Multimodal argumentation:
association of words and images in the construction of the argumentative dimension of discourse

Amanda Heiderich Marchon
Universidade Federal do Espírito Santo, Vitória, ES, Brasil
ORCID: https://orcid.org/0000-0002-6576-949X
amandahch.letras@gmail.com

Welton Pereira e Silva
Universidade Federal Fluminense, Niterói, RJ, Brasil
ORCID: https://orcid.org/0000-0002-4693-3151
weltonp.silva@hotmail.com

Carlos Eduardo Nunes Garcia
CEFET/MG, Belo Horizonte, MG, Brasil
ORCID: https://orcid.org/0000-0002-9112-7490
nunes_carlosedu@hotmail.com

ABSTRACT

Considering that discourse emerges from the multisemiotic complexity that characterizes language and, for this reason, involves different dimensions in the study of discourse materialized in texts, in this article we (re) formulated some reflections on the association between word and image as a strategy of argumentation. We assume that from the complete association between verbal and nonverbal elements, where one seems to replace the other because they have an equivalent informativeness, to the complete dissociation between these elements, where the word seems to negate the image and vice versa, there is a spectrum of association possibilities aimed at reasoning through the relationship between different modalities. To illustrate these considerations, the corpus analyzed in this article consists of seven journalistic texts of an informative nature published in the social media of media entities and on their websites. The adopted theoretical framework is mainly based on the semiolinguistic discourse theory (CHARAUDEAU, 2005; 2007; 2008), associated with the proposal of the analysis of multisemiotic texts by Kress and van Leewen (2006), as well as on previous studies in which we take as a starting point the methodological analysis of multimodal argumentation (MARCHON e GARCIA, 2020; SILVA e MARCHON, 2021). The results of the analysis point to the possibility of understanding multimodal argumentation arranged in a *continuum*, which can be applied to other research involving different discourse genres.

KEYWORDS: Argumentation; Multimodality; Journalistic discourse.

Argumentação multimodal: associação de palavras e imagens na construção da dimensão argumentativa do discurso

RESUMO

Considerando que o discurso emerge da complexidade multissemiótica que caracteriza a linguagem e, por essa razão, engloba diferentes dimensões na investigação de discursos materializados em textos, neste artigo, (re)formulamos algumas considerações no que concerne à associação entre palavra e imagem como estratégia de argumentação. Entendemos que, da associação completa entre elementos verbais e elementos não verbais, em que um parece substituir o outro por apresentarem informatividade equivalente, até a dissociação completa entre esses elementos, em que a palavra parece negar a imagem e vice-versa, há um espectro de possibilidades associativas que visam à argumentação por meio da relação entre modalidades diferentes. A fim de ilustramos essas considerações, o *corpus* analisado neste artigo é composto por sete textos jornalísticos de caráter informativo, publicados nas redes sociais das instâncias midiáticas e em seus *sites*. O referencial teórico adotado ancora-se, principalmente, na Teoria Semiolinguística do Discurso (CHARAUDEAU, 2005; 2007; 2008) associada à proposta de análise de textos multissemióticos de Kress e van Leewen (2006), bem como em trabalhos anteriores em que tomamos como ponto de partida a proposta de análise metodológica da argumentação multimodal (SILVA; MARCHON, 2021; MARCHON; GARCIA, 2020). Os resultados da análise apontam para a possibilidade de compreensão da argumentação multimodal disposta em um *continuum*, o que pode ser aplicado a outras pesquisas com gêneros discursivos distintos.

PALAVRAS-CHAVE: Argumentação; multimodalidade; discurso jornalístico.

1. Introduction

The proposal to analyze the multimodal strategies employed in argumentation is premised on the recognition that language manifests itself through inherently multisemiotic means. Facial expressions, gestures, and prosody in orality and graphic elements such as colors, font size, and type, in writing allow us to recognize that various semiotes operate alongside verbal material in our everyday interactions (KRESS; VAN LEEUWEN, 2006). Based on this understanding, Marchon e Garcia (2020) proposed the notion of multimodal argumentation, developed in Marchon e Silva (2021).

In previous studies inspired by Santaella (2012) and Barthes (1990), we described three strategies based in what we call multimodal argumentation, namely, (i) multimodal argumentation by focusing; (ii) multimodal argumentation by complementarity; and (iii) multimodal argumentation by divergence. Each of them attempts to represent in a watertight way how verbal and visual materials are interrelated depending on the intentions of the expressing subject.

Multimodal argumentation by focusing permeates texts in which the verbal and visual materials are equivalent in terms of informativeness, so that the content is reinforced and repeated between one semiotic and another. In other words, we observe that in texts that use this type of argumentation strategy, part of the images repeats some information conveyed by the verbal material and vice versa.

Multimodal argumentation by complementarity, as the name implies, presupposes that in some texts the verbal and the visual material are interdependent, since the construction of meaning in the text is possible only through the connection of the different semioses. In texts

whose argumentation is constructed through this type of strategy, we find that the information conveyed by the different semioses is different, although they are complementary.

Finally, multimodal argumentation by divergence initially creates a dissonance effect, in which the verbal and visual materials are in disharmony with each other, leaving it to the receiving instance to infer the reasons that led the producing instance to choose an image that seems to diverge from the verbal material. More precisely, this argumentative resource simulates an apparent incoherence between image and word, as if one semiosis in some way negates the information expressed by the other.

In this paper, however, we assume that the multitude of texts that employ multimodality as an argumentative strategy cannot be distributed among these three watertight categories. For this reason, we believe it is more coherent to propose a scale and gradient analysis that considers the word-image relationship under the features [+/- divergent] and [+/- focused]. In this way, multimodal argumentation strategies by focus and multimodal argumentation by divergence would represent the two poles of a *continuum*, with the former having the lowest degree of divergence and the highest degree of focus, in contrast to the latter, as we will discuss in this paper.

To set out our reasoning, in the first section we introduce the basic theoretical framework based on semiolinguistic discourse theory. This assumes that discourse emerges from the multidimensional complexity that characterizes language, and therefore seeks to capture different dimensions in the study of discourse materialized in texts. Next, we introduce the proposed *continuum* that relates words and images under the features [+/- divergent] and [+/- focused]. In the last section, we provide some considerations that summarize the analysis proposal presented in this article.

2. The semiolinguistic theory of discourse

The semiolinguistic theory of discourse assumes that every text is a unit of meaning constructed at micro- and macrotextual levels with the aim of achieving the listener's assent to the propositions presented by the speaker, and understands the text as a *speech act*, that is, an actional dialogical event between social and discursive subjects. Thus, the notion of text as discourse presupposes that it is the result of a strategic communication operation in a particular situation of social interaction. In this sense, according to Charaudeau (2008), semiolinguistics conceives language as something inseparable from its socio-historical context and, in a clearly pragmatic view, proposes that discourse emerges to fulfill certain intentions of interacting subjects and to achieve effects through its use.

Under this speaking perspective, language enables the subjects involved in communicative interaction to make the transition from a *signifying world* to a *world discursively signified* in a cooperative act, because, according to Charaudeau (2008, pp. 20-21), "the world is not given from the outset. It is given by the human strategy of signification." Figure 1 below illustrates this mechanics of meaning construction in and through discourse:

FIGURE 1. The double process of semiotizing the world
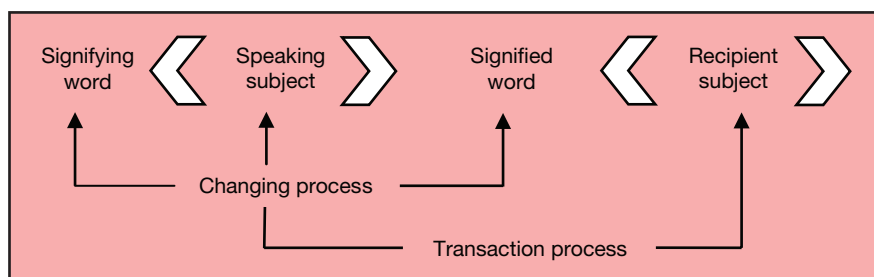
Source: Prepared by the authors; based on Charaudeau (2005, p. 14).

Since every act of communication is subject to a double process of meaning construction, based both in the production generated by the speaker and in the *interpretation* made by the listener, Charaudeau (2005) explains that the *signifying world* is the place where the raw event is located. It is in this reality, which does not yet have meaning, that the *transformation process* takes place, which consists in transforming the event from a raw state into an already interpreted state, that is, into a *world signified* in and through discourse. However, this construction of meaning is not autonomous, but is overdetermined by the *transactional process*, which forces the producing instance to construct its discourse according to the image it has of the receiving instance, which in turn reinterprets the utterances guided by its knowledge, values, and socio-discursive ideas. In this sense, Charaudeau (2008) argues that the speech act has implicit and explicit dimensions, that is, what is explicitly said does not correspond to everything that is communicated, leading the author to state that meaning depends on the communication situation:

> The purpose of the speech act (for both the uttering and the interpreting subject) should be sought not only in its verbal configuration, but in the play that a given subject develops between this and its implicit meaning. Such a play depends on the relation of the protagonists to each other and their relation to the discourse circumstances that bring them together (CHARAUDEAU, 2008, p. 24).

In this scenario, it should be noted that meanings do not exist outside the communication situation and are not only properties of one of the partners involved in the interaction, since the other must also be taken into account for the speech act to be efficient. For this reason, Charaudeau (2008, pp. 23-24) argues that every speech act is circumscribed by a multiple expectation that connects the speaker and the listener in an open, variable play of relations:

> Consequently, the knowledge that the protagonists of language construct about the various purposes of communicative exchange is not only related to the references or expectations that each of them experiences. This knowledge also depends on the knowledge that these communicating subjects assume between themselves and form the *filters that construct meaning*. (CHARAUDEAU, 2008, p. 31, Author's emphasis)

Thus, every speech-mediated act of interaction is the result of linguistic-discursive operations performed by subjective instances in well-defined situations. This model of communication proposed by semiolinguistics attempts to articulate discourse as a network of textual and social

events carried out through two complementary activities, namely: production and interpretation, activities determined by strategy spaces – the different types of discursive configurations available to the speaker – and restrictions – minimum conditions that must be met in order for communication to take place.

According to Charaudeau (2008), the activities of production and interpretation take place in a discursive *mise-en-scène* (staging) in which subjective entities participate: the communicating I (Ic), the speaking I (Is), the interpreting You (Youi), and the receiving You (YOUr). The first, the Ic, the subject responsible for production, has an intention and a speech project when it engages in an interaction with another partner, the YOUi. Both, social beings, act in the *external space* of the communication situation, the locus of the conditions of production and interpretation, action or doing of the discourse. These two subjects include the speaking subject (Is) and the receiving subject (YOUr), who belong to the internal space of the communication situation, the locus of the speaking produced by operations and/or linguistic-discursive maneuvers during the production of the discourse, as shown in Figure 2:

FIGURE 2. The speech act and the subjects
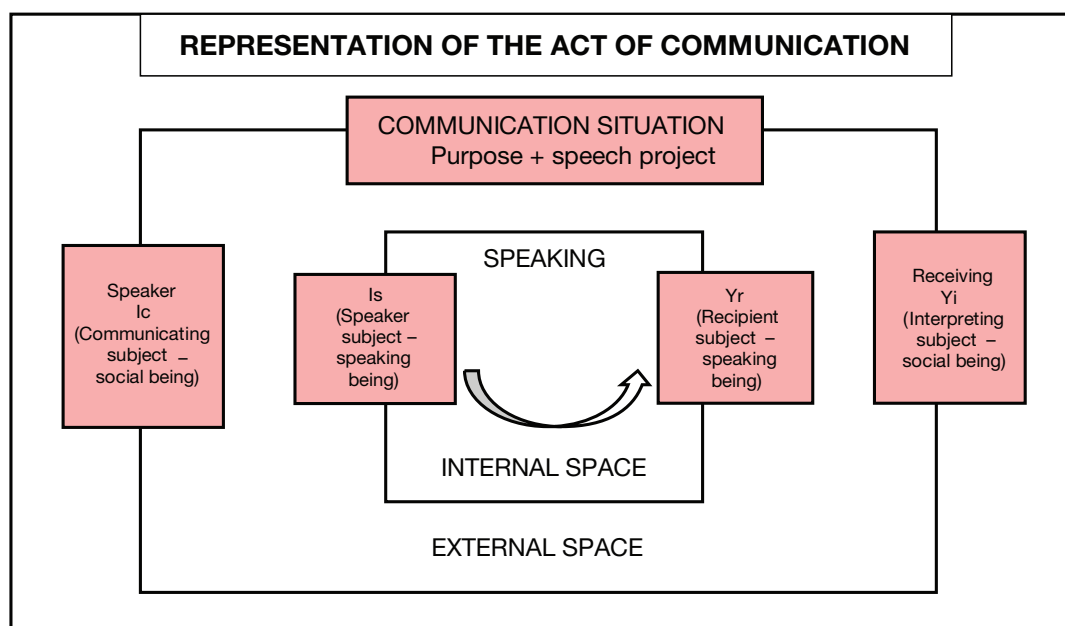
Source: Prepared by the authors; based on Charaudeau (2008, p. 52).

Figure 2 shows that the interlocutive process results in an unfolding of speaking places performed by four subjects, not just two, as in Jakobson's (1969)[1] theory. All communicative acts are thus considered as role-plays, in the same sense as in theater: the social subjects

---

[1] In proposing the classical functions of language (emotional, conative, poetic, meta-linguistic, phatic, and referential), Jakobson (1969) noted that the communicative act presupposes a *sender* and a *receiver*. We assume that the Russian linguist envisages in the linguistic circuit only the beings that for Charaudeau (2008) are synonymous with the social beings, namely: the communicating I and the interpreting You.

are the protagonists who play different speaking roles determined in and by the conditions of the communicative situation. Although the speech act is determined by the speech project generated by the speaker, it is an act of controlled freedom because it is subject to situational, discursive, and formal constraints imposed by the communication contract (CHARAUDEAU, 2007).

The communication contract presupposes that social beings belonging to the same social body establish a relationship of cooperation and complicity among themselves so that they can agree on speech representations. According to Charaudeau (2004, p. 132), the contract enables:

> the partners in a linguistic exchange to recognize each other with the identity features that define them as subjects of this act (identity), to identify the goal of the act that overdetermines them (purpose), to understand what constitutes the thematic goal of the exchange (purpose), and to take into account the relevance of the material constraints that determine this act (circumstances).

Specifically regarding the communication contract that determines the production of media discourse, Charaudeau (2007, p. 86) emphasizes that the communication situation involving the production of news and reports is inscribed in a double contract: one of information (based on the goal of making know) and another of capture (based on the goal of making feel). The first is based on a social logic that consists in informing the citizen; the second is guided by a commercial logic that consists in producing a consumer object to capture the masses and compete.

These goals of making known and making feel are oriented towards an argumentative intentionality. Apart from informing about something, the media employ several discursive strategies to influence the public's opinion. These maneuvers are implemented both by selecting angles on the facts in the news and by mobilizing readers' emotions – Charaudeau (2010) calls this incitement of feelings the pathemization effect. In the semiolinguistic theory of discourse, therefore, argumentation is understood in a broad sense, for it is inherent in all discourse, as the speaker attempts, directly or indirectly, to enforce on the listener the claims about the world presented in the communicative interaction. In this understanding that "argumentation is not limited to a sequence of sentences or propositions linked by logical connectors" (CHARAUDEAU, 2008, p. 203. Author's emphasis), semiolinguistics comes close to Ruth Amossy's argumentation theory of discourse. For this discourse analyst, argumentation can take place in an explicit way, presenting an argumentative view or orientation, or in an implicit way, understood as the argumentative dimension of discourse:

> The notion of argumentation here undergoes a maximum expansion, encompassing both the genres of discourse that explicitly defend or refute a thesis and those that are content to instruct ways of seeing and thinking. All speech aims at conveying a point of view, a way of responding to a situation, or a sense of an issue. The argumentation schemes in this case are more or less elaborated and more or less visible. Often the orientation of the discourse is shown in speech features (modalities, axiological forms, connectors, etc.) without the appearance of a formal argument. Sometimes this argument is only uttered in elliptical form and one has to resort to interdiscourse to reconstruct it. (AMOSSY, 2016, p. 174)

In this article, we start from this broader notion of argumentation and assume that every discourse has an argumentative dimension that can be analyzed. In the next section, we will see how the association of different semioses can produce argumentation – a strategy we call multimodal argumentation – even in genres that are not considered prototypically argumentative, such as news texts.

3. Multimodal argumentation: proposing a *continuum*

As we have seen, within the framework of semiolinguistic discourse theory, the transformation process in which the biosocial world is signified by verbal and visual material is assumed to be subordinate to the transactional process in which language is transformed into discourse by a subject's action. Thus, it is the intentionality of the speaker, always aimed at influencing the listener, that leads him to use language – and we add: images – in one way or another (CHARAUDEAU, 2007). Expanding the notion of text and understanding it not only as a complex of verbal forms, we can see that the choice of iconic elements, of images, depends not only on linguistic and lexical choices, but also on the principle of influence that governs communicative exchanges.

In this paper, we assume that the word-image relationship in terms of divergence and focus can be represented in scalar and gradient form, by a *continuum* whose poles are represented by the multimodal material [- divergent; + focused] and [+ divergent; - focused]. In this way, the different texts that want to use the multimodal argumentation strategy can position themselves at different points of the continuum that we propose here, in addition to the extremes.

Thus, we believe that the strategies of multimodal argumentation by focusing (MAF) and multimodal argumentation by divergence (MAD) are inversely proportional. In other words, the higher the focusing relation, the lower the divergence relation between the semiotes that make up a given text. This results in the first argumentation strategy being positioned at the left pole of the *continuum* proposed in this article, while the second is positioned at the right pole, as shown in Figure 3:

FIGURE 3. Relationship between words and images

MAF   MAD

[+ focused]   [- focused]
[- divergent]   [+ divergent]

Source: Prepared by the authors.

Unlike what was presented by Marchon e Silva (2021), multimodal argumentation by complementarity is no longer presented as a separate category. The change is necessary because verbal and visual semioses complement each other to some extent, either in the repetition of the content expressed by one of the semioses, which represents the features [+ focused; - divergent]

characteristic of multimodal argumentation by focusing, or in the apparent dissociation, which represents the features [- focused; + divergent] characteristic of multimodal argumentation by divergence, logically going through texts in which these features are not so obvious.

To illustrate our proposal, we begin by analyzing a journalistic text in which there appears to be an equivalent relationship between the verbal and visual material. The text of example (1), which can be found at the link[2] in the footnote at the bottom of this page, was published by the newspaper *O Globo* on October 10, 2021, and presents as verbal material the headline: "Videos: New dust storms hit at least four Brazilian states." As visual material, a photo of a huge reddish-brown dust cloud approaching a residential area is included. The image, which consists of a printout of one of the videos mentioned in the headline, shows that trees and buildings look small in front of the approaching cloud, which repeats the verbal material and reinforces the size of the announced storm.

This relationship leads the reader to understand the photo as evidence for the claim in the headline, an understanding that leads us to classify the multimodal argumentation as [+ focused] because part of the verbal statement is focused by the image, and as [- divergent] because there is no implied dissonance between verbal and visual material. Following this reasoning, Text 1 could be located near the left pole of the *continuum*.

Although the image reinforces what is in the caption, one can only infer from the context that there was a dust storm, but it is not known when or where. It is the geographical location "four Brazilian states" that captivates the interlocutor and makes him realize that it was something spatially and temporally proximate, since the means of communication presents, through verbal material, the date and location of the facts reported.

As for the complementary relationship between words and images, now understood as constitutive of any multimodal text, we note that new, albeit subtle, information is added to the text by the visuals, such as the height of the cloud of smoke and the reddish-brown color, which can evoke feelings of fear and anxiety in the reader. Comparing the height of the clouds in relation to the buildings, houses, trees, and cars present in the image not only illustrates the headline, but also engages the reader emotionally. According to Charaudeau (2007), in terms of the dual purpose of the media communication contract, the reader can be addressed in two ways: as an intellective target or as an affective target:

> The intellective target is assumed to be able to evaluate their interest in what is proposed to they, the credibility they give to the body that informs, her own ability to understand the message, i.e., to have access to it. An intellective target is one to which one attributes the ability to think. [...]
> An affective target, in contrast to the previous one, is one that is assumed not to evaluate anything rationally, but rather unconsciously through emotional reactions. Therefore, the media instance hypothesizes what is most appropriate to touch the target subject's affectivity. (CHARAUDEAU, 2007, p. 80-81)

---

2 Available at: <https://oglobo.globo.com/brasil/videos-novas-tempestades-de-poeira-sao-registradas-em-ao-menos-quatro-estados-brasileiros-25221233>.

As we can see, in this work, besides the notion of text constituted by the association between different semioses, we also have a comprehensive view of the phenomenon of argumentation, which is understood not only as logical argumentation that leads a premise to a conclusion, but as any discursive construction that aims at influencing the listener to a greater or lesser extent, in line with the postulates of Amossy (2016) and Charaudeau (2007).

Next, we will discuss another textual example in which the word-image relationship presents itself [+ focused; - divergent]. In the center of the image given as an example (2), on a light blue background, we see the current American president speaking firmly from a pulpit with his fists closed; to his right, the flag of the United States. Next to the photo, which can be seen in the link inserted in the footer, we read "Operation in Afghanistan was 'extraordinary success' (sic), says Biden"[3].

The text comes from the website of *Exame* magazine and was published on August 31, 2021. We can note that the verbal content resumes the U.S. president's speech about the decades-long occupation of the U.S. army in Afghanistan. Moreover, there is an associative relationship between the headline and the image showing Biden giving a speech from a pulpit. We note that the chosen image suggests that the president employed an energetic tone in his speech, given his posture with clenched fists and stern facial expression, which (re)enacts the very moment of the speech about the success of the American operation in Afghanistan. We also note that the U.S. president's gaze is directed at the reader of the text, who is strategically induced to participate in the discursive construction. According to the assumptions outlined by Kress and van Leeuwen (2006) in the *Grammar of Visual Design*, Biden, as the represented participant of the image, makes visual contact with the reader, an interactive participant, in order to engage in dialogue.

Thus, we assume that during the transformation from the signifying world to the world signified, the discourse production analyzed uses verbal and nonverbal elements that construct the image of a president speaking not only to his nation, represented by the United States flag on the left side of the image, but also to the world.

We note, however, that the same image has been used on other occasions by the same media outlet to illustrate other speeches by the U.S. president, as in a headline that announced, "Biden visits New York and New Jersey devastated by Hurricane Ida" (Example 3) and in another that read, "Biden insists wealthier pay 'fair share' of taxes" (Example 4). The texts, released on September 4 and September 14, 2021, respectively, can be viewed at the links provided in the footnotes[4].

Although the President may have addressed different topics in the speech in which he was photographed, it is certain that the image does not accurately reflect the three distinct moments in which he speaks about (i) the supposed success of the U.S. occupation in Afghanistan, (ii) the

---

3   Available at: https://exame.com/mundo/operacao-no-afeganistao-foi-sucesso-extrordinario-diz-biden/
4   Available at: <https://exame.com/mundo/biden-visitara-nova-york-e-nova-jersey-castigados-pelo-furacao-ida/>.
Available at: <https://exame.com/mundo/biden-insiste-que-mais-ricos-paguem-parcela-justa-de-imposto/>.

visit to New York and New Jersey after Hurricane Ida, and (iii) the payment of fairer taxes by the wealthiest.

Therefore, we note that although the image of Biden in a political speech illustrates each of the headlines, the image was certainly chosen strategically. By choosing the photo of the U.S. president giving a speech with clenched fists and a stern face, the media entity paints the picture of a strong and decisive political leader, and gives the impression that he really did speak in a concise manner on the topic mentioned in the headlines. It should be noted, however, that unlike Example 1 about the dust storm, Examples 2 through 4 focus on the secondary information in the headline, since the primary information is the content of Biden's statements, not Biden himself. Therefore, we propose to position these texts slightly further to the right on the *continuum* we have presented.

Next, we discuss a text from the newspaper *O Globo* published on the media instance's Facebook profile on May 3, 2019. In the verbal part of the text, located below the image, the caption "North Korea reduces food portions to 300 grams per day per person, says UN" and the caption, located above the photo, "some families consume protein only a few times a year" suggest a food shortage scenario. In contrast, the image shows North Korean leader Kim Jong-Um in the center smiling in a large wheat field, the basis for the production of many foods, accompanied by three uniformed military personnel who appear on the far right of the photo. You can see the image of example (5) at the link[5] provided in the footnote below.

The brightness of a sunny day in the photo combined with the leader's smile indicates the obvious discrepancy between the linguistic and pictorial material. The information that there is food rationing, which in itself is a negative fact, and that this reasoning limits the amount of food per person to only 300 grams per day, contradicts the image of a smiling leader in a field full of food. Thus, we can say that the word-image relationship is [- focused; + divergent], suggesting that the text can be positioned at the right end of the proposed *continuum*. The association of the semioses strategically aims at evoking the polemical effect of indignation in the interpreting subject.

Analyzing the arrangement of elements in the text according to the parameters proposed by Kress and van Leewen (2006), we found that the verbal statement at the beginning of the text indicates that the consumption of protein by North Koreans is not a tacit reality, in contrast to the dictatorial imposition of reduced food portions – information expressed at the end of the text, i.e., a positioning anchored in the tangibility of the situation described. Since "an image can be read even if it does not correlate with the verbal, thus confirming its status as language" (PASSOS DA SILVA, 2021, p. 75), we also highlight that in the construction of its discourse, the media instance presents the leader of North Korea in a central position, as the protagonist of the measure that reduces food consumption in the country. The presence of the military and the absence of civilians in the image reflect the country's ruling system that silences the population and disregards their basic needs.

---

[5]   Available at: <https://oglobo.globo.com/mundo/coreia-do-norte-reduz-porcoes-alimentares-300-gramas-ao-dia-por-pessoa-diz-onu-23639930>.

Thus, we conclude that in transforming the signifying world into a signified world, the speaker not only reports the fact that is being reported, but also evaluates it. By creating a discourse that only simulates neutrality effects, the speaker, through words and images, tries to make the interpreter see reality from the point of view of the producing instance, even though, theoretically, the informing text should perform the primary function of informing, with impartiality.

There are also texts that are not at either end of the *continuum*, but occupy a place in between. In example (6), which you can see in the link[6] below, the verbal and nonverbal elements are not dissonant, but they also do not directly reflect the same information.

In the text from *CNN Brazil* published on Facebook on June 22, 2022, the following verbal message is under the image: "MPF[7] recommends that the hospital perform a legal abortion on a rape victim." The nonverbal part shows a swing with two chairs, one of which is empty and on the other is probably a little girl, whose image we only partially see through the framing of the photo – a strategy to preserve the identity of the child. The entire image is in dark tones, especially the gray of the checkered floor and the black of the shadows of the empty chair and the girl playing in the other rocking chair.

When we compare the verbal and nonverbal material, we notice that there is no direct relationship – that is, [+ focused; - divergent] – between the headline and the image, because the only element that is repeated in both is the reference to at least one child. We also observe no dissonance effect in the word-image relationship that prevents the text from being positioned at the other end of the continuum - [- focused; + divergent].

By employing a strategy of polemical mobilization, the swing in the photo, functions as a symbol of childhood, which in a situation such as the one described in the headline, is at least cut off, as shown by the image of the swing with only part of the child visible, or completely distorted, as indicated by the empty swing on the right side of the picture. Furthermore, the color palette of the photograph indicates the trauma described in the headline, as the predominance of black and white shows a monochromatic scale that symbolically does not match childhood, which is generally represented by many colors. By using his own universe of beliefs and knowledge that mobilizes certain socio-discursive imaginaries related to childhood (swing; many colors) and the terror experienced by the child (monochromatic scale; empty swing), the subject interpreting the messages can be made to be moved by the text. According to Charaudeau (2007), the socio-discursive imaginaries correspond to the symbolic representations discursively disseminated in a given social group.

As for the analysis of the arrangement of the visual elements, according to Kress and van Leewen (2006), the elements that occupy the upper part are classified as ideal and represent the most abstract part of the information conveyed, while the elements in the lower part are classified as real and materialize the concreteness of the information expressed. Following this

---

6   Available at: <https://www.cnnbrasil.com.br/nacional/mpf-recomenda-que-hospital-realize-aborto-legal-em-crianca-viti-ma-de-estupro-2/>.
7   MPF= Public Prosecutor's Office.

reasoning, we understand that the swing with the child in the upper part of the image represents the idealization of a childhood characterized by fun, while the empty image of the swing in the lower part of the photo shows the reality of a childhood marked by violence, a stolen childhood. The association of the constituent semiotes of this text indicates that the listener instance is addressed as a target-affective according to Charaudeau (2007).

Therefore, based on the analyzed texts, we believe that we can schematically represent the positioning of the analyzed texts in the proposed continuum as follows:

FIGURE 4. Arrangement of the texts according to the relationship between words and images
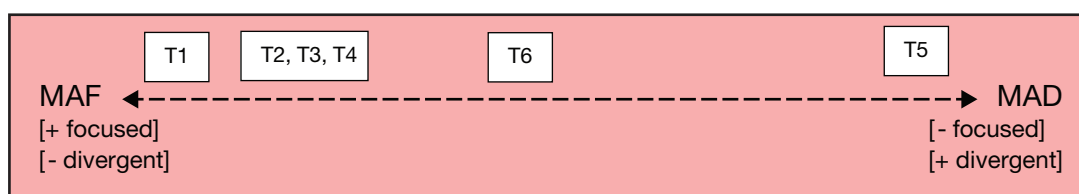
Source: Prepared by the authors

However, the presented scheme should be understood as a formal representation of the understanding about the different relationships between semiotes present in a multimodal text and not as a kind of ruler or tool for objective measurement.

4. Conclusions

In this paper, we use the assumptions of semiolinguistic discourse theory (CHARAUDEAU, 2007; 2008; 2010) associated with Kress and van Leewen's (2001) proposal of multisemiotic text analysis, similar to other studies we have developed. Under this prism of analysis, we understand that the purpose of discursive production by the speaker, both narrowly and broadly, is to make the listener agree to a certain extent with the claims about the world that are formulated in speech exchange.

Based on these theories and our previous studies, we have reformulated the notion of multimodal argumentation, understood here as the strategic use of the combination of different semioses as an argumentation strategy. In this paper, we replaced the proposal of multimodal argumentation in three fixed categories, namely multimodal argumentation by focusing, multimodal argumentation by complementarity, and multimodal argumentation by divergence, with a scalar analysis of the relationship of focus and dissonance established between word and image in discourse.

Due to the complexity of the material analyzed, this proposal will be further refined in future work. However, we assume that the discussion in this article indicates that it is useful to consider multimodal argumentation from a scalar perspective, a perspective that can also be used to analyze other discursive genres constituted by the combination of word and image.

AUTHORS' CONTRIBUTION

All authors participated in all stages of the article (conceptualization, formal analysis, research, methodology, revision and editing).

FINANCING

This work was not funded by any institution that promotes research.

CONFLICT OF INTERESTS

The authors declare that there are no conflicts of interests in this work.

REFERENCES

AMOSSY, Ruth. Is it possible to integrate argumentation into discourse analysis? Problems and challenges. **ReVEL**, special issue vol. 14, n. 12, 2016. Available at: <www.revel.inf.br.>. Accessed: June 20, 2022.

BARTHES, Roland. **The obvious and the obtuse**: **critical essays III** (L. Novaes, Transl.). Rio de Janeiro, RJ: Nova Fronteira, 1990.

CHARAUDEAU, Patrick. A semiolinguistic analysis of text and discourse. *In*: PAULIUKONIS, Maria Aparecida Lino; GAVAZZI, Sigrid (Orgs.). **From language to discourse**: reflections for teaching. Rio de Janeiro: Lucerna, 2005.

CHARAUDEAU, Patrick. **Media discourse**. Translated by Angela M. S. Corrêa. São Paulo: Context, 2007.

CHARAUDEAU, Patrick. **Language and discourse**. **Modes of organization**. Coordination of the translation team. by Angela M. S. Corrêa & Ida Lúcia Machado. 1st ed., 1st reprint. São Paulo: Context, 2008.

CHARAUDEAU, Patrick. Television patemization as a strategy of authenticity. *In*: MENDES, Emília; MACHADO, Ida Lúcia. **As emoções no discurso**, v. II. Campinas: Mercado das Letras, 2010.

CHARAUDEAU, Patrick. Communication contract. *In*: CHARAUDEAU, Patrick; MAINGUENEAU, Dominique. **Dictionary of discourse analysis**. São Paulo: Context, 2014 [2004].

JAKOBSON, Roman. **Linguistics and communication**. Trad. Isidoro Blikstein and José Paulo Paes. 4 ed. São Paulo: Cultrix, 1969.

KRESS, Gunter; VAN LEEUWEN, Teo. **Reading images**: **the grammar of visual design**. London: Routledge, 2006.

MARCHON, Amanda Heiderich; GARCIA, Carlos Eduardo Nunes. Intertextualidade multimodal como estratégia argumentativa. **Revista (Con)Textos Linguísticos**, Vitória, v. 14, n. 27, p. 1-1, dez./2020.

PASSOS DA SILVA, Luzia de Cássia Almeida. **The representation of political women on the covers of Veja:** a semiolinguistic analysis. PhD Thesis in Vernacular Literature. Rio de Janeiro: Federal University of Rio de Janeiro, 2021.

SANTAELLA, Lúcia. **Reading Images** (How I Teach Collection). São Paulo, SP: Melhoramentos.

SILVA, Welton Pereira e; MARCHON, Amanda Heiderich. **Argumentação multimodal: uma proposta teórico-metodológica**. Acta Scientiarum. Language and Culture, v. 43, 2021. Disponível em: <https://doi.org/10.4025/actascilangcult.v43i1.56894>.