

Blended Classic Joint Attention and Multimodal Deixis

Atenção Compartilhada Clássica Mesclada e Dêixis Multimodal

Mark Turner

Institute Professor at Case Western Reserve University

Maíra Avelar

Universidade Estadual do Sudoeste da Bahia - UESB – Vitória da Conquista - Bahia – Brasil

Milene Mendes de Oliveira

University of Potsdam, Potsdam, Germany



Resumo: In this paper, we aim at presenting the concept of Blended Classic Joint Attention, BCJA (TURNER, 2015). At first, we discuss the concept of Classic Joint Attention (TOMASELLO, 1995, THOMAS; TURNER 2011) and then illustrate BCJA using a scene from a video made by Jout Jout, a Brazilian youtuber. After that, we discuss the concept of multimodal deixis and present our data which consist of four samples collected from the Distributed Little Red Hen Lab. Two of them are from Brazilian TV news and the other two from the US talk-show “The Daily Show”. We analyzed the multimodal verbo-gestural compounds in which the spatial deictics “here” and “aqui” co-occur with manual gestures and eye gaze. Results show that in less prototypical examples, with fixed expressions, there is no interaction with the viewer. On the other hand, in more prototypical examples, the eye gaze is aimed at the viewer, while the manual gesture is aimed at the immediate interlocutor.

Keywords: Blended Classic Joint Attention. Multimodality. Spatial deixis.

Resumo: Neste artigo, temos por objetivo apresentar o conceito de “Atenção Compartilhada Clássica Mesclada”, ACCM (TURNER, 2015). Primeiramente, discutimos o conceito de “Atenção Compartilhada Clássica” (TOMASELLO, 1995, THOMAS; TURNER, 2011), e, em seguida, ilustramos a ACCM utilizando cenas de um vídeo realizado pela youtuber brasileira Jout Jout. Na sequência, discutimos o conceito de dêixis multimodal e apresentamos nossos dados, que consistem em quatro amostras coletadas no *Distributed Red Hen Lab*. Dois vídeos pertencem a telejornais brasileiros e dois, ao programa de auditório norte-americano “The Daily Show”. Analisamos compostos verbo-gestuais em que os dêiticos espaciais “aqui” e “here” co-ocorrem com gestos manuais e olhares. Os resultados demonstraram que, em exemplos menos prototípicos, em que ocorrem expressões verbais fixas, não há interação com o telespectador. Por outro lado, em exemplos mais prototípicos, os olhares são direcionados ao telespectador, enquanto os gestos manuais são direcionados ao interlocutor imediato.

Palavras-chave: Atenção Compartilhada Clássica Mesclada. Multimodalidade. Dêixis espacial.

1. Introduction

Joint attention--i.e., the shared focus of two or more people on the same object--is a phenomenon that has been discussed by cognitive linguists for about two decades. With the development of multimodal resources related to media such as TV and Youtube, there is a growing need to discuss multimodal mechanisms related to the shared space. Situations worthy of notice are, for instance, those in which there is an immediate interlocutor (such as an interviewer and an interviewee in a talk show) but also the viewer, who is absent in the immediate scene, but is considered part of the broader interaction scene. In this regard, there might be an attempt not only to address this viewer, but also to persuade her through the use of joint attentional mechanisms.

In this paper, we intend to demonstrate how these mechanisms of Blended Classic Joint Attention work. To do so, we first explore the concepts of joint attention and "Classic Joint Attention" (THOMAS; TURNER, 2011), and then explain and exemplify mechanisms of BCJA. After that, we describe and analyze instances of multimodal deixis and focus our analysis on some examples of the deictic expressions "here" and "aqui," extracted from The Distributed Little Red Hen Lab videos of interviews given on Brazilian and American TV programs.

2. Classic Joint Attention and Blended Classic Joint Attention

Joint attention can be described as a cognitive phenomenon at the basis of human interaction. A basic scene of joint attention, as defined by Tomasello (1995), is the perception that two or more people in a certain environment are attending to the same thing (be it an object or an idea) at the same time. Writing about the development of the cognitive capacity for joint attention, Tomasello (1995, p. 107) explains that "In joint attention, the child coordinates her attention to the object and the adult at the same time that the adult coordinates her attention to the same object and the child... The

coordination that takes place in joint attentional interactions is accomplished by means of an understanding that the other participant has a focus of attention to the same entity as the self." This is connected with the understanding of the other as an intentional agent and the understanding that participants are sharing an intentional relation to the world. The following picture illustrates the process of joint attention:



Figure 1: Joint attention

Tomasello (1995) argues that when humans use language to communicate referentially, they are attempting to manipulate the attention of another person or persons. Thus, in this view, linguistic reference may implicitly communicate the following intention to an interlocutor: 'I intend for you to attend to X,' that is to say, 'I intend for you to perceive X in such and such a way.'

Apart from Classic Joint Attention, Turner (2015) presents the concept of Blended Classic Joint Attention, which is especially useful for the analysis of TV videos, films, and other multimodal means of communication. This specific type of classic joint attention is 'blended' (in reference to Fauconnier and Turner's blending theory from 2002) because of its interplay with the actual studio situation (cameras, lights, production crew, etc.). Thus, the input space 'actual experience' gets blended with the input space 'classic joint attention.' Therefore, whenever a news anchor looks at the camera and utters 'it is good to have you here,' she is leaning on the audience knowledge of how Classic Joint Attention is established. Due to blending, the audience can understand what the anchor means even if 'you' here has a very different

meaning from its prototypical use in classic joint attention, when two people are talking to one another. In Classic Joint Attention, interactants know that they are jointly attending to something. Moreover, they are aware that they all know this and that they are engaging with each other through joint attention even if they are not communicating about it (TOMASELLO, 1995). In classic joint attention (THOMAS; TURNER, 2011), two or more people together are jointly attending to something which they both perceive in the same human environment and they are communicating about it. For instance, as in the case of the two girls in Figure 2 talking about the flowers they just picked:



Figure 2: classic joint attention

Thomas and Turner (2011) argue that the human ability to communicate, which includes linguistic and gestural means, is particularly dedicated to this basic scene of classic joint attention. In such a scene of Classic Joint Attention, participants have an understanding of “the ground;” i.e., they understand a great deal about “the speech event, its setting, and its participants” (LANGACKER, 1985, p.113) without the need to refer to that understanding. In face-to-face communication, linguistic clues of Classic Joint Attention include deixis and certain constructions, while non-verbal clues include eye contact and gestures (such as pointing).

The following example, presented in Mendes de Oliveira (2017), which illustrates instantiations of BCJA, shows scenes of a video created by a famous Brazilian youtuber called Júlia Tolezano (nickname, Jout Jout), whose channel is called “Jout Jout Prazer”. Jout Jout is the opposite of the traditional youtuber: the ideas for her

video content come from ordinary experiences like a five-minute video on how she is happy about having abandoned wearing bra with push-up pads, or how women and men should not feel bad about their cellulite. The scenes were extracted from a video about a Brazilian movie “Que horas ela volta?”¹. The youtuber is emphatically talking about a character of the movie, Jéssica, who is Regina Casé’s (a famous Brazilian actress) daughter (TOLEZANO, 2016). She is reporting on a specific scene of the movie.

1. Regina Casé, Jessica's mother, says:
2. Jessica, listen: I think you are too arrogant. You think you are superior to everyone here.
3. And then Jessica says:
4. Queen... she says:

¹ In English, the literal translation would be: “What time does she come back?”. In English-speaking countries, it was launched under the title “The second mother.” The movie is about the relationship of a middle-class family, their maid, and their maid's daughter.



Figures 1-6: Blended Classic Joint Attention

In scene 1, Jout Jout turns her eye gaze to the camera and uses a palm up open hand gesture, metaphorically presenting an idea and informing that she is going to make a direct quote. When doing so, in scene 2, she rotates her torso and directs her eye gaze to the narrative space created by her speech (SWEETSER, 2017). The pointing-up gesture is related to the discursive marker “listen,” pragmatically related to an attitude of reprehension. In scene 3, Jout Jout turns her eye gaze again to the camera, doing a palm up open hand with both hands when informing that she is going to directly quote Jéssica’s saying. In the fourth scene, Jout Jout makes a comment that works as a parenthesis, about Jéssica, calling her a “queen.” This comment is made with a palm up open hand gesture, with both forearms horizontally opened, as if she were performing the greatness of the described character. In the scene 5, while performing Jéssica answering the former reprehension made by her mother, Jout Jout’s face is replaced by Jéssica’s, whose eye gaze replaces Jout Jout’s gaze. The gesture made with both hands and three joints extended also demonstrate the greatness that the performer attributes to the performed character. In the last scene, Jout Jout performs an enacting gesture (MÜLLER 2013a) of “slapping,” right before saying “in everyone’s face, Jéssica.” She also turns her eye gaze to the camera, at the same time as she simulates a dialogue with Jéssica.

In the aforementioned scenes we have very explicit blends that are made explicit by the use of special effects. The aspect of viewpoint (SWEETSER, 2017) is also made explicit in the BCJA. This character’s viewpoint is blended with the speaker’s viewpoint. This is shown by a type of special effect in which a picture of the character Jéssica is put on the screen. The performer reproduces an utterance used by the character in direct speech. Her voice is changed in the video; the text is presented as a quote and its authorship is shown. However, this authorship is connected with the viewpoint of the youtuber (“Jessica, my owner”). The special effects represent, very explicitly, this blend between youtuber and character.

In terms of BCJA, the ‘object of attention’ is this character and several clues are given to the addressee so as to confirm where the focus of attention is. Therefore, there is an exaggeration--a redundancy--that creates a comic effect. This is actually a character viewpoint (from a viewer’s perspective), because when the youtuber performs the palm down open hand gesture (scene 5), she is reporting on Jessica’s performance on the movie. More specifically, she is talking about how the behavior of this character affects the viewers of the movie. We have a very lively BCJA moment in which the viewer is not only listening, but also being slapped in the face by the youtuber.

3. Blended Classic Joint Attention and Multimodal Deixis

In this paper, we specifically focus on the speech-gesture compound to make considerations about deixis and multimodality. Based on Avelar and Ferrari’s (2017) analyses of the adverb “aqui” (‘here’), we can argue that the gesture “pointing down” would be the prototypical one for the verbal occurrence of “here.” As previously seen, pointing gestures are valuable resources for establishing Classic Joint Attention, whereas eye gaze is a valuable resource for building the space in which the subject establishes a connection with the physically absent interlocutor (SWEETSER, 2017) in Blended Classic Joint Attention. In the next subsection, we are going to analyze how these and other resources are

used to establish Classic and Blended Classic Joint Attention.

3.1 Data collection: The Distributed Little Red Hen Lab

The Distributed Little Red Hen Lab constitutes a large multimodal database, hosted by UCLA and directed by professors Francis Steen and Mark Turner. It comprises thousands of hours of TV news from all over the world. As stated by Turner (2017): “The Red Hen project aims to integrate a multimodal data collection of global television news with data enhancement, multimodal data mining, and cognitive analysis of multimodal communicative practices.” Specifically, TV news can be considered as a common communicative practice which aims at grabbing the interlocutor’s attention by compressing, in a quick format, complex meanings (TURNER, 2017)

To perform our illustrative analysis about the deictic “here,” in English, and the deictic “aqui” in Brazilian Portuguese, in 2017, we collected some video samples from the talk-show “The Daily Show,” exhibited by the Comedy Center channel and from Band and Globo channels exhibited in the same year in Brazil. In the first illustrative sample, Trevor Noah, the show’s host, introduces a guest by saying: “Our guest is here, everybody.” In the second sample, Noah says: “Here is what everyone struggles with,” while performing a gesture with a right open hand, vertically turned to the immediate interlocutor, with all fingers bended. In the third sample, we have the governor of São Paulo, Geraldo Alckmin, saying that “I want here to highlight the government efforts with the reforms.” In the last sample, we have an interviewee that says: “He hit his head here,” while self-pointing to her nose. All the samples can be visualized in the next subsection.

3.2 Data analysis

At first, we analyzed the prototypical uses of “here” and “aqui.” The uses considered as prototypical are the ones in which the subject points to a physical

referent on the immediate environment, such as in scenes 1 and 2 below:



Figures 7-8: Prototypical samples of spatial deictics

In scene 7, it is possible to visualize a “palm presenting” (PP) gesture (KENDON, 2004), in which the host uses a palm horizontal open hand, palm pointing to the immediate interlocutor, to point to the guest and introduce her to the audience and, at the same time, to the viewers to whom addresses his gaze. Here, the host reproduces a scene of classic joint attention, namely, when person A is introducing person B to person C. The gaze to the camera is supposed to imitate the gaze to person C.

In the second sample, scene 8, the interviewee makes a self-pointing gesture (KENDON, 2004) to inform the interviewer about the exact physical location where a fracture was generated: her own nose. At the same time, she laterally looks at the camera, and not in the interviewer’s direction. This example shows that also people who are not trained into BCJA techniques (as TV hosts and anchors certainly are) attempt to simulate joint attention with the TV viewers.

We are also going to discuss non-prototypical uses of “here” and “aqui,” considering that these uses occur when the subject points to the space, in fact

pointing to an abstract referent (MCNEILL; CASSELL; LEVY, 1993, AVELAR; FERRARI, 2017). These uses can be visualized below, in scenes 3 and 4:



Figures 9-10: Non-prototypical samples of spatial deictics

In scene 3, Noah is looking directly to his immediate interlocutor, the interviewee, while performing a precision gesture (KENDON, 2004), as if he were grabbing a 3D object (MÜLLER, 2013a) when saying “here.” In this case, the deictic is part of the micro-construction (TEIXEIRA; OLIVEIRA, 2012, PINHEIRO, 2017) “here is...” that is related to the interlocutor’s attention and contributes to the discourse sequencing. Therefore, this micro-expression has an interactional function, once this kind of expressions “migrated from more prototypical uses to uses with a higher degree of subjectivity, moving the interlocutor’s focus of attention from the space to the text².” (TEIXEIRA; OLIVEIRA, 2012, p.34). ‘Here’ refers to the object of attention in Classic Joint Attention. In the shared focus of attention, ‘here’ does not refer to a certain physical space or location, but to the text that is about to be presented.

² Original excerpt: “[Essas expressões]migraram de usos mais prototípicos, até empregos com maior grau de subjetividade, deslocando o foco de atenção do leitor/interlocutor do espaço para o texto”.

In scene 4, the former governor Geraldo Alckmin is performing a palm up open hand gesture (MÜLLER 2013b), with horizontal palms facing each other and directed toward the own body. Considering the eyegaze, Alckmin is also directly looking at his immediate interlocutor. In this case, we would consider this occurrence as an “in-between” case in which “here” still conserves some of its deictic properties but is more related to a discursive reference (as in: “here, in this interview”)³. In either case, it functions as an emphasis to the topic being discussed by the governor: the object of attention in CJA.

3.3 Discussion

In the analyzed data, we have BCJA mechanisms used in the first two samples, related to prototypical uses of the deictic “here/aqui”: the eyegazes are directed to the camera in an attempt to simulate a look into the viewers’ eyes, while the gestures address the immediate interlocutors. On the other hand, when the deictic uses are less prototypical, mechanisms related to CJA only (and not BCJA) seem to gain prominence: it is the immediate interlocutor who is aimed at by the subjects. Our analysis shows that one of the reasons for the absence of BCJA mechanisms seems to be the use of here/aqui as expressions containing the deictic work as discursive markers which are giving sequence to speeches addressed to the immediate interlocutor. Consequently, a hypothesis which could be further tested is that the more conventionalized the use of the deictic (as in fixed expressions) is, the less specific and directed the use of the speech-gesture compound seems to be.

By briefly presenting these illustrative analyses, we intend to foster further research into multimodal deixis and BCJA in videos.

³ In this occurrence of ‘here,’ we could identify the conceptual metaphor “TIME IS SPACE” (LAKOFF; JOHNSON, 1980).

References

- AVELAR, M.; FERRARI, L. Integração experiencial e dêixis: O papel discursivo dos gestos. *Cadernos de Estudos Linguísticos*, v.59, n.1, Campinas, p. 73-89 - jan./abr. 2017.
- FAUCONNIER, G.; TURNER, M. *The way we think: conceptual blending and the mind's hidden complexities*. New York: Basic Books, 2002.
- KENDON, A. *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press, 2004
- LAKOFF, G.; JOHNSON, M. *Metaphors we live by*. Chicago: Chicago University Press, 1980.
- LANGACKER, R. W. (1985). Observations and speculations on subjectivity. *Iconicity in syntax*, v. 1, p. 109.
- MCNEILL, D.; CASSELL, J.; LEVY, E. T. Abstract deixis. *Semiotica*, 95(1-2), p. 5-20, 1993.
- MENDES DE OLIVEIRA, M. Blended classic joint attention in user-created Youtube videos. Oral presentation. International Conference on Multimodal Communication. Osnabrück, 2017.
- MÜLLER, C.: Gestural Modes of Representation as techniques of depiction. In: MÜLLER, C.; CIENKI, A.; FRICKE, E.; LADEWIG, S.; MCNEILL, D.; TESSENDORF, S. (Eds). *Body - Language – Communication*. v. 1. Berlin/Amsterdam/New York: 2013a, p. 1687-1701.
- MÜLLER, C.: Forms and uses of Palm Up Open Hand gestures: a case of a gesture family? In: MÜLLER, C.; CIENKI, A.; FRICKE, E.; LADEWIG, S.; MCNEILL, D.; TESSENDORF, S. (Eds). *Body - Language – Communication*. v. 1. Berlin/Amsterdam/New York: 2013b, p. 233-256.
- PINHEIRO, H. *Uma análise multimodal do “aqui” em dados multimodais*. Dissertação de Mestrado. Universidade Estadual do Sudoeste da Bahia. Vitória da Conquista, 2017.
- SWEETSER, E. Viewpoint and multimodality. Oral presentation. 6th Conference on Metaphor in Language and Thought. Salvador, 7 October 2017.
- TEIXEIRA, M.; OLIVEIRA, M. Por uma tipologia funcional dos marcadores discursivos com base no esquema construcional Verbo Locativo. *Veredas* (UFJF. Online), v. 16, p. 19-35, 2012.
- THOMAS, F. N.; TURNER, M. *Clear and simple as the truth: writing classic prose*. Princeton University Press, 2011.
- TOMASELLO, M. Joint attention as social cognition. *Joint attention: Its origins and role in development*, 1995, p. 103-130.
- TURNER, M. Blending in Language and Communication. In: Dąbrowska, E. and Divjak, D (Eds.) *Handbook of Cognitive Linguistics*. Berlin: De Gruyter Mouton, 2015, p. 211-232.
- TOLEZANO, J. Um bando de tapa na cara. [Video file]. Retrieved from <https://www.youtube.com/watch?v=r-tuxDWfEpE>. Date of access: 20 October 2016.

COMO CITAR ESSE ARTIGO

TURNER, Mark; AVELAR, Maíra; MENDES DE OLIVEIRA, Milene. Blended Classic Joint Attention and Multimodal Deixis. *Signo*, Santa Cruz do Sul, v. 44, n. 79, p. 03-09, jan. 2019. ISSN 1982-2014. Disponível em: <<https://online.unisc.br/seer/index.php/signo/article/view/12710>>. Acesso em: _____.
doi: <https://doi.org/10.17058/signo.v44i79.12710>.