

Multimodal Metadiscourse: A Model Development and Application to Spoken Political Discourse

الخطاب عن الخطاب متعدد الأنماط: تطوير نموذج وتطبيقه علي الخطاب السياسي الشفهي

Reham Mohamed El-Said Khalifa

Department of English Language

Faculty of Arts

Damietta University

reham_khalifa@du.edu.eg

المستخلص

تلعب عناصر الخطاب عن الخطاب دورًا حيويًا في التواصل الفعّال من خلال مساعدة الآخرين على الفهم والتقييم والاستجابة للمعلومات التي ينقلها الكاتب أو المتحدث. إلا إنها ليست مقتصرة على التعبيرات اللغوية فقط؛ بل يمكن أن تشمل مجموعة متنوعة من صيغ التواصل المختلفة. ومع ذلك، كانت الأبحاث السابقة تركز بشكل أساسي على الجانب اللغوي منها، مما يفتقر إلى إطار شامل لتحليل الخطاب عن الخطاب بشكل دقيق في سياق متعدد الأنماط. لذلك، هدفت الدراسة الحالية إلى تحديد نموذج مناسب لدراسة جوانب الخطاب عن الخطاب في الخطابات الشفهية، وبشكل خاص في السياقات السياسية، من منظور متعدد الصيغ (الأنماط). من أجل بناء هذا النموذج، تم استعراض الأدبيات السابقة لجمع سمات الخطاب عن الخطاب المناسبة للخطاب الشفهي. تم دمج هذه السمات في نموذج شامل وتم التحقق من مدى مناسبتها من خلال تقييم الخبراء وتطبيقها على مثالين من الخطابات السياسية الشفهية. المثال الأول تمثل في خطاب سياسي باللغة الإنجليزية للرئيس الأمريكي السابق دونالد ترامب خلال حملته الانتخابية التمهيدية لانتخابات الولايات المتحدة ٢٠٢٤، والمثال الثاني تمثل في خطاب سياسي باللغة العربية للرئيس المصري عبد الفتاح السيسي في القمة الدولية للحكومات والتي انعقدت في الامارات العربية المتحدة في يناير ٢٠٢٣. وأظهرت النتائج أن النموذج المقترح قابل للتطبيق على الخطابات السياسية الشفهية وقد ساهم النموذج في توضيح خصائص الخطاب عن الخطاب التي تم استخدامها من قبل الرئيسين في خطاباتها السياسية.

الكلمات المفتاحية: الخطاب عن الخطاب، الخطاب متعدد الأنماط (الصيغ)، الخطاب السياسي، الخطاب الشفهي

Abstract

Metadiscourse plays a pivotal role in effective communication by aiding others in understanding, assessing, and responding to the information conveyed by a writer or speaker in the intended manner. It is not confined solely to linguistic expressions; rather, it can encompass a combination of various modes within multimodal contexts. Nevertheless, prior research has predominantly focused on linguistic metadiscourse, lacking a comprehensive framework for a thorough analysis of metadiscourse in a multimodal context. Hence, the current research aimed to identify an appropriate model for scrutinizing spoken discourses, particularly in political contexts, from a multimodal perspective. To construct this model, existing literature was reviewed to compile metadiscourse characteristics suitable for spoken discourses. These characteristics were integrated into a comprehensive model and validated through expert assessment of and application to two instances of spoken political discourse. The first example was an English political speech delivered by Donald Trump during his preliminary campaign for the 2024 USA elections. The second was an Arabic political speech by the Egyptian President Abdel-Fattah Alsisi at the International Government Summit in the UAE. The findings demonstrated that the suggested model is applicable to spoken political discourse, effectively uncovering the metadiscourse features employed by these two politicians in their speeches.

Keywords: *metadiscourse, multimodality, spoken discourse, political discourse, multimodal metadiscourse*

1 Introduction

"Metadiscourse", a term first coined by the American Linguist Harris in 1959, provides a means of comprehending language in use. It encompasses the writer's endeavors to influence how a receiver perceives a text (Hyland,2005, p.3). This term has emerged within the realm of applied linguistics (Ädel & Mauranen, 2010), and it has subsequently been researched in a number of languages and genres (Hyland, 2005, p.6). It has drawn scholars from a variety of academic disciplines, including linguistics, rhetoric, discourse analysis, pragmatics, and language pedagogy (Vande Kopple, 2012). The current research focusses on the analysis of metadiscourse in spoken political discourse from a multimodal perspective.

2 Defining metadiscourse

The word metadiscourse consists of the prefix "meta", which means beyond in Greek and "discourse". In this sense metadiscourse is used to denote the realization of language that goes beyond the boundaries of mere discourse (Bernad-Mechó, 2018, p.123). Different alternatives exist for the term "metadiscourse." For instance, Rossiter (1974) uses "metacommunication" to encompass verbal and non-verbal elements, including tones and intonations. Lautomatti (1978) refers to metadiscourse as "discourse connectives," while Mauranen (1993) employs the term "text-reflexives." Schiffrin (1980) uses "metatalk," and Jakobson (1985, pp.113-121) uses "metalanguage." The term "metadiscourse" is employed by Williams (1981, p.40), Vande Kopple (1985),

Crismore et al. (1993), Hyland (2005), and Ädel (2006, pp. 215–218).

Metadiscourse can be described as the linguistic framework used within communication, whether written or spoken, to acknowledge and address the act of communication itself and the surrounding context. It transcends the mere transmission of information and encompasses the individualities, attitudes, and assumptions of the communicators (Williams, 1981, p.40). Metadiscourse plays a crucial role in adapting communication to suit the intended audience, taking into account their knowledge, potential objections, and cognitive processing requirements. It provides valuable insights into the communicative setting and the perceptions of the participants involved (Chen & Li, 2023).

The concept of "recipient design" is highlighted by Hyland et al. (2022) as the foundation of metadiscourse. This concept reflects the importance of tailoring communication to the specific individuals engaged in the interaction, acknowledging that language is influenced by the dynamics and distinctions between people. Metadiscourse serves as a mechanism for expressing and constructing these interactions, whether in spoken or written form. Individuals engage in a process of negotiation, consciously making decisions about the impact they have on their listeners or readers. Hyland (2005, p. 29) further explains that the concept of metadiscourse is relative, meaning that certain elements within a text function as metadiscourse in relation to other parts of the same text. Therefore, what qualifies as metadiscourse in one context may not be considered as such in another context.

Additionally, Crismore et al. (1993) define metadiscourse as linguistic elements present in texts, both written and spoken, that do not directly contribute to the propositional content. Instead, they facilitate the organization, interpretation, and evaluation of the information provided for the listener or reader. This definition draws upon Halliday's functions of language, distinguishing metadiscourse from language solely focused on conveying propositions. It also underscores the organizational function of metadiscourse as a tool for structuring and presenting language.

Therefore, it could be concluded that the term "metadiscourse" refers to the tools and techniques used in communication to recognize and discuss the process of communication itself. It goes beyond simply sharing information and takes into account the unique qualities of the individuals involved. Metadiscourse is crucial in customizing communication to engage the desired audience, taking into consideration their understanding and cognitive requirements. It is influenced by the interactions and differences between people and serves as a means to organize and enhance the comprehension and assessment of information.

3 Spoken vs. written political discourse

Spoken discourse and written discourse are distinct forms of communication. Spoken political discourse differ from written political discourse in several ways. Firstly, spoken political discourse, such as speeches or debates, is more immediate and spontaneous, allowing speakers to react to current events or the audience's reactions in real-time. Written political discourse, like

articles or policy papers, often has a longer production process, making it less immediate (Kenzhekanova, 2015). Secondly, spoken discourse often targets a broader audience, including the general public. In contrast, written discourse can target specific readers with a deeper interest in the topic, allowing for more complex arguments and detailed analysis (Farahani & Kazemian, 2021). Thirdly, spoken discourse has a lower content density, with complex phrasing, clause structures, and grammatical intricacies. In contrast, written discourse has a higher lexical density, represents the target product, and has less complex and more eligible grammatical structure (Halliday, 1985, 76-87). Fourthly, spoken discourse relies on different modes to deliver the target message such as the speaker's tone, gestures, and body language, while written discourse depends solely on the written word, requiring skilled writing to express nuanced ideas (Sindoni, 2013, pp. 2-5). Fifthly, written political discourse allows for careful editing and fact-checking, reducing the likelihood of errors. Spoken discourse, especially in live situations, may be more prone to immediate mistakes or unprepared statements (Kenzhekanova, 2015). Due to the distinctive characteristics of spoken political discourse, the use of metadiscourse in spoken discourse may be different from that in written discourse. Therefore, it would be useful to investigate how metadiscourse is used in spoken political discourse.

4 Metadiscourse and multimodality

In contemporary communication, humans engage in complex interactions that go beyond the mere exchange of words (Kraus & Slater, 2016). The study of multimodality offers a rich

perspective on how meaning is created and conveyed through multiple modes of expression (Lyons, 2016). The concept of multimodality has its roots in the social semiotics theory, which explores the various ways that meaning may be created in interaction. The word "semiotics" is derived from the Greek word "semeion," which means "sign" and denotes the smallest unit of meaning produced by the fusion of several modes, including visual, audible, and physical ones (Gualberto & Kress, 2019).

Jewitt and Henriksen (2016, p. 145) explained that Social Semiotics has its roots in Functional Linguistics, most notably in the work of Halliday (1978), and was further expanded as a theory of multimodal sign-making in the works of Hodge and Kress (1988), Van Leeuwen (2005) and Kress and Van Leeuwen (2006). The term multimodality is used to describe the interplay between different modes of expression to convey meaning comprehensively and nuancedly (Jewitt & Henriksen, 2016, pp. 146-148). Modes, whether verbal or non-verbal, are semiotic resources for meaning creation that are socially formed and culturally imparted (Kress, 2010, p. 79). Verbal cues include all forms of written or spoken words. Non-verbal cues, including gestural cues, gaze, body language, sound pauses and syllabic duration, play a significant role in amplifying, refuting, or reinforcing verbal messages (Martin & Zappavigna, 2019). Social semiotics explores the function of each mode as well as how various modes are connected to one another to provide an in-depth analysis of meaning in a multimodal discourse (Kress, 2010, p. 59).

Metadiscourse is a linguistic phenomenon that focuses on the interaction between the writer/speaker and the readers/audiences. Therefore, different modes could be employed to emphasize this interaction. The multimodal analysis of metadiscourse has attracted the attention of different researchers. For example, Carriro–Pastor (2022) sought to analyze metadiscourse textually and visually in the written essays of learners of English for academic purposes. Ribeiro et al. (2022) examined the definitions of words in glosses from a multimodal viewpoint as the metadiscourse in glosses consists of a text accompanied by visual aspects like layout, color, font, and picture. Alyousef (2015) investigated metadiscourse features in texts written by international postgraduate students. He analyzed these features verbally based on Hyland's model (2005) and visually based on Van Leeuwen's model (2005). Similarly, Isalambo and Kenneth (2020) explored textual and visual metadiscourse used in the Covid–19 campaign from a multimodal point of view using Hyland's model and Kress and Van Leeuwen's model (2006). Bernad–Mechó (2018) conducted a multimodal analysis of spoken academic discourse, analyzing it verbally using Ädel's model (2010) and non-verbally using Kress and Van Leeuwen's model.

5 Research problem

Since its emergence in 1959, researchers have shown interest in metadiscourse. However, most of the research has primarily focused on analyzing metadiscourse in written academic discourse (e.g., Crismore et al., 1993; Toumi, 2009; Ebrahimi, 2018; Qin and Uccelli, 2019), with some studies exploring spoken discourse, but only on the verbal level like Ilie (2003), and Abusalim et al.

(2022). When it comes to analyzing metadiscourse from a multimodal perspective, the focus has mainly been on written discourse (e.g., Alyousef, 2015; Isalambo and Kenneth, 2020), with few studies that have delved into the multimodal analysis of metadiscourse in spoken discourse such as Bernad-Mechó (2018).

Reviewing research on multimodal analysis of metadiscourse revealed that there is no comprehensive model available for analyzing metadiscourse from a multimodal perspective, particularly in spoken political discourse. This study is concerned with suggesting a comprehensive model that could be used as a framework to analyze metadiscourse in spoken political discourse, with the intention of testing its applicability by using it to analyze samples of spoken political discourse.

6 Research questions

The current research seeks to answer the following questions:

1. What is the suggested multimodal model that could be adopted to analyze metadiscourse in spoken political discourse?
2. How could this model be applied to the spoken political discourse?

This question could be subdivided into the following questions:

- a) According to the suggested model, what are the forms of metadiscourse employed verbally within the two provided samples?

- b) Based on the suggested model, what varieties of metadiscourse are employed non-verbally in the two samples of spoken political discourse?

7 Methods

A survey of previous studies revealed that many researchers have been interested in multimodal analysis of metadiscourse. However, there is currently no comprehensive model for analyzing metadiscourse in general and specifically in spoken political discourse using multimodal methods. Most of the earlier research utilized Kress and Van Leeuwen's visual design approach, which is suitable for analyzing static images and pre-designed cinematic shots. However, this approach may not fully encompass the complexities of spontaneous spoken political discourse, which includes both verbal and non-verbal cues. Thus, there is a need to develop a procedural model that can encompass all aspects of spoken political discourse from a multimodal perspective. The next section will focus on extracting various features of metadiscourse in spoken political discourse and organizing them into a model suitable for multimodal analysis.

To create such a model, previous research related to verbal and non-verbal features of metadiscourse were reviewed to sum up all the verbal and non-verbal features of metadiscourse. Then, these features were collected in a comprehensive model. Exploring the validity of the model passed through two phases. The first one is sending the model to a number of experts in the field of linguistics to find out whether the selected features are suitable, unsuitable or

should be modified. The second phase was applying the final version of the model to two samples of spoken political discourse. The first sample is a speech delivered by Donald Trump, the ex-American President, on 16 July, 2023, as a preliminary meeting for his 2024 election campaign. In this speech, Trump tried to affect his audiences to get their support in his election campaign. The second sample is a speech delivered by Abdel Fattah el-Sisi, The Egyptian president since 2014, in the conversational session in the international summit of governments in the UAE. In this speech, el-Sisi tried to affect various types of audiences, the presidents present in the summit, the Egyptian populace who listen to this speech and the investors who follow the summit.

8 Developing the model

8.1 Verbal metadiscourse

Various models have been created to examine the verbal aspects of metadiscourse. The prevalent models for verbal metadiscourse encompass Vandekopple's, Hyland's, Ädel's, Crismore et al.'s, and Dafuz-Milne's models. The subsequent section provides an overview of these models.

8.1.1 Vandekopple's model

Vandekopple (1985) created a classification of metadiscourse that was based on the suggestions of Lautamatti (1978) and Williams (1981) (Alzarieni et al., 2019). This classification includes seven categories which could be classified into two major classes: textual and interpersonal (Toumi, 2009). According to Vandekopple (1985), Textual metadiscourse encompasses four categories which

are *textual connectives*, *code glosses*, *illocution markers* and *validity markers*. *Text connectives* consist of six subcategories. The first one includes sequence words like "first," and "second," which indicate the order of the text. The second subcategory encompasses words or phrases that express logical and/or temporal connections. The third subcategory involves reminders of prior content (e.g., "as previously mentioned", and "as said at the beginning"). The fourth subcategory is material announcement, such as "as we will see later." The final subcategory within text connectives is topicalizers, like "for example" and "as for."

The second textual category in Vande Kopple's model, *code glosses*, assists readers in understanding the meaning of textual components, such as the use of brackets to define something. The third category is *illocutionary markers* which include words or phrases that clarify to the audience the speech or discourse acts performed by the writer like "we might summarize" or "we can hypothesize." The fourth category is *validity markers*, which communicate the writer's opinion on the validity of the presented propositional content. It consists of three subcategories: hedges, emphatics, and attributers.

The second class of Vande Kopple's model is the interpersonal metadiscourse which encompasses three categories: *narrators*, *attitude markers* and *commentary*. *Narrators* involve phrases like "Mrs. Wilson said that" or "according to James." *Attitude markers* include phrases like "surprisingly" or "I find it interesting." The last category is *commentary*, which directly addresses the readers. It includes four subcategories: remarks on the readers'

potential mood, view, or reaction; suggestions for a mood or procedure; expectations-setting; and comments on the readers' actual relationship to the writer.

8.1.2 Crismore et al.'s model

Crismore et al. (1993) adopted Vande Kopple's model but with some modifications. As in Vande Kopple's model, Crismore et al.'s model consisted of two major classes: Textual and interpersonal. However, the textual class is sub-classified into two main categories: *textual markers* and *interpretive markers*. The *textual markers* include: *logical connectives*, which joins the parts of the sentence and help the readers interpret like the use of "moreover" and "consequently", *sequencers*, which arrange the text in logical sequence like "first" and "second", *reminders*, which refers to previously mentioned information like "as I mentioned before" and *topicalizers*, which introduce new topics like "As for X"

The second category of the textual metadiscourse is the *interpretive markers* which cover three types: *code glosses*, *illocutionary markers* and *announcements*. *Code glosses* are used to explain the text, while *illocutionary markers* are used to clarify the speech act performed in the text. *Announcements* are used to preview information (e.g., "as I will explain in the following").

Interpersonal metadiscourse include *hedges*, *certainty markers*, *attributors*, *attitude markers* and *commentary*. *Hedges* are used to show the validity of the text from the writer's point of view, whereas *certainty markers* show the writer's full commitment to the text. *Attributors* refer to the text source, *attitude markers* indicate

the writer's stance on statements like "unfortunately" and "I hope", and *commentary* is used to directly refer to the reader.

8.1.3 Hyland's model

Hyland (2005, pp. 38-43) in his model of metadiscourse outlined three key tenets. According to Hylands' model, metdiscourse stands apart from the text's propositional act, it conveys writer-reader interactions, and it alludes to both internal and external relations within the text.

Hyland and Tse (2004) Claim that all the categories of metadiscourse are fundamentally interpersonal as readers' knowledge, experiences and needs should be considered. Therefore, Hyland (2005, 43-60) adopted a new terminology for classifying metadiscourse. He substituted textual metadiscourse with interactive metadiscourse and interactional metadiscourse in place of interpersonal metadiscourse. Interactive metadiscourse in Hyland's model refers to the forms of metadiscourse used by the writer to guide readers through discourse by taking into account their knowledge, needs, and interests. There are five subcategories of interactive metadiscourse. *Transition markers*, which are used to describe relationships between phrases and sentences, are the first interactive sub-category. *Frame markers*, such as "let's talk about," "to conclude that," and others, are used to highlight subject switches, establish discourse aims, and arrange arguments. *Endophoric markers*, which make reference to other parts in the text, are the third subcategory. The fourth subcategory is *evidentials*

which are used to identify the information's source (such as "according to"). The last interactive subcategory is *code glosses*, which are intended to clarify or rephrase text for users (using expressions like "in other words" or "for example").

In Hyland's model, interactional metadiscourse which are used to engage the reader and demonstrates attitude, is the second major category of metadiscourse. This category consists of two main subcategories: *stance* and *engagement*. *Stance* metadiscourse include hedge words like "may," "perhaps," "possible," and others that imply reluctance to commit and encourage discussion. *Engagement*, the second interactional metadiscourse, has four subcategories. One interactional subcategory known as *boosters* places a strong emphasis on certainty (e.g., "definitely," "in fact"). *Attitude markers*, such as "unfortunately" and "I agree," are the second interactional metadiscourse subcategory. *Self-mention*, which directly alludes to the author, is the third subcategory. The fourth set of markers is called *engagement markers*, and it consists of readers' pronouns, personal asides, questions, and instructions, and allusions to previously learned information (such "consider" and "note that").

8.1.4 Ädel's model

Ädel (2010) proposed a taxonomy of spoken and written metadiscourse that consists of four basic types and 23 subclasses. The first one is metalinguistic remarks, which include *commenting on linguistic form*, *repairing*, *clarifying*, *reformulating* and *managing terminology*. The second major category is discourse

organization which is divided into *managing topics* and *managing phorics*. *Managing topics* include introducing topic, adding to topic, delimiting topic, marking asides, and concluding topic. *Managing phorics* include endophoric marking, reviewing, previewing and contextualizing.

The third major category is speech act labels which embrace *arguing*, *exemplifying* and *speech act labeling*. The last major category is references to audiences which include *managing comprehension*, *managing the message*, *managing audience discipline*, *imagining scenarios* and *anticipating the audiences' response*.

According to Ädel, the use of metadiscourse varies across spoken and written discourse. For instance, the discourse functions of repairing, marking asides, and contextualizing were more prevalent in written discourse, while managing comprehension/channel and managing audience discipline occurred only in the spoken discourse where there is direct presence of audiences.

8.1.5 Dafouz–Milne's model

Dafouz–Milne (2008) elaborated a comprehensive taxonomy of written and spoken metadiscourse that depend to a great extent on previous taxonomies of Vande Kopple, Crismore et al. and Hyland. Similar to Vande Kopple and Crismore et al., Dafouz–Milne classified metadiscourse into textual and interpersonal. Textual metadiscourse is classified into seven subcategories. The first one is *logical markers*, which are used to show the semantic link between discourse parts, include additives, adversatives, consecutives and conclusive. *Sequencers* are a second type of

textual metadiscourse that are used to identify positions in a succession of texts. The third one is *reminders* which allude to previous sections in the text. *Topicalizers*, the fourth category, signal a change in the topic. *Code glosses* are the fifth type, and they are used to clarify, restate, or provide examples of language. *Illocutionary markers*, which explicitly indicate the discourse acts performed by the writer, make up the sixth subcategory. *Announcements*, which make reference to portions of the text that will come later, are the final textual subdivision. Dafuz–Milne subdivided interpersonal metadiscourse into *hedges*, *certainty markers*, *attributors*, *attitude markers* and *commentary*.

8.1.6 Towards a comprehensive model for verbal metadiscourse

Analyzing various models of verbal metadiscourse reveals that each one focuses on specific aspects of metadiscourse. While Ädel's model is unique in focusing on spoken verbal discourse, it does not encompass all metadiscourse aspects, and other models also highlight features applicable to spoken discourse. Consequently, there's a necessity for a comprehensive model encompassing all aspects of spoken metadiscourse. To fulfill this need, the previously mentioned metadiscourse aspects, which align with spoken discourse, are integrated into a comprehensive model, facilitating the analysis of spoken discourse, particularly in the realm of politics.

Verbal metadiscourse could be divided into two primary categories: interactive metadiscourse (Hyland, 2005) and interactional metadiscourse (Crismore, 1993; Hyland, 2005; Ädel, 2010). The first one is interactive metadiscourse which encompasses

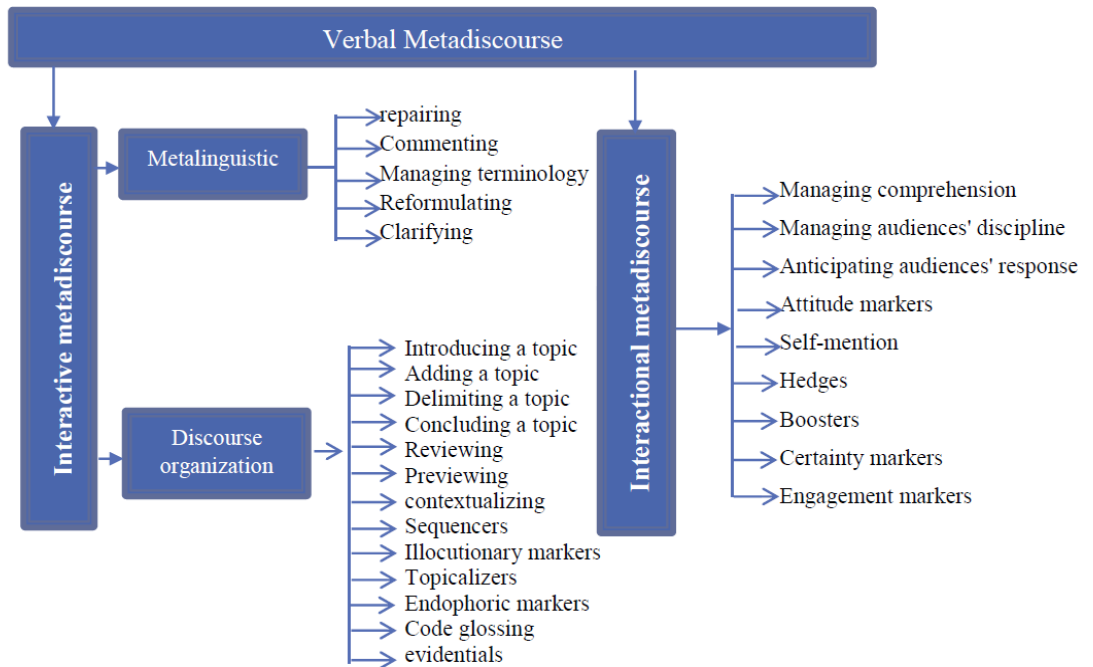
metalinguistic comments and *discourse organization*. The metalinguistic comments can be further classified into five sub-categories: *repairing* (correcting errors), *reformulating* (offering alternative phrasings for better comprehension), *commenting on linguistic forms* (using expressions like "but I mean that"), *clarifying* (making meaning explicit for better understanding), and *managing terminology* (defining and explaining terms) (Ädel, 2010). Similarly, discourse organization markers fall into 13 sub-categories that encompass various functions. These categories involve *introducing a topic*, *delimiting a topic*, *adding a topic*, *concluding a topic*, *making asides*, *previewing*, *contextualizing* (Ädel, 2010), *reviewing* or *reminding* (Dafouz-Milne, 2008; Ädel, 2010), *transition markers* or *sequencers* (Hyland, 2005; Dafouz-Milne, 2008), *endophoric markers*, *evidentials* (Dafouz-Milne, 2008; Ädel, 2010), *illocutionary markers* (Vande Kopple, 1985; Dafouz-Milne, 2008; Ädel, 2010), and *topicalizers* (Crismore, 1993; Dafouz-Milne, 2008).

The second subcategory of verbal metadiscourse is referred to as "interactional metadiscourse" which pertains to the ways writers/speakers employ to engage with their readers/audiences (Hyland, 2005). Interactional metadiscourse encompasses several aspects, such as *managing comprehension channel* (checking the effectiveness of communication channels), *managing audiences' discipline* (providing instructions or feedback), *anticipating audiences' reactions*, *using imaginative scenarios* (conveying ideas through vivid experiences, such as asking readers to imagine a situation) (Ädel, 2010), *using attitude markers* (Vande Kopple,

1985; Crismore, 1993; Dafouz–Milne, 2008), *self-mentioning*, *expressing uncertainty* (through hedging), *using attributors*, *conveying certainty* (through boosters or certainty markers) (Crismore, 1993; Hyland, 2005; Dafouz–Milne, 2008), and *engagement markers* which engage the recipients of the speech through the use of rhetorical questions, direct address to the recipients, inclusive expressions, and personalizers (Hyland, 2005). All the previous categories could be summarized in figure 1.

Figure 1

A comprehensive view of verbal metadiscourse



8.2 Non-Verbal Cues

Face-to-face interactions are enriched by a plethora of nonverbal behaviors, which play a vital role in shaping the exchange. These behaviors encompass a wide range of non-verbal

cues (Cafaro, et al., 2019). Non-verbal cues serve various purposes during face-to-face communication. These cues have the potential to replicate spoken messages, contradict them, replace them (such as conveying more vivid emotions through one's eyes), enhance them by adding to or complementing verbal communication (like a boss combining praise with a pat on the back for greater impact), and emphasize points through actions like pounding the table (Dash, 2022).

Querol-Julián & Fortanet-Góme (2012), Windsor (2017), and Cafaro et al. (2019) classified non-verbal cues into non-vocal features or kinesics and vocal or paralinguistic features. Non-vocal features/kinesics help create dynamic communication channel between participants through the use of gestures, facial expressions, postures, and gazes. On the other hand, vocal/paralinguistic features cover features like intonation, stress, pitch, rhythm, loudness, syllabic duration, laughter and pauses. In the next section, each of the previously mentioned features is analyzed.

8.2.1 Non-vocal features

A. Gestures

Gestures refer to actions carried out using hands, arms, and shoulders (Poggi, 2008). Gestures and speech are found to be intricately linked to each other, and the human brain processes both in comparable ways (Gullberg et al., 2008). In communication, gestures play a linguistic role by providing referential content through deictic expressions, occupying specific parts in an utterance's structure, and influencing or modifying speech acts

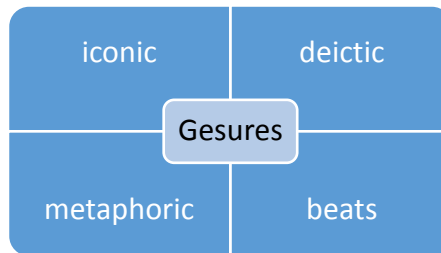
(Gullberg et al., 2008). Trotta and Guarasci (2021) indicated that gestures are utilized in political face-to-face discourse to supplement speech, complement incomplete verbal information, integrate the speech in order to emphasize the message, or reinforce previously expressed concepts, thus avoiding redundancy. Several studies have attempted to present a typology for gestures. However, due to their multifaceted nature, McNeil (2005, p. 42) put forth a dimensional framework to categorize gestures, involving four main dimensions: iconicity, metaphoricity, indexicality, and temporal highlighting.

The first dimension which is *iconicity* refers to the connection between a sign and an object, where the sign's form is perceived and understood as resembling the object it represents (Mittelberg & Evola, 2014). The second dimension is *metaphoricity*. Metaphoricity and iconicity share similarities as they both involve pictorial content, but metaphoric gestures convey abstract content rather than physical objects (Gullberg et al., 2008). The third dimension is *indexicality* that describes the use of emblematic or deictic gestures to position objects and actions in relation to a reference point in space. While the extended index finger is the prototypical example of pointing, other body parts can also be utilized (McNeil, 2005, pp. 39-40). The last dimension is *temporal highlighting*, often referred to as beats or batons, involves quick and straightforward hand movements resembling time-beating actions. These flicks of the hand, performed in an up and down or back and forth manner, appear to beat time along with the

speech's rhythm (McNeil, 2005, p. 40). Based on the previous analysis of gestures, a summary of the dimensions of gestures is presented in figure 2.

Figure 2

The four dimensions of gestures



B. Gazes

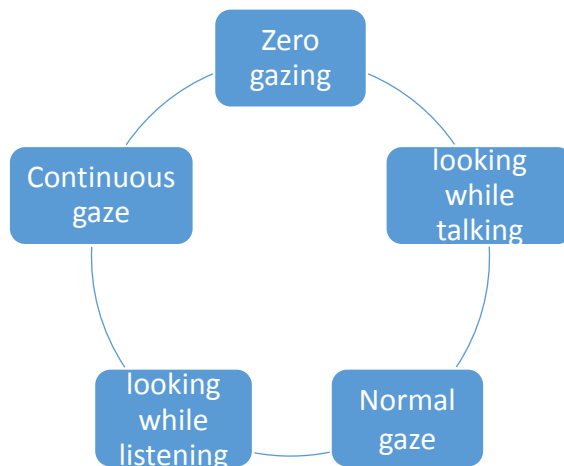
Gazing refers to the act of orienting one's eyes towards a particular point in the visual surroundings (Hessels, 2020). A gaze is a conscious activity that results as a reaction to someone or something in the surroundings (Lund, 2007). Eye gaze is considered a unique stimulus in face-to-face interactions, and it serves a significant cue in the development of social engagement (Bailly et al., 2010). Gazes serve various purposes, such as relaying or acquiring information, eliciting specific responses from listeners (Lund, 2007), expressing interpersonal attitudes or prosodic

accompaniments of speech, reducing information overload, demonstrating intimacy or avoidance (Argyle et al., 1981), and managing turn-taking (Bailly et al., 2010).

The most commonly used gazes in face-to-face interactions include *continuous gaze*, *looking while talking*, *looking while listening*, *normal gaze*, and *nearly zero gaze*. The two patterns looking while listening and normal/spontaneous gazing is found to be the most appreciated patterns of the listeners. Looking while talking and zero gazing are found to be the least favorable patterns by the listeners. Continuous gazing is found to indicate speaker's activity, but it is less favorably as it may make the speaker more intrusive or trying to make the situation more intimate (Argyle et al., 1974). The types of gazes are summarized in figure 3.

Figure 3

The different forms of Gazes in face-to-face interaction



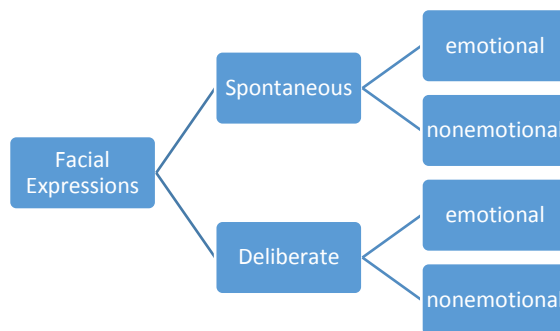
C. Facial expressions

Facial expressions refer to the changes in a person's face brought about by their intentions, inner emotions, or social interactions with others (Tian et al., 2005, p. 247). Human communication heavily relies on facial expressions, as they play a crucial role in conveying thoughts, ideas, and emotions. The face serves as a powerful tool for expressing these elements in our interactions (Frank, 2001).

Facial expressions could be categorized into two main dimensions which are spontaneous–deliberate dimension and emotional–nonemotional dimension (Hager & Ekman, 2005, p. 45). Although metadiscourse is intentionally employed to tailor communication for intended audiences (Chen & Li, 2023), spontaneous facial expressions could be adopted to adjust to any momentary incidents during face-to-face communications. Figure 4 summarizes the categories of facial expressions.

Figure 4

The main categories of facial expressions

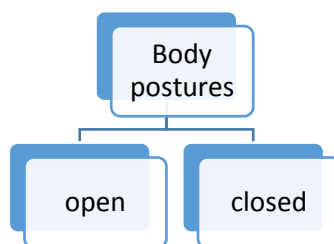


D. Body postures

Body posture refers to the manner in which the body is aligned, encompassing elements like standing or sitting orientation, the relative positioning of body parts, and the extent of space the body takes up (Collier, 2016). Body postures may be divided into two categories: open and closed. An open body posture becomes evident as the speaker takes on an open position, with outstretched palms, nodding, and leaning towards the audience (Pease & Pease, 2004, pp. 160–163). This open position reflects openness to conversation, friendliness, expressiveness, and interactivity (Bull, 1987, p. 29). Closed body posture is expressed with speaker's hands, legs, and arms crossed to the body (Pease & Pease, 2004, p. 162). This closed body posture may express disinterest to listen, discomfort with conversation and uncertainty. It reflects hostile and defensive attitudes (Collier, 2016). Forms of body postures are summarized in figure 5.

Figure 5

The forms of body postures



8.2.2 Vocal/Paralinguistic features

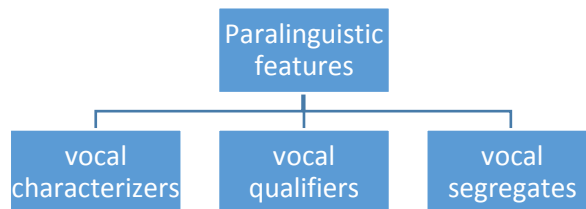
Paralinguistic features, sometimes referred to as vocalics, constitutes a facet of meta-communication that has the potential of modifying meaning, adding subtle connotations, or communicate

emotions through the use of intonation, pitch, volume, prosody, rhythm, etc. (Dash, 2022). Paralinguistic cues are used to communicate dominance and strength, emotional state, responsiveness, (Ethier, 2010, p. 18) and/or speaker's confidence (Guyer et al., 2021).

Paralinguistic features could be classified into three primary categories. The first category is "vocal characteristics" which refers to sounds that can function physiologically as a stress reliever and contribute to relaxation, especially during tense situations. These sounds can be understood as having a distinct meaning and can stand alone as symbols (e.g., laughter, yawning, crying, shouting, whining) (Bancroft, 1995). The second category is "vocal qualifiers" which include *intensity* (over loud to over soft), *volume* (high or low), *pitch height* (over high to over low) and *intonation* (rising to falling) (Dash, 2022). The last category of paralinguistic feature covers a large group of speech non-fluencies known as "voice segregates" or "voice interjections" and it includes sounds like "uh-huh," "urn," "ah," and variations thereof are included. This category of sounds functions as regulators of the flow of verbal communication (Bancroft, 1995). Figure 6 summarizes these categories.

Figure 6

Categories of paralinguistic features



8.3 Towards a comprehensive view of multimodal metadiscourse

After a comprehensive review of previous related research, it was determined that multimodal metadiscourse in spoken discourse could be divided into two categories: verbal metadiscourse and non-verbal metadiscourse. Verbal metadiscourse encompasses interactive and interactional elements. Interactive metadiscourse involves metalinguistic comments and tools for organizing discourse, while interactional metadiscourse encompasses all verbal forms that engage the listener in the conversation.

Non-verbal metadiscourse comprises non-vocal or kinesic attributes and vocalic or paralinguistic attributes. Non-vocal or kinesic attributes include all aspects of metadiscourse that rely on gestures, gazes, body postures, and facial expressions. Conversely, vocalics or paralinguistic features pertain to all sound elements (excluding words) that aid in listener engagement, further categorized into vocal characterizers, vocal qualifiers, and vocal segregates.

8.4 Validating the model: phase 1

To validate the suggested multimodal metadiscourse model in spoken political discourse, the model was sent to 10 experts in the field of linguistics via email (see appendix 1). Among these, six experts responded. One expert accepted the model as it is. Two experts considered "reviewing" and "reminders" as the same category without needing separation. Three experts believed that incorporation of "evidentials" into discourse organization is

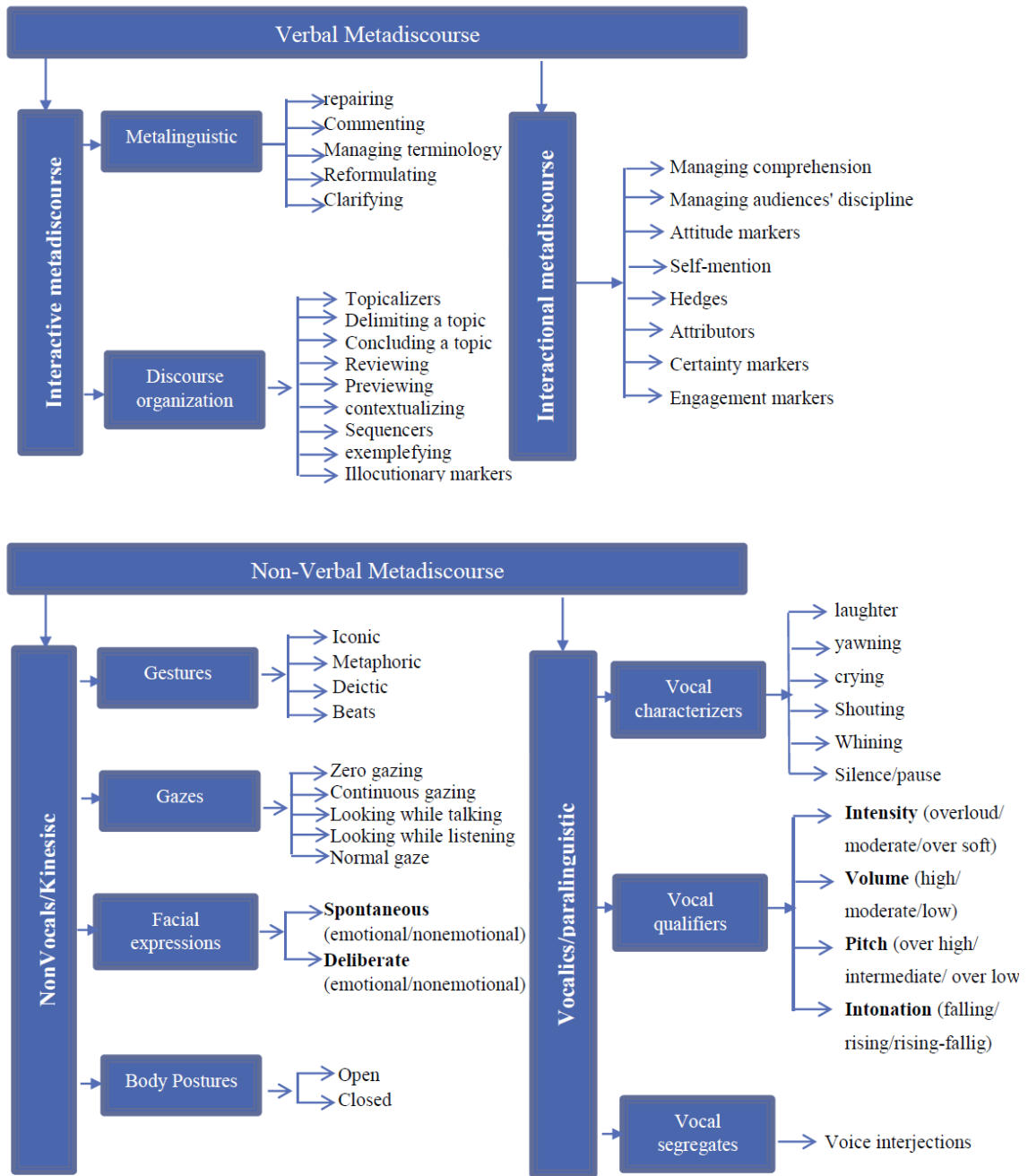
irrelevant. They advocated using "attributors" instead of "evidentials", classifying them as interactional metadiscourse elements. One expert recommended removing "endophoric markers". Four experts explained that code glossing covers punctuation and parentheses, which are not present in spoken discourse, so it should be removed. Two experts advised adding "exemplifying" to discourse organization. Five experts thought that using "inclusive expressions" covers "direct address to recipients", making a separate "direct address" category unnecessary. Similarly, four experts proposed that the "topicalizers" category covers "introducing a topic" and "adding a topic", so these two categories should be removed. Two experts found "boosters" and "certainty markers" to be equivalent, suggesting that one of them should be deleted. Moreover, three experts suggested that considering "anticipating audiences' reaction" is not an independent aspect; instead, it's encompassed within all the other metadiscourse categories. They clarified that whenever any metadiscourse category is employed, the speaker is essentially anticipating the audiences' response and selecting an appropriate approach to engage and influence them.

Regarding non-verbal cues, four experts suggested adding a "moderate" category for intensity. For volume, three experts recommended adding a "moderate volume" category. In terms of pitch, three experts advised adding an "intermediate" category, and for intonation, two experts recommended incorporating a "rising-falling intonation" category. Moreover, five experts proposed

adding "pauses"/"silence" as a category for vocal characterizers. The final version of the model is presented in figure 7.

Figure 7

The final version of the suggested model



8.5 Validating the model: Phase 2

To examine the applicability of the suggested model, it has been applied to two samples of spoken political discourse. The analysis is illustrated in the following section.

8.5.1 Application 1

The suggested model has been applied to Trump' 7-minute speech in his pre-election campaign on 16 July, 2023. In his speech, Trump aimed to refute the legal allegations against him and persuade his audience that he is the best choice for the future of the USA, contrasting himself with the current president, whom he portrayed as untruthful and tainted by corruption.

Verbal metadiscourse

Analyzing Trump's speech verbally revealed that Trump employed different forms of metalinguistic metadiscourse, which are repairing, commenting, reformulating and clarifying. These forms of are explicit in the following examples:

Ex1.1: "I never hit Biden as hard as I could" $\xrightarrow{\text{is repaired to}}$ "I just had too much respect for the office of presidency"

Ex1.2: "this will be the most important election with your country" $\xrightarrow{\text{is commented on as}}$ "your freedom and your future on the line we are in trouble this country is in trouble"

Ex1.3: "We're in trouble" $\xrightarrow{\text{is reformulated to}}$ "this country is in trouble"

Ex1.4: "whether you'll have a rule of tyrant or the rule of law" $\xrightarrow{\text{is reformulated to}}$ "whether Marxist radicals burn our civilization to the ground which they're looking for"

*Ex1.5: "whether your generation inherits a fascist country or a free country" **is clarified as** → "whether you will have a rule of tyrants or the rule of law"*

As for the discourse organization metadiscourse, Trump introduced his topic saying: "*16 months from now each of you will vote in the most important election of your lifetimes*". Then he added a topic, in which he is accusing the current American president of being dishonest and corrupt, saying: "*this is the most dishonest president in history*". Before the end of his speech, Trump made a preview for his speech saying: "*we will cast out the Communist Marxist fascists we will throw off the sick political class that hates our country and we will route the fake news media we will defeat crooked Joe Biden and we will drain the swamp once*". To declare the illocution of his speech, Trump used direct statements to indicate that he is stating facts (e.g., "*I say things about him over the last three weeks*"). Trump concluded his speech by extending blessings, appreciation and gratitude by saying "*God bless you all thank you very much*".

Regarding interactional metadiscourse, Trump initiated speaking subsequent to ensuring the microphone was functioning properly to establish an open line of communication. Throughout his speech, Trump utilized attitude markers on multiple occasions (for example, "*I should have to be honest*", "*we have to have it changed*", "*I would have never said*", and "*I think*"). Trump made reference to himself using the pronouns "*I*" and "*my*" approximately 25 times within his address. Similarly, he incorporated various engagement markers, employing inclusive language such as "*you*,"

"*your*," and "*we*" 42 times during his speech. Furthermore, he employed personal identifiers around four times (such as "*Charlie*," "*Biden*," "*each of you*," and "*all the people of attorney*"). Moreover, Trump made use of markers of certainty several times, employing repetition (for instance, "*I have pride*"), using the modal verb "will" (such as "*we will demolish the deep state*", "*we will expel the war mongers*"), and by using statements like "*he is corrupt horrible incompetent*". Trump ascribed certain actions to others using the attributor "*they*" nine times in his speech (like "*they took away my constitutional rights*").

The Non-verbal metadiscourse

Regarding the non-vocal cues, during the 7-minute speech, Trump refrained from using gestures for approximately 3 minutes. He incorporated beat gestures at 3:35 and 4:47 into his speech. At 3:23, Trump used a metaphoric gesture by opening his hands upwards, symbolizing his confidence and openness. At 3:56, he employed his index finger as an iconic gesture to indicate "No". Trump also utilized gazing while speaking and consistent gazing throughout his speech. This gazing while speaking demonstrated his attentiveness to all audiences. Trump frequently shifted his gaze direction to engage all audiences and display inclusiveness. Similarly, he maintained an open body posture throughout, conveying his dominance and strength. This stance also communicated his confidence and resilience in the face of accusations directed at him. Additionally, Trump displayed a fixed deliberate emotional facial expression, reflecting his emotional stability and indifference towards the allegations against him. This

expression also reflected his unwavering determination to contest in the 2024 elections.

As for the paralinguistic features, Trump utilized two vocal characterizers. Initially, he employed shouting during his speech to convey his dominance and power. Additionally, he strategically incorporated two instances of silence or pauses: one at the speech's outset and the other at the 4-minute and 23-second mark. These pauses aimed to command audiences' discipline and maintain control over his discourse. In terms of vocal qualifiers, the intensity of Trump's voice is over loud, the volume is high, the pitch is over high, and the intonation is predominantly falling. This falling intonation was employed to match Trump's use of declarative sentences and also underscore his self-confidence. Due to the shortness of Trump's speech, 7 minutes, no vocal segregates were detected.

8.5.2 Application 2

To make sure of the validity of the suggested model, it has been applied to an extended political speech. The selected speech is delivered by Alsisi, on 13 February, 2023, and is delivered in Arabic, a language distinct from English. The speech's duration is 49 minutes. Within this speech, President Alsisi, who has been Egypt's president since 2014, attempts to persuade a diverse audience. He addresses other presidents present at the meeting, seeking their economic assistance. He also communicates with the Egyptian populace to inform them of his accomplishments and garner their support. Additionally, he appeals to both Arabic and foreign investors, aiming to attract investments that would boost

Egypt's economy. Applying the suggested model to Alsisi's speech unveiled the presence of most of the forms of metadiscourse, whether they were expressed verbally or non-verbally.

The verbal features

Alsisi made use of all forms of metalinguistic metadiscourse, as outlined in the provided examples.

Ex2.1. أنا مش بتكلم مع متخصصين ولا مثقفين (*I do not talk to specialists or intellectuals*) **is clarified to** → أنا بتكلم مع الناس مع الشعوب (*I am talking to all to public and to peoples*)

Ex 2.2. مش حكاية مصر (*Tell Egypt's story*) **is repaired to** → مش حكاية مصر حكاية حكومة مصر (*It is not the story of Egypt, it is the story of the Egyptian government*)

Ex 2.3. التحديات كلها علي خط واحد (*All of the challenges are on one line*) **is reformulated to** → معنديش أولويات (*There is no priority*)

Ex2.4. ٢٠١١ هو تاريخ حاسم وكاشف في تاريخ مصر (*2011 is a decisive and revealing date in the Egyptian history*) **is commented on as** → كادت (*Egypt came close to being lost, as many other countries in the region have faced*).

Ex.2.5 different terminologies has been manipulated within Alsisi's speech like تحديات متوازية (*parallel challenges*) **is explained to** → لا توجد (*there is no priorities and all of the challenges should be dealt with in the same time*).

As for the discourse organization, Alsisi employed "topicalizers" when introducing new topics, as seen in examples 2.6

and 2.7. He also utilized "delimiting a topic" to draw attention to specific points, exemplified as in example 2.8. After elaborating on each topic, Alsisi concluded with "ending the topic" to highlight his viewpoint, as demonstrated in example 2.9. To reinforce previously discussed points, Alsisi incorporated "review" or "reminders" in order to remind the audiences of what is said before. For instance, he initially discussed the 2011 chaos and subsequently refreshed the audience's memory, as seen in example 2.10. Additionally, Alsisi employed "exemplifying" to make his perspective explicit as in examples 2.11 and 2.12. Finally, he provided a "preview" towards the speech's end, referring to the Egyptian government's efforts to safeguard the nation, distinguishing it from other unstable countries in the region. Alsisi made his "illocutions explicit" in different occasions as in examples 2.13 and 2.14 to reflect his openness.

Ex 2.6 هاتكلم عن التجربة المصرية (*I will talk about the Egyptian experience*)

Ex 2.7 ودلوقتي هاتكلم على مشكلة الكهرباء. (*Now, I will talk about the problem of electricity*)

Ex 2.8 لما أتكلم عن ٢٠١١ مش هاتكلم عن كل حاجة هاتكلم بس عن حالة الفوضى والتشردم. (*When I talk about 2011, I won't talk about everything, I'll only talk about the state of chaos and division that the Egyptians experienced*)

Ex 2.9 الدولة لما بتقع يا فيصل مش بترجع ثاني. (*When the country falls down, Faisal, it does not come back again*).

Ex 2.10 الحالة اللي حصلت في ٢٠١١ كانت حالة صعبة. (*what happened in 2011 was very difficult*).

Ex 2.11. على سبيل المثال (*for example*)

Ex 2.12. أنا هاذكر مثال (*I'll mention n example*)

Ex 2.13. أحب أشكر (*I want to thank*)

Ex 2.14. أحب أحبي الشيخ بن زايد (*I want to greet Sheikh Bin Zayed*)

Concerning interactional metadiscourse, Alsisi utilized various aspects of it in his speech. On multiple occasions, he inquired if his interviewer comprehended his statements by asking, "Do you understand me?" He also endeavored to maintain control over the conversation by requesting the interviewer's attention, saying, "Listen to me, Faisal!" Alsisi expressed his perspective through "attitude markers", seen in instances like 2.15 and 2.16. He employed self-mentioning, using "I," around 20 times to highlight his efforts in safeguarding the nation. Similarly, he made use of different engagement markers. Rhetorical questions appeared roughly 15 times, as in examples 2.17 and 2.18. Furthermore, he utilized the inclusive "we" 32 times. Of these, 6 encompassed attending presidents, 20 referred to the Egyptian government, and 6 embraced the entire Egyptian populace. "Personalizers" were also frequent; sometimes, he directed his speech towards the interviewer, using names like "Faisal," and other times, he referred to specific Egyptian officials such as the minister of electricity. This personalization was achieved through phrases like those found in example 2.19. Alsisi employed diverse "certainty markers", as shown in examples 2.20 and 2.21. Additionally, he utilized "attributors" frequently, similar to example 2.22.

Ex 2.15. لازم مشكلة الكهرباء تتحل (*The crisis of electricity has to be solved*)

Ex 2.16. لا بد من التواصل المستمر وايضاح الصورة (*continuous communication and providing clear explanations are urgently required*)

Ex 2.17. يعني ايه خط واحد؟ (*What does one-line mean?*)

Ex 2.18. ده معناه ايه؟ (*What does it mean?*)

Ex 2.19. كل واحد فينا (*every one of us*).

Ex 2.20. بلا شك (*undoubtedly*)

Ex 2.21. بكل تأكيد (*certainly*)

Ex 2.22. بيتيهيا لهم (*as they imagine*)

Non-verbal metadiscourse

In his speech, Alsisi heavily relied on gestures to bolster his perspective and sway the audience. He used iconic gestures as appeared when he brought his palms together as an alerting sign and combined it with the use of the inclusive pronoun "we" in the statement " لازم ناخذ بالننا كويس *we have to pay attention*". Additionally, Alsisi employed his index as an icon for "No", when saying " *أنا لا* ، *لا يمكن أقبيل ده* No, I cannot accept *that*". Deictic gestures become evident as Alsisi employs his index finger to signify place, exemplified by his statement " لازم أحفظ تراب البلد *I have to safeguard the country' sand*." Additionally, he utilizes his index finger to refer back in time, using the phrase " *في الوقت ده* *at that time*," indicating a previous point in time. Metaphoric gestures were also a significant part of Alsisi's speech. He used his index finger to symbolize essential actions for Egypt's preservation and clenched hands to represent determination and challenge. He used two open palms above each other as a metaphor for the accumulation of problems. He also used open hands with slight separation to symbolize resource limitations for

Egypt's development. Gestures like open hands raised upwards conveyed openness and acceptance. Throughout the speech, Alsihi incorporated beats to capture attention and emphasize his viewpoints.

During his speech, Alsihi maintained a normal gaze for the most part to prevent seeming intimidating or causing discomfort to his listeners. On occasion, he employed looking while talking, possibly to gain his audience's backing. On rare instances, he used zero gazing, allowing the interviewer to comment on his speech as shown in figure 8.

Figure 8

Alsihi' zero gaze at the minute 4:40



Facial expressions employed in Alsihi's speech are deliberate, which aimed to convey particular emotions. Throughout the speech, Alsihi maintained a smile to exude confidence and approachability, while occasionally using a frown to emphasize his decisiveness and commitment to Egypt's security. In terms of body posture, Alsihi adopted an open posture stance to mirror his honesty, transparency, and self-assuredness.

As for vocal characterizers, Alsihi incorporated laughter in his speech to create a friendly atmosphere and alleviate any tension, as depicted in figure 9. He also employed brief moments of silence for

two purposes. Firstly, to maintain the audience's attentiveness, and secondly, to allow them time to process his speech and show their support.

Figure 9

Alsisi's laughter in the minute 17:34



In terms of vocal qualifiers, Alsisi maintained a moderate vocal intensity, a medium volume, and an intermediate pitch throughout his speech. This conveyed his approachability, confidence, and sincerity, while also avoiding any sense of intimidation towards the audience. Furthermore, he used a falling intonation when presenting factual information or posing wh-questions (e.g., "هل كان في تحديات أخرى؟" *Are there any other challenges?*) to demonstrate his certainty. In rhetorical yes/no questions (e.g., "صح ولا غلط؟" *True or false?*), he employed a rising intonation, indicating his anticipation of the audience's response. Alsisi utilized rising-falling intonation when listing items (e.g., "one, two, three"), using rising intonation at the beginning to convey incompleteness and transitioning to falling intonation at the end to signal the conclusion of his presentation. Concerning vocal segregation, Alsisi employed it only on rare occasions, like at the minute 12:22, when he utilized the interjection "/tu?tu?/" to represent the meaning "No."

8.5.3 Discussion

Regarding Trump's speech, he utilized various forms of metadiscourse in order to gain support from his audiences. He employed four of the metalinguistic metadiscourse. For instance, when he criticized the current American president, accusing him of corruption and dishonesty, he employed "repairing" by revising his utterances to express respect for the presidency office despite the accusation. Similarly, he frequently used "reformulating" to clarify his statements and enhance their impact. Trump also employed "commenting", such as when he commented on the significance of the 2024 election by stating that this date is crucial for restoring America's freedom. Additionally, he adopted "clarifying" to eliminate any potential ambiguity or misunderstanding.

In Alsisi's speech, he effectively incorporated different forms of metadiscourse to secure the backing of all his listeners. He utilized metalinguistic metadiscourse, employing methods like "clarifying" to eliminate any vagueness in his speech and to address any potential difficulties within his message, "commenting" to bolster the impact of his messages, "reformulating" to make his points clear and more profound, "repairing" to prevent anticipated confusion, and "managing terminology" to ensure that his audiences understood the terminology he used in his speech.

These findings are consistent with those of Ädel (2010), who discovered comparable results when examining both spoken and written academic discourse. Ädel noted that "repairing" is employed for rectifying or negating a prior contribution, "reformulating" is utilized to propose an alternative term or expression to enhance the

expansion, "commenting" pertains to linguistic form, word choice, and/or meaning, "clarifying" serves to prevent misinterpretation, and "managing terminology" is employed to provide definitions and designate terms or labels for discussed phenomena.

Both Trump and Alsisi employed "topicalizers" to introduce or add new topics, making transition between topics seamless and easy to follow. This aligns with the findings from Crismore et al. (1993), Dafuz–Milne (2008), and Bernad–Mechó (2018). However, Alsisi utilized "delimiting a topic" more frequently than Trump, as his longer speech covered various topics, allowing him to focus the audiences' attention on specific points and avoid distractions. Both Alsisi and Trump used "reviewing" to recall previous events and "previewing" to emphasize forthcoming messages in their speeches. They also utilized "sequencers" to organize their ideas coherently, aiding audiences' comprehension. Notably, Alsisi employed "illocutionary markers" more than Trump, explicitly conveying his intentions, whereas Trump relied on implicit expression, particularly when directing criticism toward the incumbent American President before his official campaign launch. Additionally, Alsisi employed "exemplifying" in his speech in different occasions to clarify his viewpoints. These findings are consistent with the results of Ädel (2010), highlighting the benefits of "delimiting a topic," "reviewing", "previewing" and "exemplifying". The effectiveness of "sequencers" and "illocutionary markers" in discourse, are also supported by the studies of Crismore et al. (1993) and Dafuz–Milne (2008).

As for interactional metadiscourse, both Trump and Alsisi made sure that there is no problem with the communication channel to "manage comprehension". Alsisi explicitly "managed audience's discipline" by periodically requesting the interviewer's attention to convey his message effectively. Both Trump and Alsisi used "attitude markers" to express their attitudes on the presented topics. Similarly, they both employed "self-mention", however, Trump used it extensively, mentioning himself 25 times in his 7-minute speech to promote his presidential candidacy. In terms of "engagement markers", Trump relied solely on *inclusive expressions* and *personalization*, while Alsisi used *inclusive expressions*, *personalization*, and *rhetorical questions* to enhance audience engagement. Notable, Trump avoided using "hedges" in his speech, whereas Alsisi employed them many times. Additionally, both employed "attributors" and "certainty markers". These findings align with Alyousef's (2015) study, which observed extensive use of interactional metadiscourse in multimodal contexts. They also corroborate the conclusions of Hyland (2005) and Bernad-Mechó (2018), who noted that interactional metadiscourse fosters audience involvement in both written and spoken discourse.

Regarding nonverbal kinesics/non-vocal metadiscourse, Trump used "gestures" sparingly, while Alsisi incorporated a variety of "gestures" throughout his speech to enhance the delivery of his message and to engage the audiences. This conclusion is consistent with the findings of Gullberg et al. (2008) and McNeil (2005). In terms of "gaze", Trump maintained continuous gazing and looking

while speaking, whereas Alsisi employed normal gaze patterns. Continuous gazing, as suggested by Argyle et al. (1974), conveyed Trump's assertiveness and potentially created a challenging vibe. In contrast, Alsisi's normal gaze, as indicated by Bailly et al. (2010), fostered a sense of intimacy with his audience.

Both Trump and Alsisi displayed "deliberate emotional facial expressions" and maintained "open body postures". According to Frank (2001) and Pease and Pease (2004), these actions were employed to convey confidence, transparency, and openness. Concerning paralinguistic features, Trump primarily used shouting, while Alsisi occasionally incorporated laughter. Shouting, as per Bancroft (1995), might convey dominance and authority, whereas laughter added a friendly atmosphere.

Both Trump and Alsisi used moments of silence strategically to maintain control over their audiences and continue their speeches. Trump utilized a high volume, high pitch, loud intensity, and falling intonation to project confidence, power, and dominance. In contrast, Alsisi employed moderate intensity, moderate volume, and intermediate pitch to maintain a friendly atmosphere with the audience and express his willingness to engage in discussion. Alsisi also employed various intonation patterns to prevent his speech from becoming monotonous and tailor his delivery to each topic. Notably, Alsisi used vocal segregates to create a sense of closeness and informality with his audience.

9 Conclusion

This research sought to answer two main questions. The first one tried to find out a model that could be used to analyze metadiscourse in spoken political discourse from a multimodal viewpoint. To address this question, various models of metadiscourse were explored in the literature. The research concluded that multimodal metadiscourse can be categorized into two main types: verbal and non-verbal. Verbal metadiscourse can be further subdivided into interactive and interactional categories. Interactive metadiscourse can be delineated into metalinguistic features and discourse organization features, while interactional metadiscourse encompasses managing comprehension, managing audiences' discipline, attitude markers, attributors, self-mention, hedges, engagement markers and certainty markers.

Non-verbal metadiscourse is divided into non-vocals or kinesics and vocalic or paralinguistic markers. Non-vocals or kinesics is further divided into gesture, gazes, body posture, and facial expressions, whereas vocalic or paralinguistic markers are further divided into vocal characterizers, vocal qualifiers and vocal segregates. Once these various forms of metadiscourse were defined, they were sent to experts in linguistics via e-mail to gather their perspectives and receive recommendations for potential adjustments. All feedback provided by these experts was carefully considered during the development of the final model.

The second research question aimed to investigate the application of the model to analyze political speeches. To address this, two political speeches were chosen as the research samples.

The first speech was delivered by Donald Trump during his preliminary candidacy campaign for the 2024 USA election. The second was an Arabic political speech by the Egyptian President Abdelfattah Alsisi during the international summit for governments held in the UAE on February 13, 2023. The second research question was divided into two sub-questions. The first sub-question aimed to use the proposed model to analyze the verbal metadiscourse used in both speeches. It was found that both politicians employed various forms of verbal metadiscourse, including interactive and interactional features. The second sub-question explored the model's ability to detect non-verbal metadiscourse features. The results indicated that both politicians in the selected samples utilized different non-verbal forms as part of their metadiscourse. It was evident that the different forms of metadiscourse employed in both speeches represent the speakers' efforts to convey their perspectives and engage the audiences effectively. Therefore, it was concluded that this model is applicable to analyzing spoken political discourse.

Limitation

The current research findings are relevant to spoken discourse, specifically in the context of political speech. However, their suitability for other forms of spoken communication, such as everyday conversations, films, academic discussions, etc., has not been explored yet.

Further research

Further research is required to investigate the validity of the suggested model in other forms of spoken discourses. Moreover, more extensive research can delve into how well the suggested model can be used for quantitative analysis in larger corpuses of spoken political discourse.

References

- Abusalim, N., Zidouni, S., Alghaza, S., Rababah, G., & Rayyan, M. (2022). Textual and interpersonal metadiscourse markers in political discourse: A case study. *Cognet Arts & Humanities*, 9. <http://doi.org/10.1080/23311983.2022.2124683>.
- Ädel, A. (2006). *Metadiscourse in L1 and L2 English*. John Benjamins.
- Ädel, A. (2010). Just to give you kind of a map of where we are going: A taxonomy of metadiscourse in spoken and written academic English. *Nordic Journal of English Studies*, 9(2), 69-97.
- Ädel, A., & Mauranen, A. (2010). Metadiscourse: Diverse and divided perspectives. *Nordic Journal of English Studies*, 9(2), 1-11.
- Alyousef, H. S. (2015). An investigation of metadiscourse features in international postgraduate business students' texts: The use of interactive and interactional markers in tertiary multimodal finance texts. *SAGE Open*, October-December, 1-10. <http://doi.org/10.1177/2158244015610796>

- Alzarieni, M. M., Zainudin, I. S., Awal, N. M., & Sulaiman, M. Z. (2019). Interactional metadiscourse markers in the abstract sections of Arabic patents. *Arab World English Journal (AWEJ)*, 10(2), 379–393.
<https://dx.doi.org/10.24093/awej/vol10no2.29>
- Argyle, M., Ingham, R., Alkema, F. & mccallin, M. (1981). The different functions of gaze. In A. Kendon (Ed.), *Nonverbal communication, interaction, and gesture: Selections from SEMIOTICA* (pp. 283–296). De Gruyter Mouton.
<https://doi.org/10.1515/9783110880021.283>.
- Argyle, M., Lefebvre, L., & Cook, M. (1974). The meaning of five patterns of gaze. *Eur. J. SOC. Psychol*, 4 (2), 125–136.
- Bailly, G., Raidt, S., & Elisei, F. (2010). Gaze, conversational agents and face-to-face communication. *Speech Communication*, 52, 598–612.
- Bancroft, W. J. (1995). *Research in nonverbal communication and its relationship to pedagogy and suggestopedia*. Eric. Chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://files.eric.ed.gov/fulltext/ED384243.pdf
- Bernad-Mechó, E. (2018). *A multimodal approach to metadiscourse as an organizational tool in lectures* [PHD Dissertation]. Jaume I University, Spain.
- Bull, P. E. (1987). *Posture and gesture*. Pergamon Press.

- Cafaro, A., Pelachaud, C., & Stacy Marsella, S. (2019). Nonverbal behavior in multimodal performances. In S. Oviatt, B. Schuller, & P. R. Cohen, *The handbook of multimodal-multisensor interfaces: language processing, software, commercialization, and emerging directions* (pp. 219–262). Association for Computing Machinery and Morgan & Claypool. <https://doi.org/10.1145/3233795.3233803>.
- Carrió–Pastor, M. L. (2022). Teaching Multimodal Metadiscourse in English for Academic Purposes. *Lidil*, 65, 1– 19. DOI: 10.4000/lidil.10575.
- Chen, L., & Li, C. (2023). Interactional metadiscourse in news commentaries: a corpus-based study of China Daily and the New York Times. *Journal of Pragmatics*, 212, 29–40.
- Collier, E. S. (2016). Body posture. In T.K. Shackelford, & V.A. Weekes–Shackelford (Eds.), *Encyclopedia of evolutionary psychological science*. Springer International Publishing. DOI: 10.1007/978-3-319-16999-6_1417-1.
- Crismore, A., Markkanen, R., & Steffensen, M. S. (1993). Metadiscourse in persuasive writing a study of texts written by American and Finnish university students. *Written Communication*, 10(1), 39–71.
- Dafouz–Milne, E. (2008). The pragmatic role of textual and interpersonal metadiscourse markers in the construction and attainment of persuasion: A cross-linguistic study of newspaper discourse. *Journal of Pragmatics*, 40, 95–113. <http://doi.org/10.1016/j.pragma.2007.10.003>.

- Dash, B. B. (2022). Significance of nonverbal communication and paralinguistic features in communication: A critical analysis. *International Journal for Innovative Research in Multidisciplinary Field*, 8(4), 172-179. <http://doi.org/10.2015/IJIRMF/202204029>.
- Ebrahimi, S. J. (2018). The role of metadiscourse markers in comprehending texts of reading comprehension books published in Iran and Oxford university press. *International Journal of Applied Linguistics & English Literature*, 7(3), 90-96. <http://dx.doi.org/10.7575/aiac.ijalel.v.7n.3p.90>.
- Ethier, N. A. (2010). Paralinguistic and nonverbal behaviour in social interactions: a lens model perspective [PHD Dissertation]. University of Waterloo, Canada.
- Farahani, M. V., & Kazemian, R. (2021). Speaker-audience interaction in spoken political discourse: a contrastive parallel corpus-based study of english-persian translation of metadiscourse features in TED Talks. *Corpus Pragmatics*, 5, 271-298. <https://doi.org/10.1007/s41701-021-00099-z>.
- Frank, M. G. (2001). Facial expressions. *International Encyclopedia of the Social & Behavioral Sciences*, 5230-5234. <https://doi.org/10.1016/B0-08-043076-7/01713-7>.
- Gualberto, C., & Kress, G. (2019). Social Semiotics. In R. Hobbs, & P. Mihailidis (Eds.), *The International Encyclopedia of Media Literacy*, (1-9). Johnwiley & Sons, Inc.
- Gullberg, M., de Bot, K., & Volterra, V. (2008). Gestures and some key issues in the study of language development. *Gesture*, 8(2), 149-179. <http://doi.org/10.1075/gest.8.2.03gul>.

- Guyer, J. J., Brinol, P., Thomas I. Vaughan-Johnston, T. I., Fabrigar, L. R., Lorena Moreno, L., & Petty, R. E. (2021). Paralinguistic features communicated through voice can affect appraisals of confidence and evaluative judgments. *Journal of Nonverbal Behavior*.
<https://doi.org/10.1007/s10919-021-00374-2>.
- Hager, J. C., & Ekman, P. (2005). The asymmetry of facial action is inconsistent with models of hemispheric specialization. In P. Ekman & Rosenberg, E. L., *What the face reveals: basic applied studies of spontaneous expression using facial action coding system* (pp. 40–58). Oxford University Press.
- Halliday, M. A. K. (1978). *Language as social semiotic: the social interpretation of language and meaning*. Edward Arnold.
- Halliday, M. A. K. (1985). *An introduction to functional grammar* (1st ed.). Edward Arnold.
- Hessels, R. S. (2020). How does gaze to faces support face-to-face interaction? A review and perspective. *Psychonomic Bulletin and Review*, 27, 856–881.
<https://doi.org/10.3758/s13423-020-01715-w>.
- Hodge, R., & Kress, G. (1988). *Social Semiotics*. Cambridge.
- Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. Bloomsbury Academic.
- Hyland, K., & Tse, P. (2004). Metadiscourse in Academic writing: A reappraisal. *Applied Linguistics*, 25 (2), 156–177.
- Hyland, K., Wang, W., & Jiang, F. (2022). Metadiscourse across languages and genres: An overview. *Lingua*, 265(103205), 10–1016.

- Ilie, C. (2003). Discourse and metadiscourse in parliamentary debates. *Journal of Language and Politics*, 2(1), 71–92.
- Isalambo, K. B., & Kenneth, O. (2020). interactive multimodal metadiscourse in covid 19 campaign posters by ministry of health, Kenya. *International Journal of English Literature and Social Sciences*, 5(5), 1595–1601. <https://dx.doi.org/10.22161/ijels.55.38>.
- Jakobson, R. (1985). Metalanguage as a linguistic problem. In Volume VII *Contributions to Comparative Mythology* (pp. 113–121). De Gruyter Mouton. <https://doi.org/10.1515/9783110855463.113>.
- Jewitt, C. & Henriksen, B. (2016). Social semiotic multimodality. In N. Klug & H. Stöckl (Eds.), *Handbuch Sprache im multimodalen Kontext*, (pp. 145–164). De Gruyter. <https://doi.org/10.1515/9783110296099-007>.
- Kenzhekanova, K. K. (2015). Linguistic features of political discourse. *Mediterranean Journal of Social Sciences*, 6(62), 192–199. <https://doi.org/10.5901/mjss.2015.v6n6s2p192>.
- Kraus, N., & Slater, J. (2016). Beyond words: How humans communicate through sound. *Annu. Rev. Psychol.*, 67, 83–103. <https://doi.org/10.1146/annurev-psych-122414-033318>.
- Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.
- Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design*, (2nd edition). Routledge
- Lautamatti, L. (1978). Observations on the development of the topic in simplified discourse. In V. Kohonen & N. E. Enkvist

- (Eds.), *Text linguistics, cognitive learning, and language teaching* (pp. 71–104). University of Turku Publications.
- Lund, K. (2007). The importance of gaze and gesture in interactive multimodal explanation. *International Journal of Language Resources and Evaluation*, 41(3–4), 289–303.
- Lyons, A. (2016) Multimodality. In Z. Hua (Ed.), *Research methods in intercultural communication: A practical guide* (pp. 268–280). Wiley–Blackwell.
- Martin, R., & Zappavigna, M. (2019). Embodied meaning: A systemic functional perspective on paralanguage. *Functional Linguistics*, 6(1), 1–33. Doi: <https://doi.org/10.1186/s40554-018-0065-9>
- Mauranen, A. (1993). *Cultural differences in academic rhetoric: A textlinguistic study*. Peter Lang.
- Mcneil, D. (2005). *Gesture and thought*. Chicago University Press.
- Mittelberg, I. & Evola, V. (2014). Iconic and representational gestures. In C. Müller, A. J. Cienki, E. Fricke, S. H. Ladewig, D. McNeill, & S. Teßendorf (Eds.), *Body – language – communication: An international handbook on multimodality in human interaction* (pp.1732–1746). De Gruyter.
- Pease, A. & Pease, B. (2004). *The definitive book of body language how to read others' thoughts by their gestures*. Pease International.
- Poggi, I. (2008). Iconicity in different types of gestures. *Gesture*, 8(1), 45–61. <https://doi.org/10.1075/gest.8.1.05pog>.

- Qin, W., & Uccelli, P. (2019). Metadiscourse: Variation across communicative contexts. *Journal of Pragmatics*, 139, 22–39. <https://doi.org/10.1016/j.pragma.2018.10.004>.
- Querol-Julián, M., & Fortanet-Góme, I. (2012). Multimodal evaluation in academic discussion sessions: How do presenters act and react? *English for Specific Purposes*, 31, 271–283.
- Querol-Julián, M., & Fortanet-Góme, I. (2012). Multimodal evaluation in academic discussion sessions: How do presenters act and react? *English for Specific Purposes*, 31, 271–283.
- Ribeiro, L. A., Carvalho, D. M., Araújo, A. D. (2022). Multimodal metadiscourse: Analysis of glosses of the definitions and examples of an english dictionary. *Entrepalavras, Fortaleza*, 11, 57–72. <https://doi.org/10.22168/2237-6321-11esp2124>.
- Rossiter, C. M. (1974) Instruction in metacommunication. *Central States Speech Journal*, 25(1), 36–42. <https://doi.org/10.1080/10510977409367766>
- Schiffrin, D. (1980). Meta-Talk: Organizational and evaluative brackets in discourse. *Sociological Inquiry: Language and Social Interaction*, 50, 199–236.
- Sindoni, M. G. (2013). *Spoken and written discourse in online interactions: A multimodal approach*. Routledge.
- Tian, Y., Kanade, T., & Cohn, J. F. (2005). Facial expression analysis. In S. Z. Li, & A. K. Jain (Eds.), *Handbook of face recognition* (pp.247–276). Springer.

- Toumi, N. (2009). A model for the investigation of reflexive metadiscourse in research articles. *Language Studies Working Papers*, 1, 64-73.
- Trotta, D., & Guarasci, R. (2021). How are gestures used by politicians? A multimodal co-gesture analysis. *Computational Dialogue Modelling: The Role of Pragmatics and Common Ground in Interaction*, 7(1,2), 45-66. Doi: <https://doi.org/10.4000/ijcol.827>
- Van Leeuwen, T. (2005): Introducing social semiotics. London
- Vande Kopple, W. J. (1985). Some exploratory discourse on metadiscourse. *College. Composition and Communication*, 36(1). 82-93.
- Vande Kopple, W. J. (2012). The importance of studying metadiscourse. *Applied Research in English*, 1(2), 37-44.
- Williams, J. M. (1981). *Style toward clarity and grace*. Scott, Foresman and Company.
- Windsor, L., (2017) Language, nonverbal, and audiovisual cues: Multimodal approaches to understanding political behavior [White Paper]. National Academy of Sciences, Washington, DC.
https://sites.nationalacademies.org/cs/groups/dbassesite/documents/webpage/dbasse_179888.pdf.

Appendix 1

Dear Professor,

I am Reham Khalifa, an associate professor within the department of English Language, Faculty of Arts, Damietta University. My current endeavor involves the development of a comprehensive model designed for the analysis of metadiscourse within the context of spoken political discourse. I approach this research from a multimodal perspective, aiming to encompass both verbal and non-verbal aspects.

The proposed model is structured into two distinct sections. The initial segment concentrates on the examination of the verbal characteristics inherent to metadiscourse. In contrast, the second portion delves into the investigation of non-verbal facets within metadiscourse. I have attached the proposed model to this letter for your review and consideration. I kindly request that you take the time to peruse the attached model. Your expertise and insights are highly valued. I would greatly appreciate it if you could provide feedback regarding the elements that align well with our objectives, any necessary modifications, and any elements that may warrant deletion.

I would like to extend my sincere gratitude for your cooperation in this matter. Your guidance and input will undoubtedly contribute to the refinement and success of this research initiative.

Thank you once again for your time and attention.

Sincerely, Dr. Reham Khalifa

Verbal metadiscourse

1. Interactive		Suitable	Delete	Modify
Metalinguistic	Repairing			
	Commenting			
	Managing terminology			
	Reformulating			
	clarifying			
Discourse organization	Introducing a topic			
	Adding a topic			
	Delimiting a topic			
	Concluding a topic			
	Reviewing			
	Previewing			
	contextualizing			
	Sequencers			
	Endophoric markers			
	evidentials			
	Code glossing			
	Illocutionary markers			
Tropicalizers				
2. Interactional				

Managing comprehension			
Managing audiences' discipline			
Anticipating audiences' response			
Attitude markers			
Self-mention			
Engagement markers			
Hedges			
Boosters			
Certainty markers			
Items that need to be added			

Non-verbal Metadiscourse**1. Non-vocals/Kinesics****Suitable****Delete****Modify**

Gestures

Iconic

Metaphoric

	Deictic				
	Beats				
Gazes	Zero gazing				
	Continuous gazing				
	Looking while talking				
	Looking while listening				
	Normal Gaze				
Facial expressions	spontaneous	emotional			
		non-emotional			
	Deliberate	emotional			
		non-emotional			
Body postures	Open				
	Closed				
2. Vocalics/paralinguistic features					
Vocal characterizers	Laughter				
	Yawning				
	Crying				
	Shouting				
	Whining				
Vocal qualifiers	Intensity	Over loud			
		Over soft			
	Volume	high			
		low			
	Pitch	Over high			
		Over low			
Intonation	rising				
	falling				

Vocal segregates	Voice interjections			
Items need to be added:				

Any further comments: