The Development of Sentiment Analysis from a Linguistic Perspective

Samar Assem[1], Sameh Alansary[*2]

[*]*Phonetics and Linguistics Department, Faculty of Arts, Alexandria University*
Alexandria, Egypt
[1]samar.assem@alexu.edu.eg

[*]*Phonetics and Linguistics Department, Faculty of Arts, Alexandria University*
Alexandria, Egypt
[2]s.alansary@alexu.edu.eg

Abstract— ***Although sentiment analysis by definition is that field of Natural Language processing which focuses on analysing texts that tackle evaluating, analysing, and detecting the state of mind of the human beings towards a range of domains, most of the studies limit it to opinion mining. Yet, opinion mining is just one sub-field of three others under the umbrella of sentiment analysis which are opinion mining, emotion mining and ambiguity detection. Noticeably, ambiguity detection is considered to be a combination of the other two sub-fields thanks to its linguistic nature that considers statistical and/or syntactic-semantic levels of analysis are not adequate to reach a satisfying level of disambiguating human language. Henceforth, the current paper proposes digging deeply to reach pragmatic and socio-pragmatic levels of analysis in order to eliminate ambiguity and avoid misjudgements over texts and social media posts specifically in the sub-tasks of detecting and disambiguating hate speech, irony and sarcasm. Finally, it suggests utilizing an eclectic linguistic model of analysis includes speech act theory and the theory of (im)politeness***
Keywords*: Sentiment Analysis, Sarcasm, Irony, Hate Speech, Socio-Pragmatic, (im)politeness.*

1 INTRODUCTION

Decades ago, our human ancestors realized the power of words as a word can spark a war. The novelist Edward Bulwer-Lytton was right when he first quoted that “a pen is mightier than a sword”. It is the ambiguous power of words that may help a person to survive or kill him at once. A word may determine the fate of a nation or the fate of the whole world. It may destroy or build, it is the means by which human beings speak up their minds. Yet, reading what is going on in a human’s mind is not an easy task as the miraculous pearl that is called the brain has a lot inside. However, with the rapid development of technology, nothing is impossible and one of the golden keys to the human brain is … SENTIMENT ANALYSIS.

Recently, sentiment Analysis has become a popular field not only among researchers and technology and AI specialists but also among most of the other domains thanks to the huge number of data and the excessive use of internet and social media platforms. Henceforth, millions and millions of users all over the world with various interests express their feelings, emotions, opinions and ideas which makes the internet one of the main sources of information. Accordingly, there is a great need to analyzing this big data for the sake of decision making in most of the domains

Meriam-Webster dictionary has defined “sentiment” as an attitude, thought or judgement derived from a feeling. Accordingly, sentiment analysis includes all the areas concerning evaluating, analyzing and detecting the state of mind of the human beings towards various topics. It is the growing field which emerged nearly 20 years ago. Historically speaking, SA has been here since the emergence of the human communication as it is related not only to people’s opinions but it is also related to their feelings, emotions and affection. Thus, it is the gate to discovering what people think of.

Technologically, the whole story could have been started with the pioneering work of Philip Stone in 1966 when he proposed “General Inquirer” for detecting the usage of words of different categories. What worth mentioning is that “General Inquirer” is considered to be the basis for the unsupervised lexicon-based methods of sentiment analysis. However, Sentiment analysis as we know it today, that combines AI to technology, has officially emerged in 2002 when [1] conducted a study focusing on the classification of online documents into positive or negative sorting them by sentiment not by topic like it used to be.

Since then, Sentiment analysis has developed to include not only polarity detection or what is commonly called “opinion mining” but it also encompasses emotion mining and ambiguous language analysis. Recently, sentiment analysis is the field that interrelates NLP and Affective Computing which focuses on developing machines to be capable of understanding and responding to human emotions [2]. Accordingly, detecting human emotions could not only be from texts but from many other modalities including facial expressions, tone of voice and speech [3]. Henceforth, sentiment analysis as a field now encompasses three main sub-fields opinion mining and emotion mining and ambiguity detection which is considered to be a combination between both of the other fields as shown in fig 1 . Each of the three fields includes many other sub-tasks as follows:

2 SUB-FIELDS OF SENTIMENT ANALYSIS

A. *Opinion Mining*

Under the field of opinion mining there are sub-tasks starting from subjectivity detection to opinion summarization passing by opinion polarity classification and opinion spam detection

- **Subjectivity Detection:** tackles determining if a text is factual /objective like "the sun rises from the east" or opinionated/ subjective which means it expresses an opinion like "I love the sun rise". Practically, it is supposed to be the first stage in a sentiment analysis project. It works on detecting subjective or/and emotional ideas or phrases and sorting out any objective data.
- **Opinion Polarity Classification**: It is also known as "opinion analysis" and it is the task in which texts are classified as positive, negative or neutral. Interestingly, in cross language analysis tasks, the model is trained on a dataset from the source language and then tested on another dataset from a different language. Reference [4] argued that retrieval-based models give an alternative to Machine Learning based strategies for word polarity detection.
- **Opinion Spam Detection:** This subtask involves detecting fake or fabricated reviews/ opinions which are either criticizing or supporting a product, an idea etc… what worth mentioning is that the study of [5] is considered to be one of the first trials with trustworthy results. Spam detection focuses on identifying three main features which are the content of the review, the expertise of the field and the metadata of the review (star rating, IP address of the users, the location information of the user) in addition to bearing in mind real world general info and knowledge about the item/idea of interest.
- **Opinion Summarization:** This sub-task is really useful when it comes to decision making as it encompasses summarizing a large number of opinions / reviews about a specific issue, topic or product that includes various polarities, perspectives or aspects.

B. *Emotion Mining:*

Reference [6] has defined emotions as distinct and consistent responses to either internal or external events that have a specific significance for the person. Interestingly, opinion mining has been tightly connected to sentiment analysis however emotion mining is still a growing field specifically in the fields of AI and Natural Language processing. Moreover, most of the studies on emotion mining ignore the relationship between opinion and emotion mining although most of the techniques and methods utilized in opinion mining can be used in emotion mining as both are strongly connected as two sub-tasks of the main task of sentiment analysis in addition to being semantically connected concepts. Additionally, logically speaking, having either a positive or a negative opinion about an idea, a person, a product etc… could cause the person to have an emotion in the same direction too [3]. Accordingly, emotion detection involves sub-tasks as follows;

- **Emotion Detection:** this task is similar to subjectivity detection as it detects if the text includes any kind of emotions. This task works well in a lot of domains specifically business and customer service. [7] conducted a study with trustworthy results detecting emotions from the customer satisfaction surveys.
- **Emotion Polarity Classification:** this task is also similar to opinion polarity classification in which the existing emotions in the text are categorized**.**
- **Emotion Classification:** This task tackles fine grained classification of existing emotion in a text into one or more of a set of defined emotions
- **Emotion Cause Detection:** This task focuses on mining the factors behind specific emotions in the text.

C. *Ambiguity Detection*

This is a confusing category that holds many challenges in sentiment analysis field. Most of the issues under this category are hard and unsolved until now thanks to many factors. First, a humorous, sarcastic, ironic or a text of hate speech lie under both of the other sub-tasks of sentiment analysis; opinion and emotion mining as showed in "Fig. 1" below. Accordingly, it could be derived from an emotion that forms either positive or negative opinion towards an entity (person/idea/product/service) which is something hard to detect. Thus, this inference agrees with the previous definition of sentiment" as an attitude, thought or judgement derived from a feeling. Second, they lack clear definitions and mostly prone to misjudgment which provoke the resentment of most of the social media users and cause mislead decisions. The reason behind their ambiguity lies in their linguistic nature.

In communication, there are only two ways people use in order to send their messages to others; either to say what they intentionally mean. Thus, both of the illocution (what is meant) and the locution (what is literally said) are identical and this is called "literal language" or to say nonliteral language which means that both of the locution and the illocution are not identical. Consequently, there is a disconnection between the locution and the illocution and here comes the challenge of ambiguation. The first way of communication is an easy task for sentiment analysis as it could be easily analyzed by only the semantic and syntactic levels. Yet, without the interference of the pragmatic and socio-pragmatic levels, analysts cannot solve the mysteries of the nonliteral language.

Furthermore, [21] states that [9] has distinguished pragmatics from semantics to be announced as a separate independent linguistic branch; accordingly [10] defines pragmatics as the study of meaning in relation to speech situations. Moreover, [11] sees pragmatics as dynamic and changeable, since it is not about meaning; it is about making meaning. It shows how people understand and get the meaning in interaction. Thus, "pragmatics is mainly associated with describing the linguistic correlates of relatively changeable features of that same individual such as relative status, social role and the way in which the speaker exploits his/her sociolinguistic repertoire in order to achieve a particular goal" [11].

This particular goal is what he/she wants to deliver, what he/she intents to say. Additionally, Pragmatics does not only compasses what is said (locution) and what is meant (illocution) but it also includes the effect of what is said on the receiver of the message (perlocution) force [12]. This last part is key to the gate of solving the problems the sub-tasks of ambiguity detection specifically hate speech detection. Henceforth, the following section tackles the most challenging sub-tasks of ambiguation detection which are Humor, Irony/Sarcasm and hate speech from a linguistic perspective.

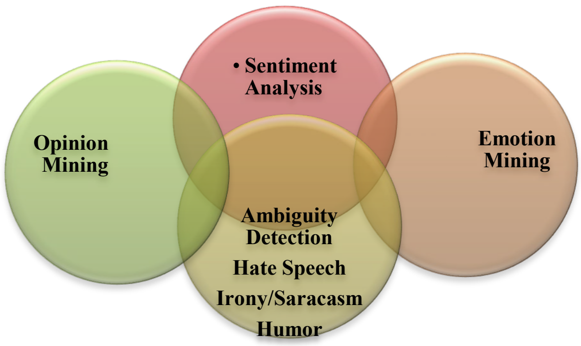

FIG 1 Subfields of Sentiment Analysis

3 CHALLENGES OF DISAMBIGUATING NONLITERAL LANGUAGE

As previously mentioned, literal message does not need a context to be inferred or understood from as what is said is what is meant. Accordingly, it equals the semantic level of analysis. If I say "give me a glass of water, please" it is a request for water. If a user writes on Facebook "I really love honest people" it indicates that she loves honest people; she simply informs the other users of that piece of information. Supposedly, no hidden messages to be inferred. Yet, what differentiate humans from robots is nonliteral language. People call it nonliteral language, figures of speech, figurative language and ambiguous language. Although how ambiguous it is, what makes our communication sounds human is the spice of language which is the nonliteral language.

A. *Hate Speech Detection*

In the age of big data thanks to social networks and social media platforms hate speech not only causes problems between users but it also influences various domains. Accordingly, it may cause a business to lose, drive a user to commit suicide or cause a war! Henceforth, researches exert great efforts attempting hate speech automatic detection. However, this is not an easy task because of many reasons. First, Social networks are full of texts of sarcastic, ironic and humorous contents which could be misjudged as hate speech. Examples from Egyptian tweets such as;

هقتلك ! مجتش ليه ع الغدا امبارح
يا جبان ! خضيتنا عليك
سافرت لوحدك يا خاين هندبحك
كل سنة و أنتم طيبين يا مسيحين

These tweets among thousands seem of hate speech content although they show how close the users are and they indicate camaraderie and intimacy.

- *Related Works*

Reference [8] used hybrid methods and incorporate twitter metadata with other sources of data. They used the word frequency vectors and the sentiment of the users towards three classes of hate speech sexism, neutral or racism according to user tweet history. Accordingly, their model obtained an f1 score of 0.93. Furthermore, [9] utilized a combination of boosted decision trees, LSTM and Random embedding obtaining an f1 score of 0.93. Additionally, [10] conducted a study detecting sentences of hate speech that gave results of 63.75% accuracy in the binary classification task. used lexicon-based methods to detect sexism in tweets and obtained accuracy of 0.704.

On the other hand, studies of Arabic Datasets showed a little but promising contribution. Some studies such as [12] have focused on hate speech against women. It built a dataset of 6550 tweets of the Levantine dialect. The tweets were labeled as either misogynistic or not and the misogynistic were labeled in 7 categories. it used ensemble technique that utilized SVM, Logistic Regression Classifiers Naïve Bayes. Additionally, [13] conducted a study using Arabic dataset in order to detect cyberbullying using sentiment analysis techniques and machine learning methods.

Accordingly, most of the previous studies tackling hate speech detection focus on either building a semantic dictionary of common hate speech words or on the binarily classifying it as. Yet, detecting hate speech is a more complicated process that needs deep analysis beyond statistical, semantic or semantic-syntactic analysis of texts specifically when hate speech is hidden behind humor, sarcasm and irony.

- *Challenges of Hate Speech Detection*

One of the unsolved issues in hate speech detection as previously mentioned is that it lacks specific definition. Yet, [14] mentioned that hate speech lies under the pragmatics field of linguistics. Additionally, it belongs to forensic linguistics because of its legal nature as the most of the linguistic approaches to hate speech mostly critique or quote legal texts. Moreover, every entity/social media platform has its own hate speech definition under its terms and conditions However, [15] has based their definition of hate speech on the various definitions of these entities (twitter, Facebook, YouTube, European Union Commission).

Accordingly, they defined Hate speech as “the language that attacks or diminishes, that incites violence or hate against groups, based on specific characteristics such as physical appearance, religion, descent, national or ethnic origin, sexual orientation, gender identity or other, and it can occur with different linguistic styles, even in subtle forms or when humor is used.” Thus, there is a consensus that hate speech definition has three main conditions 1) incitement and violence, 2) against specific groups, 3) in any linguistic level or form.

Yet, In the third condition lies the second challenge in hate speech detection as it is a very loose idea to determine what specific criteria make a text classified as hate speech? And how can and automatic hate speech detection system be built? In order to answer these questions, we suggest looking to hate speech detection from a socio – pragmatic perspective. Which means detecting and classifying hate speech in relation to speech act theory and Politeness theory with its three main perspectives.

4 POLITENESS AND THE SOCIO_PRAGMATIC APPROACH

Politeness as a linguistic phenomenon is being studied according to the relationship between language use and social behavior. [14] mentioned that studying politeness has two main approaches: the pragma-linguistic (pragmatic) approach and the socio-pragmatic approach. Scholars of pragma-linguistic approach have studied politeness in terms of language use to express the polite behavior. Thus, this approach is considered to be the modern approach of politeness. While scholars of the socio-pragmatic approach, who belong to the postmodern school of politeness, have studied it in terms of the relationship between the language use, its meaning to the speaker and the hearer, in addition to the effect of the big-picture view of how politeness relates to the social behavior and society in general.

Recently, some postmodern theorists add some changes to the socio-pragmatic approach; nevertheless, they have taken a novel approach to politeness which is “the discursive approach”. Additionally, some of them do not even mean to have a discursive perspective towards politeness although they share the same features of that approach [16].

As a matter of fact, the three politeness approaches are interchangeably tied since most of the theories of the socio-pragmatic approach have emerged as a criticizing reaction to the pragmatic theories. Furthermore, the discursive approach is a modifying phase for the socio-pragmatic approach. “Fig.2” below clarifies their relationship. Accordingly, putting an eclectic socio-pragmatic approach into accounting while designing an automatic model to detect hate speech will make a big difference. Accordingly, the proposed model to automatically detect hate speech is an eclectic one that based on a combination of pragmatic & socio-pragmatic theories of politeness.

A. Searle’s Speech Act Theory

Searle’s speech Act theory is considered to be the base of the pragmatic approach to politeness and pragmatics in general. Searle has built his theory upon Austin’s (1962) and he maintains three components of speech act. Firstly, the locution which is said by a speaker. It is simply the spoken words said by the speaker. Secondly, the Illocution refers to what is meant by the speaker or his intention. Thirdly, the perlocution refers to what the hearer does as a response to the speaker utterance or it is the effect of the utterance on the speaker. Thus, a speech act cannot be analyzed apart from the context, a concept on which Searle agrees with [16] and the discursive approach.

Searle also divided the illocution into five main types: assertives: how things are transferred in a form of information, directives: the speaker tries to get the hearer to do something, commissives: the speaker commits himself to do something and does so by the utterance, expressives: utterances in which the speaker expresses feeling and attitudes, and declaratives which bring about a change to the situation in which they are uttered.

This theory will be the starting point of hate speech detection as hate speech by definition is based on the speech act of incitement. Thus, in order to determine if the speech act is incitement or not, the model have to focus on; first the context of the whole text as the meaning is not inherited in speech acts, two collecting deeper metadata that could determine the community of practice for the users of such a text. The idea of community of practice will be discussed in detail in the section of Watts’ theory.

B. Brown and Levinson Politeness theory

In order to connect between the locution (actual words), illocution (what is meant) and perlocution (its effect on the receiver) forces of every text and solve disambiguation, they are analyzed as explicit (direct) or implicit (indirect) to clarify which strategy the user has used. Fig.”3” shows B&L’s original theory and shows B&L’s theory has been adapted to match the current study. Hence, the speech act will be classified either direct or indirect while the direct will be also binary classified bald without a pragmatic modifier or with a pragmatic modifier.

C. Aijmer’ Framework

As the nature of the social media data is a kind of turns interaction either in a form of post and comments to reply or just a comment to a comment. Thus, every turn has a function that affects the whole interaction and there is no text aimlessly written. The functions are analyzed according to Aijmer’s (1996, 2002) framework. Aijmer (1996) believes that every speech act has a specific frame according to the situation and the relationship between the users (community of practice). This frame includes the head act of the utterance, internal pragmatic modifiers or external pragmatic modifiers which function either to mitigate the force of the head act or enhance it. Pragmatic modifiers are illocutionary force indicating devices (IFIDs) modifying the illocutionary force signaled by the utterance mood (Aijmer,2002, P.25). She associates between the social situation and the speech act formula, a concept which agrees with [15]’s politic and (im)polite behavior.

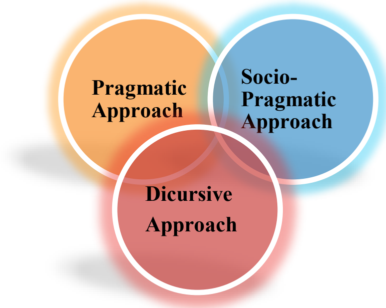

FIG2 Politeness Approaches

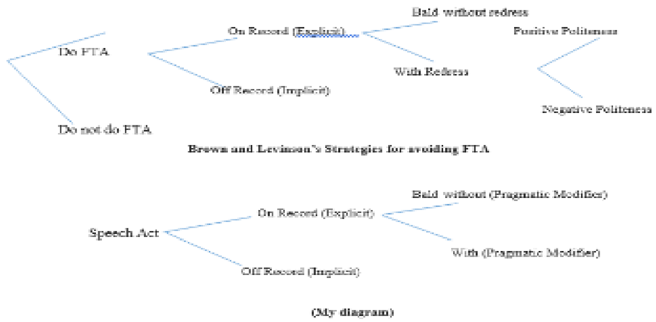

FIG3 Brown and Levinson's theory

D. Leech's Cost/Benefit Scale

The turn then is analyzed according to how much it is beneficial to the receiver or costly an idea that is derived from [16]'s cost/benefit scale of politeness with a discursive perspective. Leech suggests that politeness is based on three scales. He calls them the indirectness scale, the optionality Scales and the Cost/Benefit scale. Politeness principle arises from the cost/benefit principle as the producer should do his best to maximize benefit to receiver and minimizes cost to receiver [17]. Thus, the higher the cost to the receiver is, the less polite the act is. The three scales are connected as the higher the cost the more indirect the post/comment will be, consequently the more options the producer will offer the receiver. Leech's cost-benefit scale will be included in the analysis but with a discursive perspective; according to the context and the community of practice.

E. Watts' (im)politeness

The post then comes to pre-end stage in which it belongs to one of two categories which are whether it is accepted by the community of practice to be politic or either more or less than what is accepted by the community of practice to be (im)polite. Interestingly, [18] is one of the first theorists who have developed the notion of the community of practice. The community of practice notion refers to the language practices that are particularly developed in specific groups of interlocutors that are connected together on a specific task or activity. Accordingly, social media posts could not be analyzed without the metadata that shows the relationship between the users, the whole context of the post and the community they belong to.

The community of practice notion has a deep impact on the studies of politeness as it allows a contextualized analysis. Thus, Theorists pay attention to the construction and the evaluation of different communities to different norms as appropriate or not. Henceforth, an analyst cannot describe politeness without paying attention to a specific community of practice. [15] claims that there are not specific norms that determine the politic behavior as the social interactions are always interactive. [18] argues that the interlocutors in particular communities of practice are the only persons who can judge or evaluate an utterance as polite or not and these communities of practice are not fixed but rather always changeable.

Politeness is equivalent to giving more than what is required by the expected politic behavior and it may be evaluated positively or negatively, (im)politeness. Accordingly, any text within the acceptable range of the terms and conditions, the social context and the community of practice is considered politic. However, anything more or less than that socially imposed behavior it could be classified (im)polite according to the inference of the other participants (users).

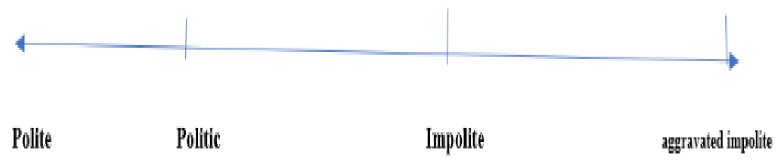

FIG 4 (im)politeness Scale

Here comes the last stage of analysis in case a text is negatively beyond the politic behavior. In this case it is impolite thus classified to whichever level of impoliteness. Accordingly, hate speech texts could be classified as aggravated impolite which is the extreme of the impolite text as shown below in Figure Fig."4". What worth mentioning also is that the two main notions that effect the politic or (im)polite behavior are the "habitus" and the "community of practice". The notion of habitus refers to the ways of behaving that we perform unconsciously in the society as they considered to be normal. They are such some unwritten rules that are developed inside a particular community of practice by the repetition of doing some actions in a specific way. For [16], the politic behavior is compatible to the nature of habitus according to the social characteristics of the situational context and it is always open to renegotiation in the social interaction. Which is in the context of social media could be to what extent the users are close.

Moreover, the points at which speakers see politic behavior to be "polite" may differ considerably from speaker to speaker, from community of practice to another and even from one situational context to another. The difference happens according to a wide range of social and contextual variables. The interlocutors may see examples of interactions to be appropriate to the norms established in previous interactions, that in any interaction will be unmarked will go unnoticed and it will be politic [19]. Figure."5" clarifies the anti-clockwise analysis framework.

Table 1 below shows the suggested analysis framework which is consisted of each of the stages a post/comment should pass by in order to be annotated as hate speech or not. It starts with classifying the head speech act according to Searle's speech act theory then annotated as direct or indirect if direct it goes to the next level if indirect in order to analyze if it has any pragmatic modifier and what exactly its function is. Next comes the stage of annotating the data into cost or beneficial to the hearer and finally comes the stage of sorting them into politic or (im)polite to be sorted in either polite or (im)polite in which lies at the end of the scale the hate speech post.

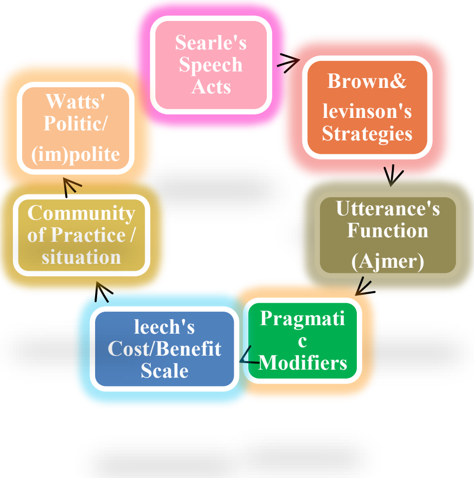

FIG5 Anti-clockwise analysis Framework

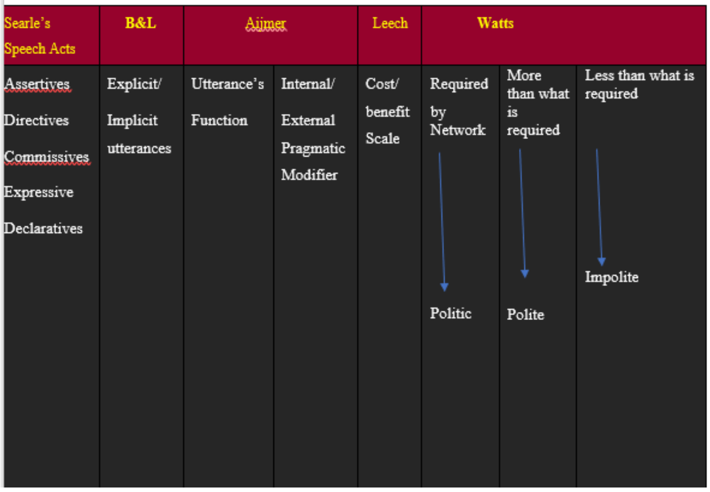

FIG 6 Analysis Framework

B. Irony/Sarcasm

Many years ago [20] suggests two levels of meaning which are included in any kind of verbal interaction: the semantic or the propositional meaning of the utterance (literal meaning) and, the intention of the speaker to make the utterance that is the illocutionary force of a speech act (nonliteral meaning). Accordingly, [20] cooperative principles (CP) theory is considered to be one of the most important contributions to the study of pragmatics. In the Gricean approach, both of the speaker and the hearer have forceful roles that are crucial in inferring the meaning. Moreover, the speaker says what he has to say in a proper attitude at the suitable time. According to [20] the CP contains four maxims and sub-maxims, as shown below in table (2), all of which describe the efficient cooperative use of language. Accordingly, any violation of the maxims signals conversational implicatures, which are non-explicit messages, intended by the speaker to be inferred by the hearer. Thus, for sentiment analysis, texts should be first classified as literal or non-literal using the four maxims. If a text is classified as non-literal as it has something that violates one of the maxims it should be passed to the second level of classification concerning irony detection as it will be shown in the next section

TABLE I COOPERATIVE PRINCIPLES MAXIMS

1- The Maxim of Quantity	2. The Maxim of Quality	3. The Maxim of relation	4. The Maxim of manner
Means making the speaker's talk informative as is required not more or less. (Be Brief)	Indicates giving true and adequate information. (Be True)	Tackles making the contributions relevant to the topic. (Be Relevant)	Refers to being brief and avoiding ambiguity (Be Clear)

Most of the confusion comes from overlapping between irony and sarcasm as most of the analysts deal with them as synonyms and adding to them satire and sometimes hyperbole. Accordingly, one of the key solutions to solve the disambiguation of this sub-task is to differentiate between all of these terms from a linguistic point of view bearing in mind that they all tightly connected to both of the other sub-fields of opinion and emotion mining.

In order to differentiate between irony and sarcasm, it is worth mentioning that sarcasm is a kind of irony among others. Accordingly, the classification of [21] of the kinds of irony is adapted in order to contribute to detecting irony.

- Canonical Irony: this could be the classification that simply defines irony as we know it. Canonical irony refers to using positive words to express negative evaluation. a text like " الفيلم حلو لدرجة اني مش محتاج منوم"indicates that the user did not like the movie although he used positive words like "حلو". On the other hand, noncanonical irony could be possible if we reversed this tweet by saying "الفيلم وحش لدرجة انه سهرني". Yet, this form is really rare in the normal human interaction because it is really rare to hide the positive feeling behind non literal language
- Cosmic/situational irony that refers to cases in which something out of the hands of the individuals happens as the actions of the individuals have nothing to do with them. These cases could be good or bad but mostly it is bad which makes a good environment for irony. Usually, this kind of irony accompanied by juxtaposition and contradiction as for example somebody tweets that "أحلي فرح دة و لا ايه المطرة نورت الليلة." this tweet seems positive

at the first glance or if analyzed semantically but when putting situational irony into consideration it shows how negative it is as the writer is very sad because of the contradiction. This kind of irony could be detected using metadata (information about the writer, the context of the text, previous texts, user's profile) in addition to giving general info to the machine about common sense that takes nothing from an ordinary human to understand such as the previous tweet.

- Verbal Irony: from a socio-pragmatic perspective, all the human being all over the entire world share some sort common ground. if common ground can be defined it could indicate that shared knowledge, beliefs and attitudes that can be utilized in the interaction between people. Accordingly, common ground could be regarding from sharing the place of residence (earth – region etc..) passing by sharing native language, political affiliation, career, occupation until we reach sharing the same family.

  Accordingly, [22] suggests that verbal irony has a tight connection with the degree of relationship between users from strangers to intimates. [22] called this degree inferability. Hence, if inferability is high, people tend to use more sarcasm/verbal irony however if inferability is low, the probability of using sarcasm/verbal irony is also low as in this case using sarcasm could be risky. Consequently, using metadata as previously mentioned serves as the context of texts. An idea which agrees with the socio-pragmatic analysis of the text specifically the idea of the perlocution force. Accordingly, it contributes in detecting if the text is positive or negative as this is also subject to if the relationship between the users allows verbal irony/sarcasm or not.

- Sarcasm is a kind of verbal irony. yet, not all sarcastic texts are ironic. thus, sarcasm is that kind of verbal irony in which the sender of the message intends to cause pain to the receiver [21] accordingly, this definition makes a close relationship between mockery and sarcasm.

Henceforth, in order to differentiate sarcasm from irony, we can say that sarcasm tend to express negative feelings towards the receiver and it could hurt his/her feeling and threaten his/her face. Interestingly, the pragmatic approach to politeness is called by some scholars "the Grice-Goffman pattern" because Goffman's notion of face plays an essential role in this approach specifically in theorizing politeness, despite not having a pragmatic framework. The notion of face is based primarily on [24]'s theory of social interaction. [24] defines face as that positive social value a man forms for himself and it is dependent on what he sees as the judgments of the other participants; his own judgments of their interaction with the discourse and frame.

Furthermore, it is the self-image which is described in accordance with approved social features and it could be shared by others, as when a person represents his job in a good way, he represents his self-image and the people who share the same job. Accordingly, [24] connects the notion of face with the English folk phrase "to lose face" which represents getting embarrassed because of some actions. For Goffman, a person's face is something that could be saved or lost in the interaction.

To sum up Goffman's approach, an individual's face depends completely on the social interaction. Thus, for this individual to have face he must be involved in an interaction. He could save or even insult the faces of the other interlocutors. Hence, every situation has a different scenario and the face of an individual is different in every situation. There will be different individuals making different judgments of the speaker based upon how he acts and other interlocutors too. Face is a temporary and momentary construction developed from an unlimited number of judgments and assumptions. Additionally,[16] argues that face, in [24]'s definition, is not a fixed property that people own; instead, it is negotiated and renegotiated during interactions, therefore it is a public property. In this way, Goffman's approach is compatible with the conceptualization of the discursive approach. In contrast; Brown and Levinson characterize face as an image that essentially belongs to the individual.

The reference [25] define politeness as couple of activities to find, maintain and save face during the interaction. For them, politeness is something which is emotionally invested, and that it could be lost, maintained or intensified and it must be always found in the conversation. Accordingly, they put some strategies through which participants can deal and cooperate with each other. Their theory of politeness proposes that people have an instinctive desire to observe the faces of both the person and the others in social interactions and they hypothesize that all members of a society wish to keep, or maintain face [25]. Using Goffman's concept of face work as a theoretical departure point and, mixing it up with Grice's theory of conversational maxims and implicature, [25] have created a theory of politeness depends on the presumption of the speakers' rationality and efficiency in the interaction and on the hearer's recognition of the communicative intention of an utterance. Accordingly, there are certain factors that control politeness in interaction: power distance, social distance and the degree of imposition an act entails. The strategies of interaction that the participants use rely on the nature and the context of the social relationship between the speaker and hearer. Speakers will be more polite when the power of the hearer over them increases, the social distance between hearer and speaker increases, and the amount of imposition on the hearer increases as well [26].

Every individual has a negative face and a positive one in the social interaction. The negative face is the desire to be completely free from imposition, while the positive face is the desire to maintain a positive self-image that is appreciated by others. Every speech act will be divided into direct or indirect, in their theory, pay their attention to Face threatening Acts (FTAs) and they design strategies to redress or soften the FTAs. Accordingly, we can classify ironic texts according to the scale in **Fig.2** below from the least FTA which is banter that happens most probably between family and friends to the most FTA which is mockery that is a version of sarcasm that is meant to hurt the feeling of others.

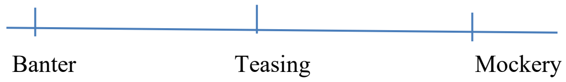

FIG 7 Scale of Irony

Thus, we can merge between fig 7 and fig 4 with some modifications to connects between (im)politeness from on hand and hate speech, irony and sarcasm form the other. As shown in fig 7, the scale could start with the polite text that save the face of the receiver and could be beneficial to him/her as it does not cause any hurt or harm. In the middle lies the politic texts which are normal and suitable to the community of practice and agreed by member of the common ground. Finally, comes the negative impolite texts under which lies the extreme negative point which is aggravated impolite where hate speech lies. However, this scale is changeable according to the degree of inferability between the users which means that either banter or teasing could be polite if there are more than what is accepted by the community of practice and give positive feeling to the receiver. On the other hand, they could be just politic if this is what is exactly required in the community of practice. Yet, when it comes to sarcasm an mockery both of them are included between impolite and aggravated impolite as by definition they both cause harm to the receiver and are meant intentionally to hurt his/her feelings.

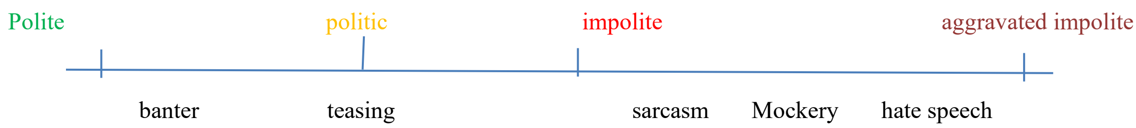

FIG 8 Levels of (im)politeness

5 Conclusion

This paper has proposed looking deeply into Sentiment Analysis. First suggested classified sentiment analysis into three sub-categories which are to some extent interrelated. Additionally, it looks from a deeper linguistic level in order to solve the ambiguation of the social network language and to fully understand the human communication. Accordingly, it inviting the researches of the sentiment analysis field to go beyond the semantic and syntactic levels of analysis and to have a deeper look into to the higher levels of analysis which are the pragmatic and socio-pragmatics levels of linguistic analysis of the social networks language in order to solve the disambiguation of the nonliteral language in sentiment analysis. Additionally, it offers a socio-pragmatic model of analysis for automatic hate speech detection using Speech Acts theory and (im)politeness in order to avoid misjudging the posts of the social networks. Finally, this paper recommends more future works about humor detection and dealing with sarcasm and irony more precisely. Last but not least, it rings a bell to conducting more papers in humor, irony, sarcasm and hate speech detection in the Arabic language from a socio-pragmatic point of view side by side with machine learning and deep learning methods.

REFERENCES

[1] L. Lee, and S. Vaithyanathan, "Thumbs up?" Proceedings of the ACL-02 conference on Empirical methods in natural language processing - EMNLP '02, Stroudsburg, 2002.

[2] H. Aka Uymaz and S. Kumova Metin, "Vector based sentiment and emotion analysis from text: A survey", Engineering Applications of Artificial Intelligence, vol. 113 no. 1, 2022, pp. 2787–2793.

[3] A. Yadollahi, G. Shahraki, and O. Zaiane, "Current state of text sentiment analysis from opinion to emotion mining", ACM Computing Surveys, vol. 50, no. 2, 2018, pp. 1–33
[4] M. Singh, A. Jakhar, and S. Pandey, "Sentiment analysis on the impact of coronavirus in social life using the bert model", Social Network Analysis and Mining, vol. 11, no. 1, pp. 22-38, 2021.
[5] N. Jindal and B. Liu, "Opinion spam and analysis," Proceedings of the international conference on Web search and web data mining - WSDM '08, vol. 8, pp. 219–230, California, 2008.
[6] E. Fox, Emotion science: Cognitive and neuroscientific approaches to understanding human emotions. London: Palgrave Macmillan, 2008.
[7] N. Gupta, M. Gilbert, and G. D. Fabbrizio, "Emotion detection in email customer care", Computational Intelligence, vol. 29, no. 3, pp. 489–505, 2012.
[10] G. Leech, Principles of pragmatics. London: Longman, 1983.
[8] G. Pitsilis, H. Ramampiaro, and H. Langseth, "Effective hate-speech detection in Twitter data using recurrent neural networks", Appl. Intell., vol.48, 2018, pp. 4730–4742.
[9] P. Badjatiya, S. Gupta and M. Varma, "Deep Learning for Hate Speech Detection in Tweets", In Proceedings of the 26th International Conference on World Wide Web Companion, Perth, Australia, 3–7 April 2017; vol 2, pp. 759–760
[10] W. Warner and J. Hirschberg, "Detecting Hate Speech on the World Wide Web". In Proceedings of the Second Workshop on Language in social media,7-9 June 2012, vol. 7, pp. 19–26
[11] S. Frenda, B. Ghanem, M. Montes-Y-Gómez, M and P. Rosso, "Online hate speech against women: Automatic identification of misogyny and sexism on twitter". J. Intell. Fuzzy Syst. 2019, vol.36, pp. 4743–4752
[12] H. Mulki, H. Haddad, C. Bechikh Ali, and H. Alshabani, "L-HSAB: A Levantine Twitter dataset for hate speech and abusive language", Proceedings of the Third Workshop on Abusive Language Online, vol. 3, 2019, pp. 251–263.
[13] S. Almutiry and M. Abdel Fattah, "Arabic cyberbullying detection using Arabic sentiment analysis", The Egyptian Journal of Language Engineering, vol. 8, no. 1, pp. 39–50, 2021.
[14] J. Culpeper, "Impoliteness and hate speech: Compare and contrast", Journal of Pragmatics, vol. 179, pp. 4–11, 2021.
[15] G. Leech, The pragmatics of politeness. Oxford: Oxford University Press., 2014.
[16] R. Watts, Politeness. Cambridge: Cambridge Univ. Press, 2009.
[17] G. leech, "Towards an anatomy of politeness in communication", International Journal of Pragmatics, vol.14, pp. 101-123, 2003.
[18] E. Wenger, Communities of Practice: Learning, meaning, and identity. Cambridge: Cambridge University Press, 2018.
[19] S. Mills. Discursive approaches to politeness. Berlin: De Gruyter Mouton, 2011.
[20] M. A. Locher and R. J. Watts, "Politeness theory and relational work," Journal of Politeness Research. Language, Behavior, Culture, vol. 1, no. 1, pp. 9–33, 2005.
[21] H. P. Grice, in Logic and conversation, New York: Academic Press, 1975, pp. 20–30.
[22] R. J. Kreuz, Irony and sarcasm. Cambridge (MA): The MIT press, 2020.
[23] G. Held, "Politeness in linguistic research", Politeness in Language. Berlin, New York: De Gruyter Mouton.2005, pp.131–154
[24] J. Thomas, Meaning in interaction: An introduction to pragmatics. London: Longman, 1995.
[25] E. Goffman, "On face-work", Interaction Ritual, Rome: Pantheon, 2017, pp. 5–46.
[26] P. Brown and S. C. Levinson, Politeness: Some universals in language usage. Cambridge: Cambridge Univ. Press, 1994.

BIOGRAPHY

Samar Assem a Passionate Linguist and a promising Language Engineer who works a lecturer of linguistics in the faculty of languages, MSA University. A PhD candidate in the department of phonetics and linguistics, the faculty of Arts Alexandria university. Her main areas of interests are Computational linguistics, corpus linguistics, pragmatic analysis and sociolinguistics. Got an Applied Linguistics diploma from the faculty of language and communication from AAST in 2020. Got an MA in Applied Linguistics with honours, faculty of Arts Ain Shams university. Honoured for being the youngest MA holder from the faculty of Languages, MSA University.

Dr. Sameh Alansary is the Professor of computational linguistics in the Department of Phonetics and Linguistics and the head of Phonetics and Linguistics Department, Faculty of Arts, Alexandria University. His main areas of interest are concerned with computational linguistics and language engineering, corpus work, morphological analysis and generation, and building formal grammars. He established the Arabic Computational Linguistics Center in Bibliotheca Alexandrina. Prof. Alansary is the co-founder and Vice president of board of the Arabic Language Technology Center (ALTEC) in Cairo, an NGO aims at providing Arabic Language resources and building a road map for Arabic Language Technology in Egypt and in the Middle East. He is and a member in many scientific organizations: (1) Egyptian Society of Language Engineering, Cairo, (2) Arabic Linguistic Society - USA, (3) Association of Computational Linguistics - USA – Europe, (4) Universal Networking Language foundation, United Nations, Geneva, Switzerland.

TRANSLATED ABSTRACT

تطور تحليل المشاعر من وجهة نظر لغوية

[1] سامح الأنصاري[2] سمر عاصم،
قسم الصوتيات واللغويات، كلية الآداب، جامعة الإسكندرية، الإسكندرية، مصر
[1]samar.assem@alexu.edu.eg
[2]s.alansary@alexu.edu.eg

ملخص

على الرغم من أن تحليل المشاعر بحكم التعريف هي فرع من فروع معالجة اللغة الطبيعية والذي يركز على تحليل النصوص التي تتناول تقييم وتحليل وكشف الحالة الذهنية للبشر تجاه مجموعة من المجالات، فإن معظم الدراسات تقتصر على تحليل الرأي. ومع ذلك، فإن تحليل الرأي ليس سوى مجال فرعي واحد من ثلاث مجالات أخرى تحت مظلة تحليل المشاعر؛ تحليل الآراء، وتحليل العواطف، اكتشاف الغموض اللغوي. ومن الملاحظ أن اكتشاف الغموض اللغوي يعتبر مزيجا من المجالين الفرعيين الآخرين بفضل طبيعته اللغوية التي تعتبر مستويات التحليل الإحصائية و/أو النحوية الدلالية غير كافية للوصول إلى مستوى مرضٍ من اللغة البشرية المعقدة. وبناء عليه، تقترح الورقة البحثية الحالية البحث بعمق للوصول إلى مستويات عملية واجتماعية براغماتية من التحليل من أجل القضاء على الغموض وتجنب سوء التقدير على النصوص ومنشورات وسائل التواصل الاجتماعي على وجه التحديد في المهام الفرعية للكشف عن السخرية وخطاب الكراهية. أخيرًا، يقترح استخدام نموذج لغوي انتقائي للتحليل يتضمن نظرية فعل الكلام ونظرية التأدب واللياقة.

الكلمات المفتاحية: تحليل المشاعر، السخرية، خطاب الكراهية، البراجماجتماعية، التأدب