

Measuring Discourse Bias on Top Speeches by Women for Women: Text Mining Analysis

DAGIIMAA BALAAZ

School of language and culture,
University of the Humanities
Ulaanbaatar, Mongolia

Abstract:- The purpose of this research article is to show how the participation of women politicians and their public speeches in social cognition and how women speakers deliver their voice at the decision-making level, using the strategies based on discourse topics and text mining techniques. For this, knowledge is an important factor in creating discourse and understanding discourse, we examined.

Knowledge is the general level of "understanding" of a writer, speaker, or interlocutor, how he or she views the issue from many angles, how well it conforms to general principles, how the real world evaluates the interrelationships of phenomena, and the timing of the issue. All forms of knowledge should be taken into account, and the use of tactics to make speech understandable is also related to knowledge. Semantic knowledge is an experience gained over a lifetime, and the ability to understand and abstract anything is a mental activity based on practice. It is measured by episodic-based personal experiences, memories, and general knowledge. Thus, we intended to show how semantic knowledge is getting the base of discourse in social cognition.

Keywords:- knowledge, topic, context model, mental models, mental representations, bias, topic modeling, discourse network, evolution of topics, influence distribution.

I. INTRODUCTION

The process of discussing, expressing opinions, and persuading others is a mental act of transmitting knowledge and expressing knowledge, and it is a manifestation of language use, thinking through language, and communication. Therefore, modern approaches to tendency of ideology are studied in many ways based on the theory of discourse. The purpose of our research is to study the speeches of top women speakers on language use, discourse theory, general and specific methods of semantics with modern text mining topical modeling, and to evaluate the specific style, tendency to persuading, and topics of speeches and discussions.

Discourse is the study of language, thought, and cognition, and studies the minds, behaviors, attitudes, and beliefs of people who is discussing the idea based on language use, knowledge, and attitudes. Therefore, the study of discourse theory and methodology of discursive approaches are important factors for language studying.

Modern linguistics, in combination with artificial intelligence and statistical methods, has chosen this style as a significant innovation in the study of factor and phenomenon analysis using the "Language-Mathematics-Statistics" approach.

The purpose of this study is for analyzing of the speakers' speeches on similar topics, the specifics of language use, and the expression of the speaker's ideas and discourse BIAS.

The meaning of bias: [1]-[2]the word 'bias' ('biās) *noun: mental tendency or inclination, an irrational preference or prejudice 1(a): an inclination of temperament or outlook especially: a personal and sometimes unreasoned judgment: prejudice (b): an instance of such prejudice: bent, tendency 2(a): deviation of the expected value of a statistical estimate from the quantity it estimates.* Synonym study for bias: bias, prejudice mean a strong inclination of the mind or a preconceived opinion about something or someone. A BIAS may be favorable or unfavorable: bias in favor of or against an idea. PREJUDICE implies a preformed judgment even more unreasoning than BIAS, and usually implies an unfavorable opinion: prejudice against people of another religion. [1](www.dictionary.com/bias) (BRI)

'Bias is commonly understood as inclination or prejudice towards a certain point of view. A discourse or text that has a bias may have a certain agenda or promote a certain ideology. In the age of "fake news", the rise of extreme ideologies and various misinformation techniques it is important to be able to identify the level of bias in discourse: be it social network posts, newspaper articles or political speeches. Bias is not necessarily a bad thing. Sometimes it can make an intention stronger, push an agenda forward, make a point, persuade, dissuade and transform. When we measure bias we measure how ideologically charged a text is, how much it wants to put forward a certain point of view' said it[29](Paranyushkin)

According to the definition and the conclusions of some researchers, bias is a purposeful discourse that seeks to bring any issue facing society to the attention of the authorities and society in a positive and in its favor. It is clear that it is a social psychological issue that influences for making decisions based on particular ideology.

II. LANGUAGE AND IDEOLOGY

The relationship between language and ideology is very close, and scholars explain it from the point of view of linguistic theory by the fact that language has specific linguistic and lexical tools for expressing ideas and attitudes [27] (Joseph & Taylor, 1990). Linguistics, lexical specific meanings, and psychological factors are communication factors, and they are discourse in themselves. From this it can be seen that a purposeful conversation on a topic is considered discourse. Therefore, the concept of creating and understanding discourse is interrelated with the ideas and opinions expressed in the context. [19] (Van Dijk in Wodak & Chilton, *Critical Discourse Analysis: History, Agenda, Theory, and Methodology*, 2009, p. 71) Ideology is a method of persuasive communication that uses language as a means of communication to achieve its goals. Ideology is “a concept that cannot be separated from any circle of independent ideas pertaining to discursive discourse” but is a way of expressing ideas, discourse and experience.

In the process of this relationship, it is a matter of the individual's position and attitude, which attracts and persuades people, and thus builds its own relationship with society. This is due to the need to establish their own discursive approach with the help of language, to use the method of speech, to convey and store ideas with a purpose and organization. [3] (Alan , Kate, & L. Althusser , 1972)At the level of thinking, discourse analysis is a process of oral and written, mental, and intellectual expression, and in this sense, discourse is a linguistic study of linguistic methods by recognizing and evaluating whether the tone of an expression as a whole changes. It is important to study the interrelationships of sentences, meanings, speech, and dialogue, to explain the meaning behind the sentences, and to take into account the social status and psychological factors of the speaker. The word "politician" means people who know how to use political language to achieve their goals. Therefore, discourse explores ways to use political language as a powerful tool for self-confidence, self-support, and motivation when needed. Sociocognitive attitudes reflect cognitive perceptions, participants' knowledge, and beliefs, and serve as a “interface” for discourse and social cognition. Through a sophisticated social and cognitive interface, they create beliefs or social representations that people share with others in their group or community [12](Dijk T. A., *What is Political Discourse Analysis*, 01 January 1998) Thus, social cognition is defined as the collective representation of a society that includes knowledge, attitudes, ideologies, values, and standards.

Discursive expression is a means of communication used to achieve ideological goals and to establish authority and dominance. It is the discourse of discourse to openly define, explain and promote the ideology and political activities of social inequality through the use of language. [30] (Phillips, Jørgensenm, 2002) Every decision, news, and statement that is made to guide the way of life of a society is entirely a political activity, and it is easy for the public to express their political views, personalities, lifestyles, values, and eloquent language tools. is a phenomenon.

“The basic of discourse and the theory of knowledge is the principle of understanding something from one's own point of view. For modernist theorists, every tendency in language has a role to play. The concepts of language use, influence, and knowledge play a key role. Therefore, discourse is inextricably linked to an individual's point of view. Methods of expression include language vocabulary selection, grammatical features, and body language. The psychological state of the people with whom you are interacting and mutual understanding are key factors. It is important to briefly explain why this psychological factor is related to the concept of “language” and linguistic issues.

III. THEORY OF KNOWLEDGE: MENTAL REPRESENTATIONS

- **Knowledge:** Knowledge is an core factor in creating and understanding discourse which is the general “understanding” level of a writer, speaker, or interlocutor, how people discuss the issue from many views, how well it meets the general basic principles, how evaluates the relationship between the real world and the phenomenon, It is the timing of the issue between the times, and how chronology is used, are the most important criteria for determining the direction of an individual's attitude and opinion. [9] (Б.Дагиймаа) Both knowledge and discourse are fundamental notions in the humanities and social sciences. Since cognitive and social psychology are the only disciplines that have paid extensive attention to the role of knowledge in discourse processing, our general perspective on this relationship will generally be sociocognitive.[17] (Dijk T. A., *Discourse and knowledge*)
- **Topics:** The macrostructure theory of discourse is a theory of topics, and topics can be understood as keywords or key ideas that contain the main idea of a complex part. In discourse macrostructure theory, keywords or topics play an important role, both externally and internally, forming a hierarchical system of meanings. The meaning of discourse is not limited to the meaning of its words and sentences. ‘Topics typically are the information that is best recalled of a discourse.
- **Context model of language:** ‘Context models Language users not only form semantic situation models of the events referred to in a discourse, but also reflexively construe dynamic pragmatic models of (each moment of) the very communicative situation in which they participate themselves, and these are called context models [20] (Van Dijk, 2008a; see also Givón, 2005). These models are crucial for the management of discourse, because they represent the way language users interpret their current environment as relevant for the current discourse [14] (Clark, 1996) [17] (Dijk T. A., *discourse and knowledge*)
- **Mental models:** For social representations such as ideologies, knowledge and attitudes to have any specific impact at all on concrete discourses and social practices, a very important cognitive interface is still missing: mental models. Mental models are representations in episodic memory and may simply be identified with people s experiences. They are representations of the specific acts/events people participate in, witness or hear/read about [18] (Dijk T. A., *Discourse, Ideology and Context*)[17].[18].[19] (Clark, 1996; see also Smith, 1982).

To conclude it, mental representations deal with knowledge which is fundament of discourse. The **topics, context model of language, mental models** have specific impact at all on discourse comprehension and representations.

IV. TOPIC MODELING AND METHODOLOGY

In modern discourse, a text mining and NLP approach are widely used to optimize knowledge modeling and algorithms mathematically the text through the semantic node and neuron network to uncover the implications and ideas. Optimizing knowledge and information structure is evolving into a multidisciplinary, multidisciplinary scientific approach to digitizing large amounts of information, visualizing it, determining correlations, weighted averages, the distribution of topics, and sentimentality.

Mathematical approaches such as frequency and approximation value clustering are used to determine the structure of discourse values, and all text analysis methods, such as determining value nodes based on neural network algorithms, are performed by data-driven machine processing.

Latent dirichlet allocation (LDA) is an approach used in topic modeling based on probabilistic vectors of words, which indicate their relevance to the text corpus and how they can be used together to provide more relevant results for the general method. (Papadimitriou, Raghavan, Tamaki and Vempala, 1999), (David Blei, Andrew Ng, Michael I. Jordan, p. Dirichlet prior discourse.

Algorithms and formulas we used: Basic IR Models – Boolean Model – TF-IDF Weighting – Vector Model – PLSA – Latent Semantic Indexing Model – Neural Network Model – Retrieval Weigh a concept (keyword) in a semantic network by frequency: Tf-idf: (tf) normalizing and describing the power of inverse document [10].[11] (Dagiimaa.B, 2022) (Dagiimaa.B, Semantic macrostructure and mental models: mathematical aproach and text mining, 2022)

Objective: Within theory of discourse macrostructure aims to measure the discourse bias using a 'language-mathematical' model for optimizing knowledge'. The “topic modelling” of the discourse is scaled by mathematical models to determine the author's ideas, positions, implications, and conclusions as well as to determine whether decisions could be made. So we have built text networking on ‘THE MOST POWERFUL SPEECHES MADE BY WOMEN, FOR WOMEN’ SPEECH BY Oprah Winfrey, Michella Obama and Hillary Clinton, the speeches to the attitude to measure the bias

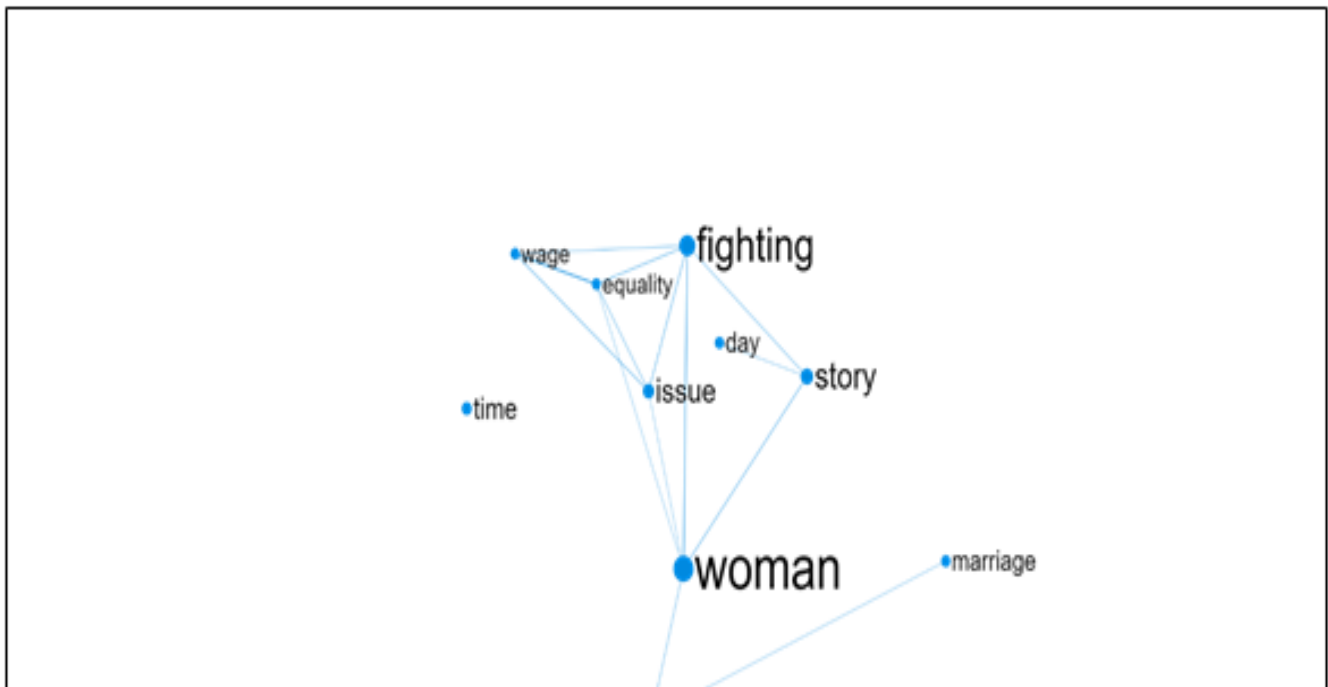


Fig.1: Main topics

Fig.1. *Speech 1st*: The most powerful speeches made by women, for women’ speech by Oprah Winfrey

- ✓ This discourse’s structure is Biased,
- ✓ Its immunity is very low, which means it is easier to spread information through it.
- ✓ It's leaning towards and promotes a certain perspective, circulating around the most influential nodes **‘WOMAN, FIGHTING, STORY, POWERFUL’**



Fig. 2: Keywords evolved over time

Fig.2 chart shows how the main topics (in Fig.1) and the most influential keywords evolved over time

- X-axis: time period (split into 10% blocks).
- Y-axis: cumulative frequency of occurrence.

their frequency. Meaning they may appear not as often as the most influential nodes but they are important narrative shifting points in 3 period.

latent brokers between the topics: The nodes that have an unusually high rate of influence (betweenness centrality) to

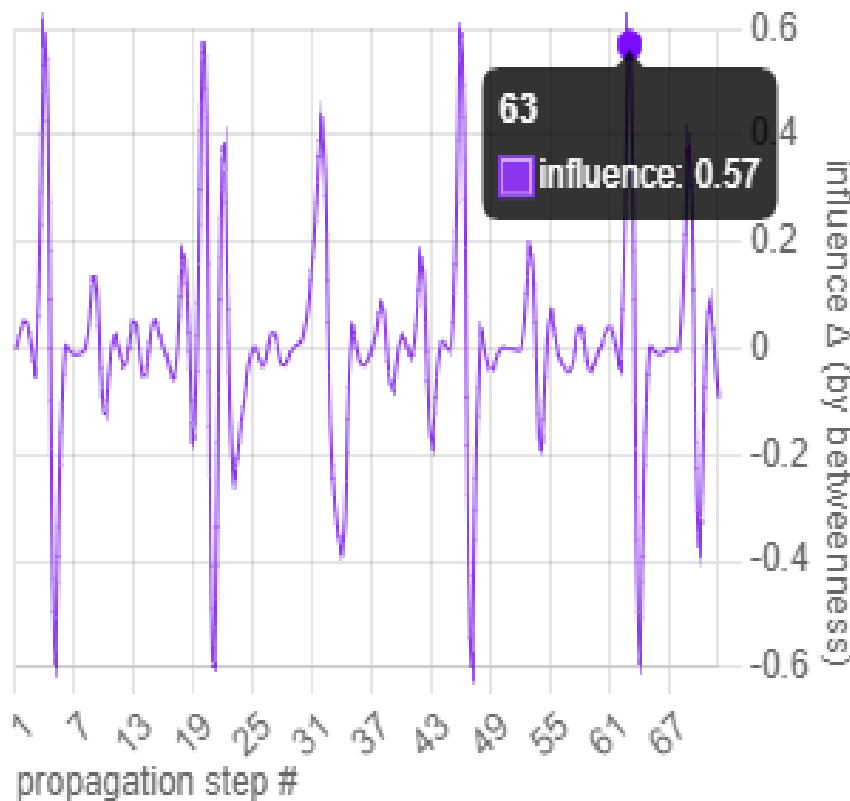


Fig. 3: Propagations: Narrative influences

Fig.3. The chart above shows how influence propagates through the network (The fluctuation shows how the narrative influences (in Fig.2) a overtime)

- X-axis: lemma to lemma step (narrative chronology).
- Y-axis: change of influence.

- Propagation dynamics: antagonistic variability | alpha exponent: 0.38 | low (based on Detrended Fluctuation Analysis of influence)
- Modularity: ks: 0.89, d: 0.40 <= cr: 0.61 | may be power law (^2)

Speech 2nd: Key note Address at young African women leader forum-June 22,2011By Michelle Obama:

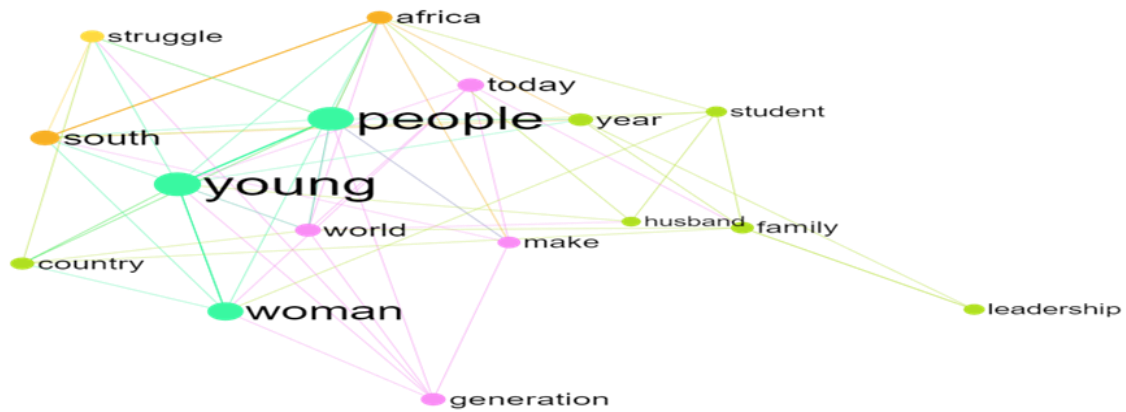


Fig. 4: Text network

Fig. 4. Shows ‘The network structure and mental representations of the speech: Speech by Michelle.O (Keynote Address at Young African Women Leaders Forum - June 22, 2011)

- Network Structure Insights: Focused (The current discourse structure in this network is which means that while there is some level of diversity, there is a focus on one topical cluster).
- Discourse structure in this network is **Biased**, which means very interconnected and is dominated by core nodes
- mind-viral immunity: Medium (**modularity**: >0.4 for medium), network will be more resilient and adaptive than a less diverse one.

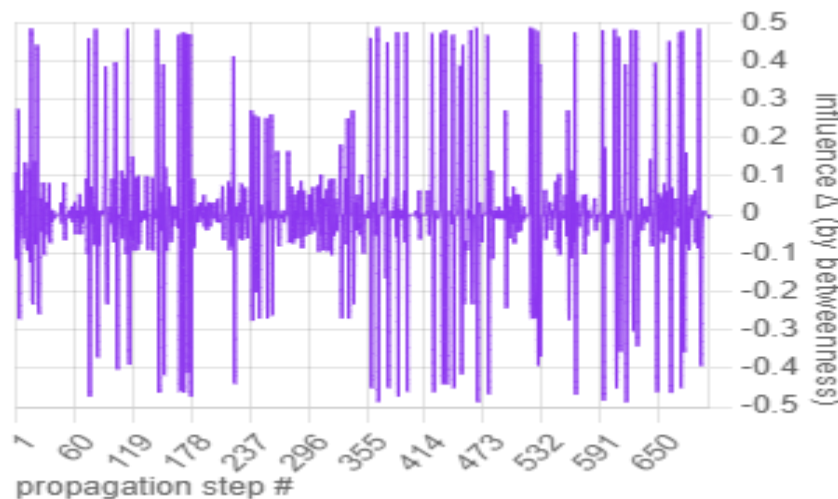


Fig.5. Influences of betweenness

Fig.5. The influence of narrative: speech by Michelle.O (Keynote Address at Young African Women Leaders Forum - June 22, 2011)

- Propagation dynamics: cyclical variability | alpha exponent: 0.57 | medium (based on Detrended Fluctuation Analysis of influence)
- plotting the log2 scales (x)
- log2 of accumulated fluctuations (y).
- ks: 1.58, d: 0.41 > cr: 0.36 | power law not identified (based on kolmogorov-smirnov test)

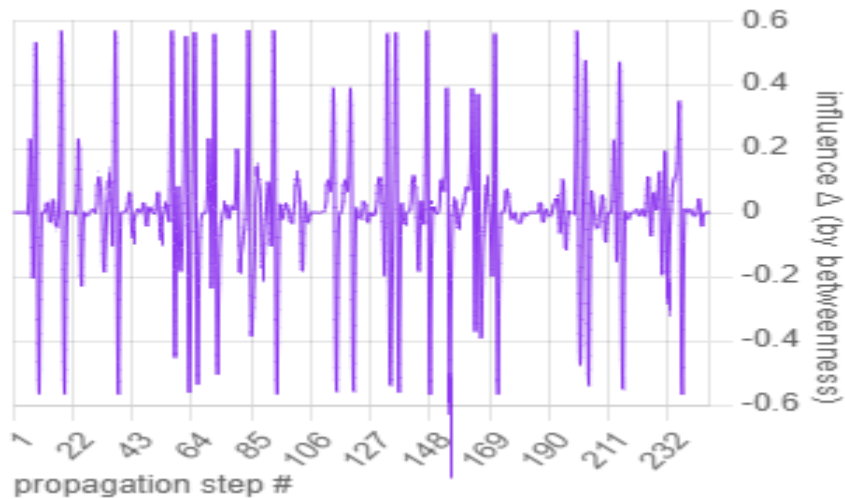


Fig. 8: The Influence of Speech

Fig.8. Shows the influence distribution of speech:

- Mind-viral immunity (Medium)
- Modularity: 0.51(Medium)
- Influence distribution: 40%
- The more even and rhythmical this propagation is, the stronger is the central idea or agenda (see alpha exponent below ~ 0.5 or less).

- The more variability can be seen in the propagation profile, the less is the reliance on the main concepts (agenda), the stronger is the role of secondary topical clusters in the narrative.

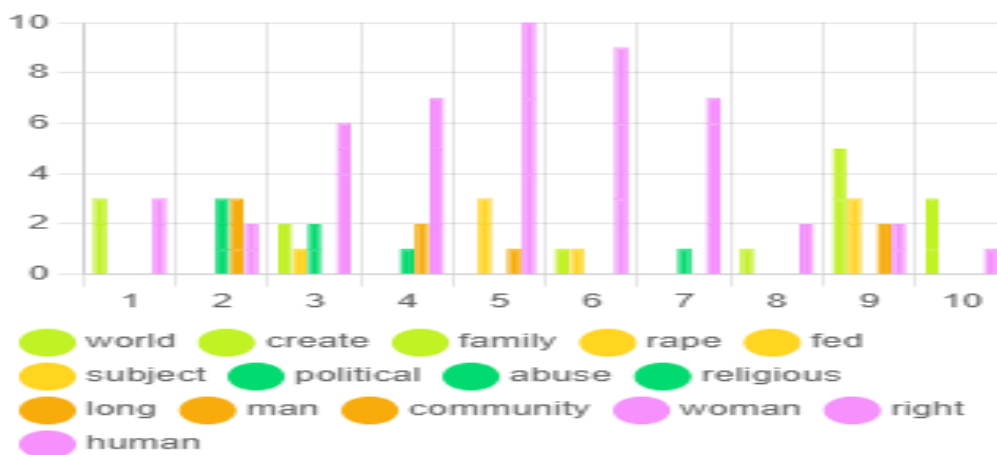


Fig. 9: Main topics

Fig.9. Shows the chart shows how the main topics and the most influential keywords evolved over time

- X-axis: time period (split into 10% blocks).
- Y-axis: cumulative frequency of occurrence.

V. SUMMARY

Politics itself creates ideological discourse in order to achieve its political goals. The purpose of political discourse is only to reach out to the public, also using the official vocabulary to express political sentiments is important. The mental representations deal with knowledge which is fundament of discourse. The **topics, context model of language, mental models** have specific impact at all on discourse comprehension and representations. **Topic modeling:** according to above charts and figures topics which are represented by semantic algorithms can represent the thought and ideology in discourse. The **topics, context model of language, mental models** have particular function of network, distribution, evolution of topic and delivery.

REFERENCES

[1.] (n.d.). BRITISH DICTIONARY DEFINITIONS FOR BIAS.
 [2.] (n.d.). www.dictionary.com/bias.
 [3.] Alan , S., Kate, N., & L. Althusser . (1972). *L. Althusser (1972), 'Ideology and Ideological State Apparatuses: Notes towards an Investigation', in Lenin and Philosophy and Other Essays, New York: Monthly Review.* Routledge. Retrieved from <https://www.taylorfrancis.com/chapters/mono/10.4324/9781315264547-25/althusser-1972-ideology-ideological-state-apparatuses-notes-towards-investigation-lenin-philosophy-essays-new-york-monthly-review-pp-85-126-alan-scott-kate-nash>

- [4.] Bassnett, & Schäffner. (2010). Political Discourse, Media and Translation. Newcastle upon Tyne: Cambridge Scholars Publishing. (Book Review). *Transcultural A Journal of Translation and Cultural Studies*, 5(1-2):219. doi:10.21992/T9791X
- [5.] Beard. (2000). A Comparative Study of Political Discourse Features in English and Arabic. *Researchgate*9(6):148). doi:10.5539/ijel.v9n6p148
- [6.] Bouvier, G., & David Machin. (2018). Critical Discourse Analysis and the challenges and opportunities of social media. *Research Gate*. doi:10.1080/15358593.2018.1479881
- [7.] Chilton, P. (2005, September 1). Analysing Political Discourse: Theory and Practice. Routledge, 2004. Paul Bayley (ed.): Cross-Cultural Perspectives on Parliamentary Discourse. (S. Harris, Ed.) *Volume 26, Issue 3, September 2005, Pages 462–467*,. Retrieved from <https://academic.oup.com/applij/article-abstract/26/3/462/181455>
- [8.] Dagiimaa.B. (2019). Analysinig political discourse and knowledge. *The journal of ES*.
- [9.] Dagiimaa.B. (2020). Semantic and knowledge. 500-95.
- [10.] Dagiimaa.B. (2022). Semantic macrostructure and mental models: mathematical aproach and text mining. *International Journal of Innovative science and reserch technology* .
- [11.] Dijk, T. A. (01 January 1998). What is Political Discourse Analysis. (J. B. Bulcaen, Ed.) *Political Linguistics*. Retrieved from <https://benjamins.com/catalog/bjl.11.03dij>
- [12.] Dijk, T. A. (2003). Ideology and Discourse. In *Ideology of dicourse* (pp. 118-46). University of the Barcelona.
- [13.] Dijk, T. A. (n.d.). discourse and knowledge. In *Discourse and knowledge* (p. 589). Cambridge University Press .
- [14.] Dijk, T. A. (n.d.). *Discourse and knowledge*.
- [15.] Dijk, T. A. (n.d.). Discourse, Ideology and Context.
- [16.] Dijk, T. A., & Woda, R. (2002). Critical discourse analysis (CDA).
- [17.] Dijk, V. (2006).
- [18.] Fairclough 2003: 165. (n.d.).
- [19.] Fairclough, & Şükran Gölbaş1. (1995). Critical Approach In Social Research.
- [20.] Fairclough, 1992:72. (n.d.).
- [21.] Joseph, & Taylor. (1990). *Ideologies of Language*. Routledge. Retrieved from <https://www.routledge.com/Ideologies-of-Language/Joseph-Taylor/p/book/9781138972278>
- [22.] Obama, M. (n.d.). Keynote Address at Young African Women Leaders Forum - June 22, 2011. <https://awpc.cattcenter.iastate.edu/>.
- [23.] Paranyushkin, D. (n.d.). Measuring Discourse Bias Using Text Network Analysis. <https://towardsdatascience.com/>.