, Peignon & Chen, 2018). This is because the social media profoundly allows for free flow and exchange of all kinds of information. Although the terms of service for the social media platforms sometimes provide guidelines on what content is prohibited, which should shape the automatic filtering tools of these platforms (Taylor, Peignon & Chen, 2018), users continually abuse the platform. Hence, one of the major bane of the social media has been the propagation of hate speech and abusive content which creates an uncomfortable atmosphere that can inflict a damaging financial and social cost on both the social network and the victims alike,

The Nigerian public can rightly be said to have been inundated with negative media usage such as character assassination and negative political campaigns given that citizens become content providers and journalists or editors are not just supporters of particular political beliefs but play a fundamental part in setting national agendas which leaves much to be desired (Alakali, Faga & Mbursa, 2016). In Nigeria, the quest for power and the desire of politicians to win elections at all cost is overwhelmingly strong and these issues converge in shaping societal discourse by ardent users of the mainstream media and social media networks, thereby influencing the exchange of abusive content on the social media. This work therefore looks at the connection between the social media, conflict and hate speech to determine the extent to which the social media either enhances or fuels the display and propagation of hate speech and conflict. The work is therefore set to address the following research questions:
1. To examine the existence of hate speeches on the social media
2. To identify some hate speeches by prominent Nigerians as discussed on the social media
3. **To analyze how hate speech could fuel conflict in the Nigerian society**

REVIEW OF LITERATURE
THE CONCEPT OF HATE SPEECH
Hate speech according to Neisser (1994, p 337), refers to "all communications (whether verbal, written, symbolic) that insults a racial, ethnic and political group, whether by suggesting that they are inferior in some respect or by indicating that they are despised or not welcomed for any other reasons". Kayambazinthu & Moyo (2002) refer to hate speech as "war waged on others by means of words" which is particularly true when it comes to hate speech on social media networks because online hate speech is mainly characterized by the use of words and symbols. A hate speech which is a type of terrorism is popularly defined as the unlawful use of violence or intimidation against individuals or groups, especially for political aims (Osibanjo, 2017). Alakali, Faga & Mbursa (2016) citing the UN Committee on the Elimination of Racial Discrimination (2013, p. 4) further identifies hate speech as follows:

(a) All dissemination of ideas based on racial or ethnic superiority or hatred, by whatever means;
(b) Incitement to hatred, contempt or discrimination against members of a group on grounds of their race, colour, descent, or national or ethnic origin;
(c) Threats or incitement to violence against persons or groups on the grounds in as in (b) above;

(d) Expression of insults, ridicule or slander of persons or groups or justification of hatred, contempt or discrimination on the grounds in (b) above, when it clearly amounts to incitement to hatred or discrimination; and

(e) Participation in organizations and activities, which promote and incite racial discrimination.

With regards to what motivates hate speech, many scholars have identified several factors including: lack of tolerance, political clashes, discrimination, enmity and the openness of social media. For instance, even before the emergence of social media, Spiegel (1999 p. 375) predicted that the internet will be another communication tool for racists and "hate-mongers" to spread their messages, which was considered by Nemes (2002 p.193) as a very important channel for those spreading hate messages. Witschge (2008 p.75-92) however, presented a more balanced understanding of the Web's potentials within political communication that whether the Web enables deliberation or not, it depends on how people utilize the opportunities provided online because of the presumed horizontal, open and user-friendly nature of the social media which affords opportunities for greater participation in the public sphere. This further subscribes to Dalhberg's (2001) counter argument that the Web might facilitate abusive postings and even contribute in silencing some voices.

According to Olga & Roiha (2016) "Hate speech' includes every stance purporting to jeopardize the rights of an ethnic, religious or national group, in clear violation of the principles of equal dignity of and respect for the cultural differences among human groups". Another concept with relevance to the present study is that of "cyber hate", which essentially refers to "hate speech online" and defined by the Anti-Defamation League as "any use of electronic communications technology to spread anti-Semitic, racist, bigoted, extremist or terrorist messages or information. Furthermore they maintain that these electronic communications technologies include the Internet (i.e., web-sites, social networking sites, 'Web 2.0' user-generated content, dating sites, blogs, on-line games, instant messages, and E-mail) as well as other computer- and cell phone-based information technologies (such as text messages and mobile phones)". Cyber hate towards an individual, not related to his or her identification with a particular group, tends to be referred to as "cyber-bullying" (Council of Europe, 2012), hence hate speech attempts to stigmatize, marginalize, ridicule, defame and intimidate people vulnerably. In this regard, Waldron (2012) emphasizes that the harm hate speech causes to the persons targeted is often forgotten in the debates on freedom of speech.

Waldron (2012) identified two types of messages that can be conveyed by hate speech. The first message is aimed at the targeted group, trying to dehumanize and diminish them. The second message is sent to others with similar views, aiming to reinforce a sense of a group under threat from "the others" and unite this group of like-minded persons. In this sense, hate speech both divides and unites at the same time (UNESCO, 2015), making a distinction between who is inside and outside of the desired society.

Olga & Roiha (2016) have identified specific characteristics and challenges of hate speech as given by (UNESCO, 2015) to include:

1. Permanence: Hate speech can remain online for long periods of time and in different formats across different platforms, and can be repeatedly linked. In this sense, the architecture of the platforms influences how long topics "stay alive", e.g. Twitter is built around the idea of "trending topics" which may facilitate a quick and wide dissemination of hateful messages, however, if topics are ignored, discussion rapidly fades; Facebook on the other hand, provides the opportunity for longer lasting discussion threads.

2. Itinerancy: Even when content is removed, it may be revived somewhere else, either on the same platform with another name or in other online spaces.

3. Anonymity or pseudonymity (false names): With the possibility to be anonymous, people tend to feel more comfortable expressing hate, as they won't be discovered or will have to deal with any consequences. Anonymity, especially on social media, may also be an obstacle to prosecution.

4. Transnationality: The transnational reach of the Internet enhances the effects of hate speech and also poses complications regarding legal mechanisms for combating online hate speech.

Although hate speech expressed online has these specific characteristics, it is a fallacy to adapt the digital dualist assumption that anything happening online is separate from the offline, as argued by Jurgensen (2011) "Social media has everything to do with the physical world and our offline lives are increasingly influenced by social media".

Although there has been an increase in widening of the concept of Hate speech and identifications within the online platform, with respect to both hate speech classification and the detection of extremist communities (Taylor, Peignon & Chen, 2018), opine that Hate Speech (HS) can be difficult to define as there are some who argue that restrictions on what constitutes Hate Speech are in fact violations of the right to free speech. Hate speech is also any offense motivated, in whole or in a part, by the offender's bias against an aspect of a group of people and under this definition, online hate speech may not necessarily be a crime, but still harm people and the offended aspects can encompass basic hate crimes, such as race, religion, disability, sexual orientation, ethnicity, or gender, which may also include behavioral and physical aspects that are not necessarily crimes (Leandro, Mainack, Denzil, Fabr´ıcio & Ingmar, 2016). Hence, hate speech lies in a complex nexus with freedom of expression, group rights, as well as concepts of dignity, liberty, and equality (Gagliardone, Gal, Aleves & Martinez, 2015). For this reason, any objective definition (i.e., that can be easily implemented in a computer program) can be contested.

Hence, there is no common international definition of the concept of hate speech, but rather several definitions exist in parallel. In legal terms, hate speech tends to refer to "expressions that advocate incitement to harm […] based upon the targets being identified with a certain social or demographic group" (UNESCO 2015). In the public debate, however, or related to issues of awareness, broader definitions tend to be applied. This is the case, for example, in the commonly used definition by the Council of Europe where "hate speech" is understood as "covering all forms of expression which spread, incite, promote or justify racial hatred, xenophobia, anti-semitism or other forms of hatred based on intolerance, including: intolerance expressed by aggressive nationalism and ethnocentrism, discrimination and hostility against minorities, migrants and people of immigrant origin" (Council of Europe, 1997).

CONFLICTS AND SOCIAL MEDIA

At the National Economic Council Security retreat held at the presidential villa in Abuja on the 17th of August, 2017, the Vice President of Nigeria, Yemi Osinbajo as reported by Ayitogo, (2017), stated that the Nigerian government had decided to draw a line against hate speech, it will not be tolerated, it will be taken as an act of terrorism, all of the consequences will follow. He further surmised that the intimidation of a population by words or speech is an act of terrorism that will no longer be tolerated by the President Muhammadu Buhari administration which had reportedly catered for the security and welfare of Nigerians where media practitioners compromise their role as peacemakers and promote disunity, ignite crises and trigger hatred among members of the society (Ali 2013: 1). This practice seem to have normalized the use of hate speech as the media serve as the platform through which people castigate one another using vulgar or provocative language which are often discussed and debated on the social media. In the process the journalists consider the gains from the politicians as against truthful and peace journalism. Indeed, the press fell to the trap of reporting hate speech by quoting directly from interviews, press statements, advertorials and sometimes from alleged online sources. A case in point is the 2015 general elections where popular media outlets in Nigeria, like AIT, Channels, Thisday, Vanguard and The Nation inter alia were flushed with campaigns by several political parties displaying cross abuse of the right of free speech including hate speech and other types of foul language (Olowojolu 2016, p. 8). This hate speech and foul language in Nigeria are mostly inclined to religion or politics (Alakali, Faga & Mbursa, 2016).

We recall how the social media played a critical role in "mobilization, empowerment, shaping opinions and influencing political debates and influencing change" during the Arab Spring by facilitating communication

and interaction among participants of political protests and how protesters used the platform to organize demonstrations (both pro- and anti-governmental), disseminate information about their activities, and raise local and global awareness of ongoing events and how the governments used social media to engage with citizens and encourage their participation in government processes to monitor internet traffic or blocked access to websites or the entire internet (Salem & Mourtada, 2011).

ADDRESSING THE EFFECTS

Hate speech and foul language is reportedly prevalent on social media platforms in Nigeria even though there seems to be an adequate legal provisions to curb the phenomenon (Alakali, Faga & Mbursa 2016), which are often times not imbibed thereby leading to conflicts. This challenge manifests in a number of variations given that online hate speech has been rapidly recognized as a serious problem by the authorities of many countries, leading to reactions by several organizations and agencies. For example, the Council of Europe supports an initiative called "No hate speech Movement" and UNESCO had reported a study by Gagliardone, Gal, Aleves & Martinez (2015) entitled "Countering Online Hate Speech" aimed at helping countries deal with the problem. This is because promoting hate speech and foul language on social media has moral and legal consequences in the Nigerian society and this new trend of disseminating and publishing information has undermined the ethics of journalism profession and has negative implication on social media networks in Nigeria (Alakali, Faga & Mbursa 2016). It is also not surprising that most, if not all, existing efforts in this field are motivated by the impulse to ban hate speech in all forms given the dangers it poses on human personality and privacy.

Leets (2002, p.223) holds the view that hate speech violates the individual's dignity, resulting in humiliation, distress and psychological or emotional pain. Similarly, Nemes (2002 p.220) further submits that hate speech can provoke pain, distress, fear, embarrassment and isolation to individuals. While hate speech towards groups of people can bring inequality problems and isolation, it creates the feeling of fear and discourages them from participating in the community and expressing their opinions. Adding to the argument, Nielsen, (2002 p.265-280) avers that the degradation and humiliation brought by hate speech can silence the 'victims' and therefore reinforce existing hierarchies in society; while Parekh, (2006 p.213) says it can also lead victims to become aggressive and dangerous. It is evident from the above that hate speech is harmful and as such, should be curtailed, especially on the social media.
However, scholars like Cornwell & Orbe (1999) hold divergent views that any attempt to limit hate speech would result in undue censorship which was reinforced by Brinks (2001), who stated that the regulation of hate speech might bring equality but it would affect liberty. Similarly, Downs & Cowan (2012 p.1354) hold the view that "if speech is restricted, it silences those who may benefit largely from its expression". Post, Hare, & Weinstein (2009, p.123) observed that hate speech regulation imagines itself as simply enforcing the given and natural norms of a decent society but from a sociological or anthropological point of view, law actually only enforces the mores of the dominant group that controls its content.

According to Gagliardone, Gal, Aleves, & Martinez (2015 p.13-15) online hate speech is not essentially different from similar expressions found offline; however, there are some specific characteristics as well as challenges unique to online content and its regulation. They summarized these characteristics as permanence, itinerant, anonymity or pseudonym and transnationality. On permanence, hate speech can remain online for long periods of time and in different formats across different platforms, and can be repeatedly linked. In this sense, the architecture of any particular platform influences how long topics 'stay alive'. For instance, Twitter is built around the idea of trending topics, which may facilitate quick and wide dissemination of hateful messages, however, if topics are ignored, discussion rapidly fades; Facebook on the other hand, provides the opportunity for longer lasting discussion threads. Notwithstanding, online hate speech content may particularly be itinerant, which means that even when it is removed from one platform it may find expression elsewhere, possibly on the same platform under a different name or on different online spaces. If a website is shut down, it can quickly reopen using a web-hosting service with less stringent regulations or via reallocation to a country with laws imposing higher threshold for hate speech. The itinerant nature of

hate speech also means that poorly formulated thoughts that would not have found public expression and support in the past may now arrive on spaces where they can be visible to large audiences.

Regarding anonymity or pseudonymity (false names), the possibility of anonymous posting on online social media networks tend to make perpetrators of hate speech more comfortable to express their feelings, because their hidden identities dissipate their fears of having to deal with any consequences of their action. Anonymity especially, on social media may also be an obstacle to prosecution. Citron & Norton (2011) added that the internet itself facilitates anonymous and pseudonymous discourse, which can just as easily accelerate destructive behaviour as it can fuel public discourse.

Moreover, the transnational reach of the internet enhances the effect of hate speech and poses complications regarding legal mechanisms for combating online hate speech. In addition, King and Sutton (2013) have added that the climate of online hatred is characterized by targeted discrimination, prejudice and violent attacks, which tends to cluster in time and space and drastically increase after so called 'trigger events'.

The 2015 European Commission against Racism and Intolerance (ECRI) report highlights online hate speech as one of the main trends of the previous year, emphasizing that "hate speech through social media is rapidly increasing and has the potential to reach a much larger audience than extremist print media were able to reach previously"; Also UNESCO has recently focused on this growing issue, mapping and analyzing the existing initiatives to combat online hate speech in their comprehensive 2015 report "Countering online hate speech" (Olga & Roiha, 2016).

Alakali, Faga & Mbursa (2016) examined the phenomenon of hate speech and foul language on social media platforms in Nigeria, and found that promoting hate speech and foul language on social media have moral and legal consequences in the society and to journalism practice and that there is need for people to know and understand that hate speech and foul language attract legal consequences against perpetrators of hate speech and foul language in Nigeria. (Alakali, Faga & Mbursa 2016), even though the implementation of such legal implications had been quite obscure due to the nature of social media platform in Nigeria whereby users can manipulate words and castigate people and often are not prosecuted.

EMPIRICAL STUDIES

Existing studies suggest that attempts have been made at obtaining a bigger picture of the problem of hate speech. According to Olga & Roiha (2016) the study by the French organisation MRAP which in 2008-2009 analyzed over 2000 Uniform Resource Locators (URLs), from presumed hate sites, links to and from the sites leading to forums, blogs, social networking sites and videos, revealed a series of highly interconnected 'hate networks, illustrating the sophistication of many hate groups in spreading their ideology.

Another study, conducted in the UK, has focused on the perpetrators of Islamophobia on Twitter, following the Woolwich attack in May 2013 (Awan, 2013). The study examined 500 tweets from 100 Twitter users, looking at the hashtags #Woolwich, #Muslim and #Islam, to analyze how Muslims are viewed by perpetrators of online abuse, aiming to provide a typology of offender characteristics. The majority (72%) of the tweets analyzed was posted by men living in the UK, and over 75% of the tweets showed a strong islamophobic sentiment. The study is yet another piece of evidence of the growing issue of online Islamophobia, as well as of the role of so called trigger events in unleashing waves of hate crimes, including online hate speech against certain collectives.

A possible response to online hate speech, suggested by some scholars (e.g. Keats & Norton, 2011), is to give a stronger responsibility to intermediaries such as Google or Facebook, in fostering a digital citizenship, by transmitting norms of respectful engagement, as the silence of intermediaries "can send a powerful message that targeted group members are second-class citizens" (Keats & Norton, 2011). Keats and Norton specifically suggest for intermediaries to advance the fight against hate speech through increased transparency and engagement in conversations with stakeholders to identify the potential harms of hate speech and understand how a particular policy may regulate it.

In 2012 the Council of Europe in preparation of its campaign against online hate speech, conducted a mapping of existing initiatives addressing cyber hate. They focused on projects or organizations addressing the specific issue of online hate speech to conclude that relatively few organizations work specifically on this issue. In this regard, the recent study by UNESCO (2015) further provides an overview of responses to online hate speech. Among the responses described in the study, campaigns alerting companies advertising on social media of hate content serve as a tool making social media platforms react and withdraw hate content through reactions from advertisers. For example, in 2013 the group "Women, Action and the Media and the Everyday Sexism Project" in the UK launched a campaign showing page advertisements of prominent companies on Facebook pages that disseminated graphic sexist content. In response to the campaign, "Nissan and the insurance company Nationwide" withdrew their ads from Facebook. Having achieved this, the organizers together with online supporters began sending written complaints and photos of different adverts on hateful pages to other major companies on their social media platforms, urging them to follow suit. As a result of this campaign, 15 major companies decided to remove their adverts from Facebook. Shortly after, Facebook removed the content, and issued a statement expressing the need to clarify their content regulation policies and promote collaboration with organizations preventing online hate speech (UNESCO, 2015).

It is also presumed that Monitoring hate speech is key to understanding the phenomenon of online hate speech, and this monitoring tends to be conducted exclusively through collecting user complaints, which does not give a complete picture of the scale of the problem (Olga & Roiha, 2016). This few existing studies demonstrate that hate speech is quite prevalent, causes violence and a series of other related problems, intrapersonally, interpersonally and massively in the society which could also make business owners or online intermediaries loose customers or clients.

THEORETICAL FRAMEWORK
SPIRAL OF SILENCE THEORY
In terms of individuals' responses to hate speech on the social media, the spiral of silence theory (Noelle-Neuman, 1974) which is the tendency of people, so called bystanders, not to speak up about policy issues in public - or among their family, friends, and work colleagues when they believe that their own point of view is not widely shared can be applied. So, they tend to follow the view of the majority to form a negative opinion against the victims of hate speech. This is the case, for example, in the study conducted in the US by Hampton, Rainie, Lu, Dwver, Shin & Purcell (2014) looking into people's willingness to share opinions on sensitive social issues on social media. The finding showed that, in the Edward Snowden case, whilst 86% of Americans were willing to have an in-person conversation about the case, only 42% of Facebook and Twitter users were willing to post about it on those platforms because people were more likely to share their views if they thought people agreed with them and people choose not to speak out for fear of isolation. Previous studies (e.g. Rainie & Smith, 2012) have found that social media users are often not familiar with their friends' beliefs and are surprised to discover their friends' actual views via social media, and hence do not want to disclose their minority views for fear of disappointing their friends, getting into arguments, or losing their friendship. Some people may prefer not to share their views on social media because their posts persist and can be found later, in stages of life when they may compromise them. This suggests that while there is a free flow of interaction and expression of opinions on social media, people with dissenting views may choose to remain silent and may be influence into negative opinion formation against the victims of hate speech.

METHODOLOGY
The study adopted a textual analysis of Hate speeches assessable on the social media which were downloaded and analysed using the Critical discourse Analysis (CDA) which is an approach founded in linguistics, that allow users to carry out more systematic analysis of texts and language and can reveal how speakers and authors use language and grammatical features to create meaning and to persuade people to think about events in a particular way; it is also an approach that seeks to draw out the ideology of a text by pointing at the details of language and grammar, usually carried out either on selected examples of texts or speeches,

looking also at some parts and figures of speech as ingredients for analysis and also lexis, quotations and concepts. Some hate speech excerpts from prominent Nigerian personalities which were shared and discussed on the social media were studied using this approach. The study has a census population of all accessible hate speeches on the social media addressed to Nigerian political office holders and still accessible online. The unit of analysis included all hate speeches in the form of words, sentences, comments and long speeches made by Nigerian personalities or political office holders which are accessible on the social media from June to August, 2018

ANALYSIS
- **EXISTENCE OF HATE SPEECH ON THE SOCIAL MEDIA**

Table 1:
Existence of hate speech on Social media

Social media	Percentages (%)
Facebook	36% (N=19)
Whatsapp	19% (N=10)
Twitter	17% (N=9)
Instagram	15% (N=8)
Bado	9% (N=5)
Imo chat	0
Others	4% (N=2)
Total	100% (N=53)

Among the few hate speeches still available on the social media, facebook dominated as the primary medium for the propagation of hate speech (36%) within the period of the study, followed by Whatsapp (19%) and the others. The interactive and participatory nature of the social media platform enhances wider use and makes it possible for users to produce their content for public discourse and for the conventional media to share their links. These scenarios make social media a very popular means of information dissemination. This work also discursively reveals the preponderance of hate messages on different formats on the social media from various invisible online hate groups and also from politicians who attack one another openly in the mainstream media, which according to experts can be measured by knowing the hate words or hate targets priori which appear as "one word" person which captures when hate is directed towards a group of people. The hate could be likened to religion, race, personality, people, or vice versa. In the social media platform, Hate speech is often expressed by Nigerians without any keywords and it is difficult for human annotators to identify hate speech if they are not familiar with the meaning of words or any context that may surround the text (Zeerak Waseem and Dirk Hovy. 2016).

Table 2:
Sources and Targets of Hate Speeches

Sources	Percentages (%)	Targets	Percentages (%)
Supporters	42% (N=22)	A given political interest	9% (N=5)
Popular politicians	58% (N=31)	Nigerian political system/ opponents	91% (N=48)
Total	100% (N=53)	Total	100% (N=53)

This work further analyzed the sources and targets of hate speeches to discover that the two dominant sources are the party supporters (42%) and politicians (58%) with the target of achieving a particular political interest (9%) and more significantly, attacking their political opponents as well as the Nigerian political system (91%). This further explains that any user has the potential to broadcast any sort of message in these systems and reach millions of users in a short period which has democratized content publication as users can publish

content, and anyone interested in the content can obtain it (Leandro, Mainack, Denzil, Fabrˊıcio & Ingmar, 2016) and then share among friends. Despite the fact that some guiding journalism codes of ethics such as the Nigerian Media Code of Election Coverage exists the Nigerian society reject the use of such messages, while hate speech and foul language filled the media landscape.

- **HATE SPEECH BY PROMINENT NIGERIAN PERSONALITIES AS ALSO DISCUSSED IN THE SOCIAL MEDIA**

Table 3:

Nature of hate Speech

Nature	Percentages (%)
Attacking/ Indicting	53% (N=28)
Defensive	19% (N=10)
Controversial	4% (N=2)
Propaganda	20% (N=11)
Neutral	4% (N=2)
Total	100% (N=53)

The social media hate speeches as found in this study are predominantly attacking and indicting of the victims (53%) and propagandistic (20%) but with significant proportions on the defensive (19%) and hardly controversial (4%). One of such texts still available and often shared and being retweeted on the social media is this

"*The Igbos are collectively unlettered, uncouth, uncultured, unrestrained and crude in all their ways…Money and the acquisition of wealth is their sole objective and purpose in life. (Reported by Daily Post, August 8, 2013)*".

Another prominent example is this "*You should not be bordered with cockroaches of politics. Cockroaches are only found in the toilet even at homes, If you see cockroach in your house, Crush them.(Reported by Premium Times on 19th November,2014)*".

These further explain that while hate speech can be reported by the mainstream media, the social media plays the complementary role of audience discussion and analysis of such messages overtime, given also that hate speech is primarily aimed to subject to ridicule, the character and name of a given party or individual.

However, the emergence of new media has broadened the spread of information and further offers an ideal platform to adapt and spread hate speech and foul language easily because of its decentralized, anonymous and interactive structure (Alakali, Faga & Mbursa, 2016). Some hate speeches by prominent Nigerians reported by the mainstream media and as displayed in the social media include:

- *If they contest (Northerners) they are wasting their time. He who pays the piper will dictate the tune. We own them. We are feeding them. They are parasites. A beggar has no choice…They are beggars and parasites. (http://www.vanguardngr.com/2014/12/north-ungrateful).*

- *It is going to be rig and roast. We are prepared not to go to court but drive them out. (Tell, 7 July 2014.)*

- *The Igbos are collectively unlettered, uncouth, uncultured, unrestrained and crude in all their ways…Money and the acquisition of wealth is their sole objective and purpose in life. (Daily Post, August 8, 2013)*

- *You should not be bordered with cockroaches of politics. Cockroaches are only found in the toilet even at homes, If you see cockroach in your house, Crush them.(Reported by Premium Times on 19th November,2014).*

- *God willing, by 2015, something will happen. They either conduct a free and fair election or they go a very disgraceful way. If what happened in 2011 should again happen in 2015, by the grace of God, the dog and the baboon would all be soaked in blood. (Reported by Lika Binniyat in Vanguard Newspaper on May 15th, 2012).*

- *We assure those cold blooded murderers that this time, their blood thirsty campaign will not go un-replied. (Reported by Clifford Ndujihe in Vanguard Newspaper, 5th December, 2014)*

- *Anybody that come and tell you 'changes', stone that person… What you did not do in 1985, is it now that old age has caught up with you that you want to come and change…You cannot change rather you will turn back to a baby. (The Nation on Sunday, 15th March, 2015)*

These above speeches from prominent politicians in Nigeria were reinforced and severally discussed in the social media with expressions of desperation, anger and violence further revealing the predominance of hate words such as *"beggars, parasites, unlettered, uncouth, uncultured, unrestrained and crude, cockroaches of politics, murderers, the dog and the baboon etc"*. As illustrated above, this is often witnessed during electioneering campaigns whereby hate speech may be prone to manipulation; accusations of promoting hate speech may be traded among political opponents or used by those in power to curb dissent and criticism (Alakali, Faga & Mbursa, 2016).

Whereas some scholars (Kwok and Wang 2013) argue that most existing efforts to measure hate speech require knowing the hate words or hate targets priori, the general description of hate speech tend to be wide, sometimes even extending to embody words that are insulting of those in power or minority groups, or demeaning of individuals who are particularly visible in the society (Alakali, Faga & Mbursa, 2016) as witnessed in the hate speech excepts listed above. Meanwhile, Leandro, Mainack, Denzil, Fabr´ıcio & Ingmar (2016) provided a measurement study of the main targets of hate speech in online social media by gathering traces from two social media systems: Whisper and Twitter to develop and validate a methodology to identify hate speech. Their results identify online hate speech forms and a broader understanding of the phenomenon, providing directions for prevention and detection approaches. They designed the hate target token as simply a "one word" person which captures when hate is directed towards a group of people as given below:

Table 4:
TOP TEN TARGETS OF HATE IN TWITTER AND WHISPER AS GIVEN BY LEANDRO, MAINACK, DENZIL, FABR´ICIO AND INGMAR (2016)

Twitter		Whisper	
Hate Target	**% posts**	**Hate target**	**% posts**
Nigga	31.11	Black people	10.10
White people	9.76	Fake people	9.77
Fake people	5.07	Fat people	8.46
Black people	4.91	Stupid people	7.84
Stupid people	2.62	Gay people	7.06
Rude people	2.60	White people	5.62
Negative people	2.53	Racist people	3.35
Ignorant people	2.13	Ignorant people	3.10
Nigger	1.84	Rude people	2.45
Ungrateful people	1.80	Old people	2.18

As indicated in this table, Leandro, Mainack, Denzil, Fabr´ıcio & Ingmar (2016) isolated top ten targets of Hate in twitter and whisper. According to them, the transformative potential of social media systems brings

together many challenges among which is the ability to maintain a complex balance between freedom of expression and the defense of human dignity, as these systems open space for harmful discussions. This is because Social media platforms allow communication at near zero marginal cost to users and allow Internet users a congenial platform to freely express their thoughts and opinions (Leandro, Mainack, Denzil, Fabrʹıcio & Ingmar, 2016).

- **HOW HATE SPEECH COULD FUEL CONFLICT IN THE NIGERIAN SOCIETY**

Table 5:
Posts igniting Reactions/comments?

Variables	Percentages (%)
Audience Support	9% (N=5)
Audience Disapproval	74% (N=39)
Tensions/ Disaffection	15% (N=8)
Neutral	2% (N=1)
Total	100% (N=53)

Looking at the comments and reactions of users towards the displayed hate speeches on the social media, this work also revealed the audience disapproval of the hate speeches (74%), with comments suggesting that the messages could lead to violence and disaffection among people (15%). The audience comments also shows that there is a very marginal support for hate speech by the audience with majority suggesting that hate speech should not be encouraged.

However, the social media had often been regarded as the platform that aids interpersonal communication, interaction and exchange of various kinds of information such that people freely express themselves and make their views much publicly known. In the process, misunderstanding and controversies arise which often leads to defamation of character and damage of the image or ridicule of the character or personality of people. Two public statements in 2017 added to the social media tensions that made violence more likely to occur in Nigeria. The first statement regarded as the Kaduna Declaration, allegedly suggests that Igbo people should leave northern Nigeria to form its own nation. This is not withstanding the series of tribal and sectional conflict in the northern Nigeria. The statement had it that "It since ceased to be comfortable or safe to continue sharing the same country with the ungrateful, uncultured Igbos who have exhibited reckless disrespect for the other federating units and stained the integrity of the entire nation with their insatiable criminal obsessions" The second statement, allegedly condemned the first stating that, "their sponsors must be arrested and charged for treason". The social media discourse afterwards suggests that the writers were concerned with hate speech "from all corners of Nigeria", regardless of religion, geography or ethnicity arguing also that "If Nigeria is to split, it must be done without violence. This shows that Hate speech has continued to increase tensions in Nigeria suggesting therefore that the Nigerian government should protect her citizens and avert the spate of and tensions occasioned by hate speech

CONCLUSION AND RECOMMENDATIONS
This paper has been able to explore the nexus between the social media, hate speech and conflict arousal to ascertain whether and how the social media enhance or debunk the spread of hate speech, given that the fight against perceived online hate speech is beginning to reach a number of concerned parties, from governments to private companies and to a growing number of active organizations and affected individuals (Leandro, Mainack, Denzil, Fabrʹıcio & Ingmar, 2016). Based on the findings of their empirical investigations, Alakali, Faga & Mbursa (2016) concluded that hate speech and foul language is prevalent on social media platforms in Nigeria, with both moral and legal consequences in the society and the journalism profession. This work further recommends in line with their stand that:

…among others things, the Nigerian government and NGOs should sponsor monitoring projects like the UMATI in Kenya to better understand the use of hate speech and monitoring agencies to be set up under the legal regime to adopt mechanisms to identify and remove hate speech content on social media platforms in Nigeria;

The Nigerian government and NGOs should sponsor monitoring projects like the UMATI in Kenya to better understand the use of hate speech and foul language online by monitoring particular social media networking sites, blogs and online newspapers;

Media organizations and journalists who are morally inclined to ethical journalism should mobilize and conscientize the public through citizenship education to shun and confront hate speech and foul language online as part of their civic responsibility;

Let there be a purposeful enforcement of the crime of hate speech by law enforcement agencies in Nigeria, to implement the provisions of the Electoral Act and the Cybercrime (Prohibition, Prevention etc);

In order to effectively regulate the use of social media to propagate hate speech and foul language, both internet providers and the various social media management teams should develop a programme of moderation and censorship of content on their platforms to remove unwanted content relating to hate speech;

Various specialized government agencies for censorship created under the Cybercrime Act should begin to collaborate with internet providers and managers of social media platforms to censor content relating to hate speech considered inimical to national security in Nigeria (Alakali, Faga & Mbursa 2016).

This work further concludes that the apart from the user generated content on the social media, the social media shares and comments on the hate contents reported by the mainstream media even after several months the news was released. In view of these, we also concur with the submission of Adesina (2018), that government, and concerned bodies should integrate education and public sensitizations so as to curtail hate speech and ensure sustainability of the nation's democracy; owners of online media should uphold credibility and integrity and all process leading to news gathering, and dissemination must be made credible, accurate and truthful.

REFERENCES

Adesina, O. (2018). *Please Release: others speak on hate speech*. March 25, 2018 Press Release. Premium times. Available @ www.premiumtimesng.com/news. (Accessed on 07/06/2018)

Alakali, Terfa T.; Faga, Hemen Philip & Mbursa, Jinatu (2016). Audience Perception of Hate Speech and Foul Language in the Social Media in Nigeria: Implications for Morality and Law. *Academicus - International Scientific Journal,* pp. 161 to 178. Available online @ www.academicus.edu.al. (Accessed on 01/08/2018)

Ali, Anthony Danladi. (2013). Media Role and the Hurdle of a Nigerian Journalist 1984-1999. *Arabian Journal of Business and Management Review*, vol.2, No. 9, pp.1-13.

Awan, I. (2013) Victims of Anti-Muslims Hate. Available at http://tellmamauk.org/wp content/uploads/2013/09/appg.pdf (Accessed on 06/06/2018)

Ayitogo, Nasir (August 17, 2017) *Hate speech will no longer be tolerated in Nigeria* – Osinbajo Premium times. Available @ www.premiumtimesng.com/news. (Accessed on 07/06/2018)

Brinks, D.O. (2001). Freedom of Expression and Hate Speech. *Millian Principles, Legal Theory*. Vol. 24.

Citron, K. D. & Norton, H. (2011). Intermediaries and Hate Speech: Fostering Digital Citizenship for Our Information Age. *Boston University Law Review*, Vol. 91.

Cornwell Nancy C. & Orbe Mark P. (1999). Critical Perspectives on Hate Speech: Explicating the Centrality of 'Dialogic Listening' in Attempts for Meaningful Dialogue. *International Journal of Listening*. 13, pp.75-96.

Council of Europe (1997) Recommendation (97)20 of the Council of Europe, *Second Summit of Heads of State and Government of the Council of Europe in Strasbourg*

Dalhberg, L. (2001). The Internet and Democratic Discourse: Exploring the Prospects of Online Deliberative Forums Extending the Public Sphere. *Information, Communication & Society Journal*, vol.4:4.

Downs, D.M. & Cowan, G. (2012). Predicting the Importance of Freedom of Speech and the Perceived Harm of Hate Speech, *Journal of Applied Social Psychology*, Vol. 42, No.6.

Gagliardone, I., Gal, D., Aleves, T., and Martinez G. (2015) *Countering Online Hate Speech*. Available at http://unesdoc.unesco.org/images/0023/002332/233231e.pdf/ (Accessed on 28/06/18)

Hampton, K., Rainie, L; Lu,W.; Dwver M; Shin I & Purcell K. (2014) Social media and the 'spiral of silence'. *Pew Research Centre*. Available at http://www.pewinternet.Org /files /2014 /08/PI_Social-networks-and-debate_082614. Pdf

Jurgenson, N. (2011) *Digital Dualism and the Fallacy of Web Objectivity*. Available at: http://thesocietypages.org/cyborgology/2011/09/13/digital-dualism-and-the-fallacy-of-web-objectivity/

Kayambazinthu, E. & Moyo, F. (2002). Hate Speech in the New Malawi. in H. Englund (ed.). A *Democracy of Chameleons: Politics and Culture in the New Malawi*. Stockholm: Elanders Gotab.

Keats Citron, D. & Norton, H. (2011) "Intermediaries and hate speech: Fostering digital citizenship for our information age", in *Boston University Law Review, Vol. 91*: 1435.

King, R.D. & Sutton G. M. (2013) "High Times for Hate Crime: Explaining the Temporal Clustering of Hate Motivated Offending", *Criminology* 51(4):871-894

Kwok, I., & Wang, Y. (2013). *Locate the hate: Detecting tweets against blacks*. In Proc. of AAAI.

Leandro Silva; Mainack Mondal; Denzil Correa; Fabr´ıcio Benevenuto & Ingmar Weber (2016). Analyzing the Targets of Hate in Online Social Media. *Association for the Advancement of Artificial Intelligence* (www.aaai.org)

Leets, L. (2002). Experiencing Hate Speech: Perceptions and Responses to Anti-Semitism and Antigay Speech. *Journal of Social Issues.* Vol. 58, No. 2, pp.341-361.

Neisser, E. (1994). Hate Speech in the New South Africa: Constitutional Consideration for a Land Recovering from Decades of Rational Repression and Violence. *South African Journal of Human Rights* 10. Pp.333-356.

Nemes, I. (2002). Regulating Hate Speech in Cyberspace: Issues of Desirability and Efficacy. *Information & Communications Technology Law*. Vol. 11,No.3.

Noelle-Neumann, E. (1974). "The Spiral of Silence - A Theory of Public Opinion." *Journal of Communication* 24(2): 43-51

Olowojolu, Olakunle. (2016) Role of Media in 2015 Presidential Election in Nigeria. *International Journal of Politics and Good Governance.* Vol. VII, No.7, 2016, pp.1-12.

Olga, Jubany & Malin Roiha (2016) Backgrounds, Experiences and Responses to Online Hate Speech: A Comparative Cross-Country Analysis. Available @ http://www. unicri.it/special_topics/hate_crimes/Backgrounds_Experiences_and_Responses_to_Online Hate_Speech_A_Comparative_Cross-Country_Analysis.pdf.(Accessed on 25/06/2018)

Osibanjo (2017). *We'll treat hate speech as terrorism, Osinbajo warns.* Available @ https://guardian.ng/news/well-treat-hate-speech-as-terrorism-osinbajo-warns/. (Accessed on 25/06/18)

Parekh, B. (2006). Hate Speech: Is There a Case for Banning? *Public Policy Research*.

Post, R., Hare, I. & Weinstein, J. (2009). *Hate Speech. in Extreme Speech and Democracy*. Oxford University Press.

Rainie, L & Smith, A (2012) Social networking sites and politics. *Pew Research Center: Pew Internet and American Life Project.* Available at http://www.pewinternet.org/files/old-media//Files/Reports/2012/PIP_SNS_and_politics.pdf

Salem, Fadi & Mourtada, Racha (2011). *Civil Movements : the impact of facebook and Twitter*. Retrieved 17 April 2018.

Spiegel, L. (1999). Hate Speech, Civil Rights and the Internet: The Jurisdictional and Human Rights Nightmare. *Albany Law Journal of Science & Technology*.

Taylor, Jherez; Peignon, Melvyn & Chen, Yi-Shin (2018) *Surfacing contextual hate speech   words within social media*. Available @ https://arxiv.org/pdf/1711.10093.pdf. (Accessed on 25/06/2018)

UNESCO (2015) *Countering online hate speech*. Unesco series on internet freedom. Available   @ http://unesdoc.unesco.org/images/0023/002332/233231e.pdf. (Accessed on   25/06/2018)

United Nations Committee on the Elimination of Racial Discrimination (2013) "*General recommendation on combating racist hate speech*", CERD/C/GC/35.

Waldron, J. (2012) *The Harm in Hate Speech*. Harvard University Press.

Witschge,.C. (2008). Examining Online Public Discourse in Context: A Mixed Method   Approach. *Javahost- the Public*, Vol.15, No. 2.

Zeerak Waseem & Dirk Hovy (2016). Hateful Symbols or Hateful People? Predictive Features   for Hate Speech Detection on Twitter. *Proceedings of the NAACL Student Research   Workshop* (2016), 88–93.

APPENDIX A
CODING SHEET

Sources of the Hate Speeches
Supporters – 01
Popular politicians – 02

Targets
A given political Interest – 01
Nigerian Political System/opponents – 02

Social Media
Faceook – 01
Whatsapp– 02
 Twitter– 03
Instagram – 04
Bado- 05
Imo Chat – 06
Others - 07

Nature of Hate Speech
Attacking/ Indicting – 01
Defensive – 02
Controversial – 03
Propaganda- 04

Post Reactions
Audience Support – 01
Audience Disapproval – 02
Tensions/ Disaffection – 03

S/N	Social Media	Sources of Hate Speeches	Targets of Hate Speeches	Nature of Hate Speech	Posts Reactions