

## Combining Multimodal Techniques to Approach the Study of Academic Lectures: A Methodological Reflection

EDGAR BERNAD-MECHÓ

Universitat Jaume I

ebernad@uji.es

This article offers a methodological reflection on the use of multimodal techniques for the study of academic lectures. Three distinct multimodal approaches have been put forward to explore the use of language holistically, namely, multimodal social semiotics (MSS), multimodal discourse analysis (MDA) and multimodal interaction analysis (MIA). These approaches differ in their main focus—the social context, the system of semiotic resources available to the speakers and the social actors, respectively—and the tools they provide to conduct multimodal analyses. To exemplify how analyses may be conducted within each of the paradigms in the context of academic lectures in English, I examine an excerpt extracted from an African-American history lecture from Yale University by a native English speaker in which he organizes his discourse in between content sections. Through the use of short multimodal transcriptions, I discuss how MSS can be used for reflections on the social contexts of academic lectures, MDA describes the use of semiotic resources employed by the lecturers, and MIA can be used to look into how lecturers structure their speech into sequences of actions. Ultimately, I suggest a combination of multimodal methodologies to obtain a broader account of the intricacies of discourse in academic settings.

**Keywords:** multimodality; academic lectures; multimodal social semiotics; multimodal discourse analysis; multimodal interaction analysis

...

## Combinación de técnicas multimodales para el estudio de clases universitarias: una reflexión metodológica

Este artículo ofrece una reflexión metodológica sobre el uso de técnicas multimodales para el estudio de clases universitarias. Existen tres enfoques metodológicos para el estudio holístico del lenguaje: la semiótica social multimodal o *multimodal social semiotics* (MSS), el análisis del discurso multimodal o *multimodal discourse analysis* (MDA) y el análisis de la interacción multimodal o *multimodal interaction analysis* (MIA). Estos enfoques difieren en sus principales focos de atención—el contexto social, el sistema de recursos semióticos disponible para la comunidad de hablantes y los agentes sociales, respectivamente—y las herramientas que proporcionan para llevar a cabo análisis multimodales. Para ejemplificar cómo se pueden llevar a cabo análisis dentro de cada uno de estos paradigmas en el contexto de las clases universitarias en inglés, examino un fragmento extraído de una clase de historia afroamericana de la Universidad de Yale impartida por un hablante nativo del inglés, fragmento en el cual el profesor organiza su discurso entre secciones de contenido. A través de transcripciones multimodales breves, trato el modo en que el MSS puede ser utilizado para ofrecer reflexiones sobre los contextos sociales de las clases universitarias, el MDA describe el uso de recursos semióticos utilizados por el profesorado, y el MIA puede ser usado para indagar en la estructuración del discurso del profesorado en secuencias de acciones. En último término, propongo una combinación de estas metodologías multimodales para obtener una visión más amplia de las complejidades del discurso en contextos académicos.

Palabras clave: multimodalidad; clases universitarias; semiótica social multimodal; análisis del discurso multimodal; análisis de la interacción

## 1. INTRODUCTION

The research on academic genres in recent decades has expanded to consider not only written but also oral varieties, and from monomodal to multimodal studies (Crawford Camiciottoli and Fortanet-Gómez 2015). Multimodality analysis stems from the principle that language is inherently multimodal; in other words, meaning is not conveyed through one means—or mode—only, but through a multiplicity of modes working together (Kress and van Leeuwen 1996). Moreover, in contrast to traditional studies in linguistics, language is no longer central in multimodal approaches, but rather is just one more element that contributes to the conveyance of meaning. As Gunther Kress states, a “multimodal approach assumes that language, whether as *speech* or as writing, is one means among many available for representation and for making meaning. That assumes that the meanings revealed by forms of DA [Discourse Analysis] relying on an analysis of writing or speech are only ever ‘partial’ meanings. The meanings of the maker of a text as a *whole* reside in the meanings made jointly by all the modes in a text” (2012, 37; italics in the original).

In academic settings, multimodality has mainly been applied to the study of conference presentations (Querol-Julián and Fortanet-Gómez 2012) and lectures (Fox and Artemeva 2013; Crawford Camiciottoli 2015, 2016; Bernad-Mechó 2017), the latter being the academic genre that has received most attention. Much of the focus within lecture analysis research has been placed on comprehension (Flowerdew 1994; Buck 2001) and the interpersonal factors influencing the learning process (Flowerdew and Miller 1992; Rounds 1987). In addition, numerous studies have been conducted with the aim of exploring lexico-grammatical features in the discourse of lectures (Biber et al. 2004; Fortanet-Gómez 2004). This work, however, focuses on linguistic aspects of the lectures, lecturers and audience, and fails to examine elements beyond language. A multimodal analysis might provide a holistic view of how communication occurs (Norris 2004).

In recent years, and with the rise of multimodal methodologies, nonlinguistic elements of lectures have become the focus of research: nonverbal embodied modes (Thesen 2016; Morell 2018; Bernad-Mechó and Fortanet-Gómez 2019), the use of visuals (Gunel et al. 2006), the use of the board (Fox and Artemeva 2013) and multimedia learning (Tan et al. 2016). Other studies have explored the relationship between linguistic and nonlinguistic elements. Inmaculada Fortanet-Gómez and María Noelia Ruiz-Madrid, for example, explore the multimodal use of questions in lectures by guest speakers to facilitate comprehension in nonnative speakers of English (2014). Belinda Crawford Camiciottoli analyzes five humanities lectures with a focus on how verbal explanations are carried out and co-occur with prosodic stress, gaze and gestures (2015) and interactional devices between lecturers and students and how this interaction enhances comprehension (2016). Finally, elsewhere I have explored the relationship between organizational metadiscourse and multimodal embodied elements in academic lectures (2017). All in all, through

the use of different methodological approaches, multimodal research sheds new light on the possibilities of conducting multimodal analyses of lectures, always with the common underlying objective of describing the strategies that contribute to facilitate students' comprehension of lectures.<sup>1</sup> However, there is still much room for expansion by combining the most widely used multimodal methodologies available, as exemplified in this article.

### 1.1. Three Approaches to Multimodality

Carey Jewitt (2014a) distinguishes three main approaches to multimodality: multimodal social semiotics (MSS; Kress and van Leeuwen 1996, 2001), multimodal discourse analysis (MDA; O'Halloran 2004, 2011) and multimodal interaction analysis (MIA; Norris 2004).<sup>2</sup> Each of these approaches provides specific analytical tools, and they differ as regards their underlying theories, the degree of importance placed on the context and the interrelationship between modes and the degree of attention paid to the sign-maker (Jewitt 2014a, 36). Thus, MSS shows a particular interest in the sign-maker in a social context, MDA focuses rather on semiotic resources as systems of choices, and MIA's main interest lies in the social actors' interaction.

The first approach, MSS, takes M. A. K. Halliday's (1978, 1985) social semiotics and systemic functional linguistics further in order to explore the process of meaning-making as a social activity that takes place in social environments (Kress 2010). Two main aspects of communication are especially relevant within this model: the study of distinct modes or semiotic resources in the creation of meaning and the significance of the social context where the latter occurs. The emphasis is therefore on the sign-maker and the semiotic choices they make in a given context. In other words, MSS is interested in how writers/speakers make meaning through the use of a multiplicity of modes in a social world. Elisabetta Adami summarizes the main purposes of MSS as follows: "social semiotics conceives of sign-making as the expression of social processes; through a fine-grained qualitative analysis of usually small samples of texts, social semiotics is interested in unveiling ideologies, social values, power roles, and identities as expressed in texts, together with how individuals actively maintain, reinforce, contest and challenge them through their sign-making choices" (2017, 455).

---

<sup>1</sup> A pedagogical application can be derived from of multimodal analyses, since the results obtained may be used in teacher-training projects.

<sup>2</sup> It could be argued that there are other more or less established approaches to multimodality, such as multimodal critical discourse analysis (van Leeuwen 2008), cognitive linguistics multimodality (Forceville and Urios-Aparisi 2009) and multimodal conversation analysis (Deppermann 2013). However, these approaches have not received as much attention as those discussed in this article. Furthermore, their main interests lie in the theories from which they emerge—critical discourse analysis, cognitive linguistics and conversation analysis, respectively—instead of offering a straightforward multimodal analysis of lectures.

The second approach, MDA, is strongly based on Halliday's systemic functional linguistics (1985) and is concerned with the study of language as the result of a combination of semiotic resources; in particular, it aims to explore models of semiotic resources, how they interact to convey meaning and how meaning can change through a resemioticization process (Iedema 2003). In MDA, language seems to be prioritized over other modes; to some extent, language is the point of departure that is then explored in combination with the other modes. This is particularly relevant in the study of lectures, since spoken language is seen as the fundamental means for transmitting content in this genre (Crawford Camiciottoli and Fortanet-Gómez 2015). Simply put, MDA analyses commonly start from a detailed comprehensive examination of short parts of texts—understood as any source that is capable of conveying meaning—with the aim of developing theories and frameworks that explain how semiotic resources work.

Finally, MIA is grounded in Ron Scollon's theory of mediated discourse analysis and is organized around the concept of mediated action (2001). In short, Scollon proposes an approach to the analysis of discourse where the analyst detaches from the text—what is said—and focuses on the actions performed by the social actors—the speakers. This approach is applied by Sigrid Norris to the multimodal study of communication (2014). Thus, MIA is interested in how social actors interact in specific situated contexts; in her own words, “the work of the actor is, therefore, everything” (Jewitt 2104a, 33).

## 1.2. Theoretical and Methodological Tools

The concept of *mode* is a key element within all multimodal studies. Kress and Theo van Leeuwen describe modes as semiotic systems with rules and regularities—images, gestures, speech, music, layout, writing, proxemics, posture, etc. (2001). Moreover, the concept of mode is closely related to that of *semiotic resource*: “the actions, materials and artifacts we use for communicative purposes, whether produced physiologically—for example, with our vocal apparatus, the muscles we use to make facial expressions and gestures—or technologically—for example, with pen and ink, or computer hardware and software—together with the ways in which these resources can be organized” (van Leeuwen 2005, 285). The concepts of mode and semiotic resource are often used interchangeably to describe the different elements analyzed in a given communicative act. Kay O'Halloran, for instance, refers to *semiotic resources* as being “used to describe the resources (or modes) (e.g. language, image, music, gesture and architecture) [...] in multimodal texts” (2011, 121).

Apart from the concept of mode/semiotic resource—widely used throughout multimodal research—*multimodal affordance* and *multimodal ensemble* are also recurrently present in MSS studies. Multimodal affordance alludes to the fact that modes can create meaning while accentuating the importance of cultural aspects.

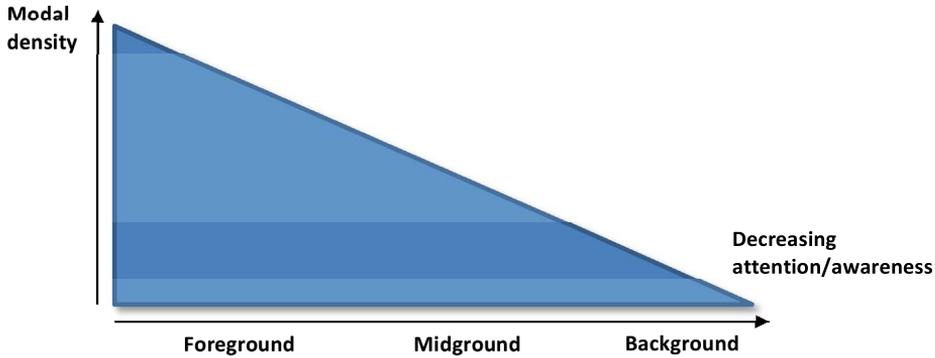
Thus, it emphasizes the importance of specifying context and social environment as intrinsically connected to the cultural aspects of a given communicative act. As Kress puts it, “meanings are socially made, socially agreed and consequently socially and culturally specific” (2010, 88). On its part, multimodal ensembles refer to the specific combinations of modes that occur in communication. The study of how different modes contribute to the creation of meaning as well as how this interrelationship itself works in the meaning-making process is paramount in MSS analysis.

The concepts of multimodal affordance and multimodal ensemble are also widely used in MDA studies, the main difference lying in the focus of analysis within each paradigm. While MSS pays close attention to the role of context in communication, MDA aims to develop frameworks to explain how specific combinations of modes work in semiotic events in order to reveal recurrent patterns. These analyses are often conducted with multilayer annotation software such as ELAN (Wittenburg et al. 2006) or Multimodal Analysis-Video (MMA-Video; O’Halloran et al. 2012). These tools allow the analyst to visually represent specific combinations of modes by creating a series of layers to describe and annotate each of the modes. In the case of MMA-Video, once a video file has been annotated, the analyst can also easily obtain quantitative data pertaining to the use of semiotic resources as well as a representation of the specific combination of modes at a given moment.

Finally, in MIA mediated action is the essential unit of analysis (Scollon 2021). Norris describes three types of actions: lower-level, higher-level and frozen. She defines the former as “the smallest interactional meaning unit” (2004, 11)—a change in gaze, an utterance or the movement of a hand. They are clearly identified with a beginning and an end. Higher-level actions, on the other hand, are successions of lower-level ones, also encompassed within a beginning and an end, such as a conversation with a friend or the action of making a coffee. Lastly, frozen actions are “higher-level actions which were performed by an individual or a group of people at an earlier time than the real-time moment of the interaction that is being analyzed” and are “frozen in the material objects” (2004, 13). For example, a cup of coffee on a table is made up of the higher-level actions making a mug, making coffee, serving coffee, etc. Frozen actions usually show characteristics of the social actors that use them; for instance, a playlist might provide an insight into a social actor’s tastes in music.

Social actors typically carry out a number of higher-level actions in any one single instance. These actions are produced across different levels of attention/awareness. For example, if a speaker is talking with a friend, they could also be making coffee at the same time and even thinking about holidays. Using the concept of *foreground-background continuum* (Norris 2004), the different levels of attention/awareness may be represented in a diagram showing which actions are foregrounded—receive most attention—midgrounded or backgrounded—receive least attention (figure 1).

FIGURE 1. Modal density foreground-background continuum of attention/awareness (Norris 2004, 99)



The concept of foreground-background continuum is closely linked to that of *modal density*. When actions receive more attention, they become modally dense. This may occur through higher modal complexity—when an action is carried out through many modes—or through higher modal intensity—when one of the modes takes on a more intense role. These concepts, in turn, are related to that of *modal configurations*, which are somewhat similar to multimodal ensembles, with the particularity that modes are represented in a hierarchical order, providing a holistic view of which ones play a greater or lesser role in a given action.

Finally, the last concept commonly employed in MIA is that of *semantic/pragmatic means*, which are “pronounced lower-level actions (e.g. a gesture, a head movement, laughter, etc.) that indicate a shift in the level of attention that a social actor devotes to particular higher-level actions” (Bernad-Mechó 2017, 44). These elements are particularly frequent when actions are foregrounded (Norris 2004) and they fulfill two functions: semantic—since they structure the social actor’s own structure of actions—and pragmatic—since the shift in the level of attention is communicated to other actors participating in the interaction.

MSS, MDA and MIA have each provided the basis for numerous studies with different foci. Choosing one approach over another in any one study will depend on the objectives and research questions. Explicit combinations of the three approaches are rare, but I argue that they are not mutually exclusive. In fact, boundaries are not always clear-cut and they can be remade and contested (Jewitt 2014a, 29). In the present study, I aim to show the affordances of each of these approaches when analyzing academic lectures. Ultimately, I claim that a combination of the three approaches will provide a deeper insight into the intricacies of communication in lectures and expand the possibilities of multimodal analysis.

## 2. METHODOLOGY

In order to demonstrate the capabilities of each of the multimodal paradigms presented in section 1, I use a twenty-five-second-long fragment from lecture thirteen—“The Road to Brown and Little Rock”—in the course African American History: From Emancipation to the Present from Yale University’s Open Yale Courses, a collection of face-to-face university lectures uploaded on Yale University’s OpenCourseWare. The fragment corresponds to minute 08:04 to 08:29. During this excerpt, the lecturer organizes the contents of the lecture by establishing a connection with the previous lecture and introducing the main topic of the present one. This fragment was chosen because organizational sections are more multimodally dense than content ones and can consequently be expected more suitable to the methodological approaches discussed here.

In order to select this lecturer from the many available on the platform, lecturing styles were considered. Following Tony Dudley-Evans (1994), I chose a conversational-style lecturer, as they follow clear class structures and use a wide repertoire of semiotic resources (Bernad-Mechó and Fortanet-Gómez 2019). The lecturer is a native speaker of English addressing a—presumably—English-speaking audience at Yale University, and hence no language accommodations are present in the lecture. The excerpt is part of a larger corpus of social studies lectures intended to explore the multimodality of organizational sections across lecturing styles (Bernad-Mechó 2018). A full verbal transcript of the excerpt is provided below following Norris’s transcription system (2004).<sup>3</sup>

08:05:23	Now I ended the last lecture before the midterm talking about A. Philip Randolph’s March on Washington Movement
08:11:20	his bluff with FDR
08:14:07	FDR’s issuing erm of Executive Order 8802
08:17:16	and the creation of the FEPC
08:21:12	I was trying to cram a lot of things into the lecture
	
08:23:25	(sounds)
08:24:05	two minutes to talk about the FEPC very quickly and move us
08:27:16	um... into the 1940s

The analysis of the excerpt was carried out in several steps. First, the modes that are most relevant in lectures were selected. In this sense, Sabine Tan et al. identify speech, gaze,

<sup>3</sup> This transcript corresponds only to the verbal mode. The full fragment utilized for the analyses ranges from minute 08:04:00 to 08:29:00, as nonverbal elements are also incorporated.

gesture and use of space (proxemics) as frequent semiotic resources used in traditional face-to-face lectures (2016). Crawford Camiciottoli reflects on the use of prosody (paralanguage), especially to stress certain lexical items (2015, 2016). Similarly, in a previous study I analyze pitch lines in a lecturer's speech to discern whether intonation indicates that an utterance is finished or is to be continued (2017). Finally, together with Fortanet-Gómez I have identified head movement and facial expression—particularly eyebrow raising and frowning—as two semiotic choices that lecturers may resort to in order to provide emphasis and engage the audience during the use of organizational metadiscourse (2019). Thus, verbal speech, gesture, gaze, head movement, posture, paralanguage, proxemics and facial expression seem to be present to different degrees among the choices made by the lecturers to convey meaning. Consequently, these are the modes examined in the multimodal analyses in the present study.

Secondly, and with the aim of expediting the analysis process, multimodal transcriptions of the fragment were created. I opted for MMA-Video for a detailed multimodal description of the modes that were selected for the study, and I also described the multimodal successions of actions and contextual information in more general image-based transcriptions by selecting screenshots that aim to accurately represent the complexity of the multimodal ensembles.

The third step in the process consisted in analyzing the transcriptions to explore the results that might be obtained by applying each of the multimodal approaches described in the previous section—MSS, MDA and MIA. Specifically, I looked into how the analysis of the modes employed by the lecturer and the ways in which they are combined may reveal contextual aspects intrinsic to the genre of lectures—MSS. I also explored the special characteristics of the systems of communication themselves—MDA. Next, I looked into the relationship between higher-level actions and how they are intertwined in the discourse—MIA. Finally, I reflected upon the results obtained by applying each of the methods and I looked into the ways in which they may be combined. The analyses, together with a discussion on the applicability of each of the multimodal paradigms, are presented in the next section.

I do not intend to extract any conclusions on the use of multimodal resources in lectures but rather to reflect on the analyses of such resources. Furthermore, it is equally necessary to reflect upon the qualitative potential of multimodal analyses in general. These analyses are usually laborious and require time-consuming, fine-grained transcriptions and processes; therefore, most of the time the results are qualitative and are based on short representative fragments from which much information may be extracted.

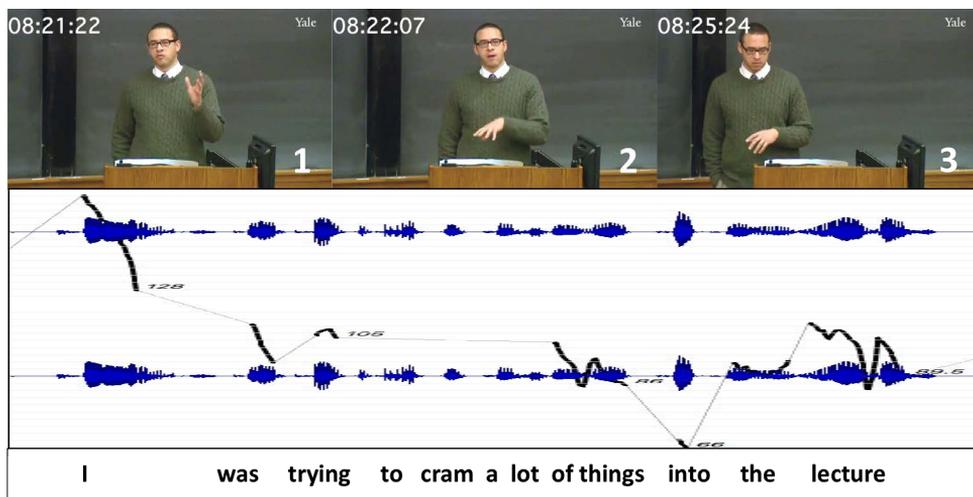
### 3. RESULTS AND DISCUSSION

#### 3.1. Multimodal Social Semiotics (MSS)

As described in the introduction, MSS analyses provide insights into the sign-making choices made by speakers, thus feeding listeners information about the communicative

event. In order to explore what a multimodal analysis may reveal about the teaching practice of a particular lecturer, I looked at a short fragment extracted from the excerpt described in section 2, where the lecturer, after reviewing the contents of his previous class and before introducing the new content, justifies the fact that he included a lot of information in his previous lecture (figure 2).

FIGURE 2. Multimodal transcription of a fragment from the excerpt



When analyzing each of the modes in this fragment individually, several aspects are worth noting. Firstly, the lecturer performs an iconic gesture with his left hand (McNeill 1992) that expresses pictorial content and shows a relationship with the contents of speech. Specifically, the lecturer moves his left hand from an upper vertical position (image 1) to a lower horizontal one (images 2 and 3) as he utters the sentence “I was trying to cram a lot of things into the lecture.” This gesture complements the verbal message, since it visually aims to represent the idea of “putting something in one place”; in this case, including many topics in one lecture. In line with David McNeill’s results (1992), the combination of iconic gestures with the verbal message indicates coexpressiveness and complementarity, since both channels communicate similar semiotic content and complement each other on different planes. In other words, the gesture offers an imagistic representation of the linguistic content. This particular combination of modes has been shown to function as an emphaser of speech content (Pak-Hin Kong et al. 2015).

Another element that stands out in this fragment is the use of gaze. The direction of gaze may provide information about the focus of attention of the speaker (Just and Carpenter 1976). Throughout this snippet, the lecturer alternates the direction of

his gaze between the audience (images 1 and 2) and his notes (image 3). In fact, this alternation is quite consistent throughout the lecture. The analysis of gaze provides information about the importance of the notes for the lecturer: although he does not read through them and tries to maintain his focus on the audience, delivering the class in a conversational manner, he is rather bound by the notes as a structuring device that provides a script for the lecture. It can be argued that the notes are an indispensable element in the lecturing process since they help the speaker keep track of the contents of the session. In figure 2, the lecturer gazes at the audience while explaining that he included a lot of content in the previous lecture (images 1 and 2), which, together with the aforementioned hand gesture, contributes to emphasizing the message and engaging the audience (Manusov and Patterson 2006; Bernad-Mechó and Fortanet-Gómez 2019). However, he turns to his notes at the end of his utterance (image 3) in preparation for the next part of the class, where he expands on a previously introduced topic.

Lastly, the analysis of the paralinguistic elements in this fragment also contributes to creating a new layer of meaning. For example, when examining the pace at which the fragment is uttered, it can be seen that the number of syllables per second is larger than in the neighboring sections of the excerpt. In other words, this sentence is delivered faster than the rest of the speech—7.78 syllables per second compared to the average rate of 5.29 across the entire excerpt. When looking at this paralinguistic feature together with the verbal content of the utterance, and taking into account the context of the sentence, it could be argued that this fragment functions as an apposition to the review of the contents of the previous class. Put differently, the lecturer interrupts the natural flow of his speech in order to make a spontaneous comment on the fact that perhaps there was too much content in the previous session.

In a nutshell, MSS analysis sheds light on the meaning-making choices and the implicatures that are intrinsic to them. By exploring how the different modes are used, the analyst learns about the underlying layers of meaning that are present in communication—the use of gestures for emphasis and the gaze to indicate attention, the relationship of the lecturer with his notes and the pace of speech to indicate the spontaneity of an utterance. Furthermore, this analysis also contributes to the description of the affordances of each mode—the types of meaning that may be conveyed through them.

### 3.2. Multimodal Discourse Analysis (MDA)

In MDA the focus is on the system of choices and the study of the combinations of modes inherent to a given communicative act. Specifically, this approach allows for the examination of the systems of meaning-making as a whole, so that the analyst can obtain quantitative data on the use of semiotic resources by using a multilayer annotation. As stated above, I employed the software MMA-Video for the analysis of

the excerpt. A series of layers were created in accordance with the modes to be analyzed: gestures, gaze, head movement, posture, proxemics and facial expression.

Table 1 shows the percentages of use of each of the semiotic resources in the excerpt. Some interesting facts that can be extracted from these data are, for instance, the frequent use of gestures—35.12% of total time—the amount of time that the lecturer directs his gaze towards his notes—55.62%—and how long the lecturer is leaning towards the lectern—32.14%. The latter two percentages are in line with the results obtained in the MSS analysis, as they reinforce the importance of his notes. Furthermore, this use of notes is in line with a conversational-style lecturer, i.e., lecturers who conduct their classes “in a relatively informal style with a certain amount of interaction with students” (Dudley-Evans 1994, 148), albeit they use their notes for guidance.

TABLE 1. Quantitative data on the use of semiotic resources in the excerpt

System	System of choices	Use over total clip (%)	
Gestures	Gesturing vs. non-gesturing	35.12 / 64.88	
	Gesture type	Iconic	2.38
		Metaphoric	2.38
		Deictic	7.14
		Beat	23.21
Single-handed vs. two-handed gestures	26.78 / 8.34		
Gaze	To the audience	43.20	
	To the notes	55.62	
	To the book	-	
	Other	1.18	
Head movement	Novel	Iconic	0.60
		Deictic	-
		Beat	-
Posture	Standing	Upright	67.86
		Towards the lectern	32.14
	Sitting on the table	Towards the table	-
		Cross-legged	-
Proxemics	Sitting on the table	Stretched	-
		Behind the lectern	100
		On the table	-
		By the table	-
		Moving in front of the audience	-
	In front of the audience	-	

System	System of choices	Use over total clip (%)	
Facial expression	Forehead	Eyebrow raising	2.38
		Frowning	-
	Mouth	Smiling	-
		Lip licking	2.38
		Swallowing	2.38

The possibility of obtaining quantitative data on the use of semiotic resources opens up the potential for analysis. For example, these results could be compared with those from lecturers with different lecturing styles or even with other sections in the same lecture.<sup>4</sup> Such analyses might reveal interesting aspects related to the use of semiotic resources in lectures. The use of MMA-Video also provides information about the specific multimodal ensemble that is occurring in any given instance. Table 2, for example, shows the multimodal ensemble of co-occurring modes that the lecturer deploys while he reviews the contents of his previous lecture. At this point—minute 08:07:00—the lecturer performs a deictic gesture and looks at the audience from an upright position behind the lectern. Deictic gestures are used to point to concrete or abstract entities (McNeill 1992). In this case, the gesture seems to point to the moment in the past to which the lecturer is referring—“the last lecture before the midterm”—thus visually reinforcing the connection that is established verbally. This gesture co-occurs with a multiplicity of semiotic resources that, when working together, contribute to the creation of the full meaning.

TABLE 2. Multimodal ensemble extracted from the excerpt

System	System of choices
Gesture (type)	Deictic
Gesture (1 vs. 2 hands)	1 hand
Gaze	To the audience
Head movement (type)	
Posture	Upright
Proxemics	Lectern
Facial expression	

<sup>4</sup> Elsewhere I provide a full description of different lecturing styles (2018).

Once again, the exploration of these multimodal ensembles allows for their comparison with ensembles from other lecturers or different sections of the same lecture, which can provide information on the use of semiotic resources in academic lectures.

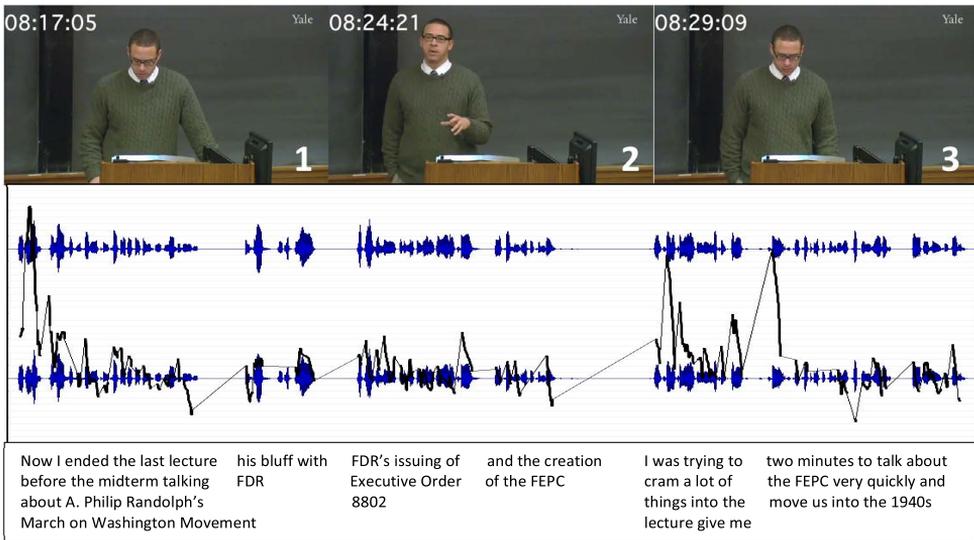
In summary, an MDA approach looks into the general use of semiotic resources and their combinations. Furthermore, the possibility of obtaining quantitative data allows the researcher to go beyond traditional multimodal qualitative studies and explore recurrent patterns or contrasting uses of the semiotic resources across the data. All in all, the study of the systems of meaning-making in MDA analyses brings to the fore critical information that may be fundamental to the genre being examined.

### 3.3. Multimodal Interaction Analysis (MIA)

An MIA approach to the dataset allows for the exploration of how lecturers structure their performances into higher-level actions. Higher-level actions, in turn, may be analyzed in terms of modal density—which modes intervene in the production of each action and to which degree. Besides, the analysis of specific modal densities will reveal which actions are given higher or lower attention/awareness in the process of lecture structuring. This level of attention is indicated in a foreground-background continuum (see section 1.2). Once the higher-level actions are identified and a level of attention is attached to each of the actions, conclusions may be extracted.

Figure 3 below shows the transition in the excerpt from the revision of the previous lecture to the introduction of a content section in the present lecture. Three main higher-level actions can be identified in this fragment: reviewing (the contents of the last lecture), introducing the topic and lecturing. The first higher-level action, reviewing, is realized through high modal density: as the lecturer verbally reviews the topics dealt with in the previous class, he gazes towards his notes and he leans against the lectern (image 1). In this higher-level action the lecturer seems to be quite focused on the notes. From a paralinguistic point of view, the boundaries of the higher-level action are clearly signaled. On the one hand, the first elements in the list of topics from the previous class are marked with a pronounced rising intonation—“A. Philip Randolph’s March on Washington Movement,” “his bluff with FDR” and “FDR’s issuing of Executive Order 8802”—that indicates an incomplete action. On the other hand, the last element—“the creation of the FPC”—is uttered with a falling intonation followed by a pause, which indicates the end of the enumeration. The following sentence—“I was trying to cram a lot of things into the lecture”—appears to be a remnant of the higher-level action of reviewing since it is delivered rapidly before moving into the next higher-level action.

FIGURE 3. Transition of higher-level actions in the excerpt



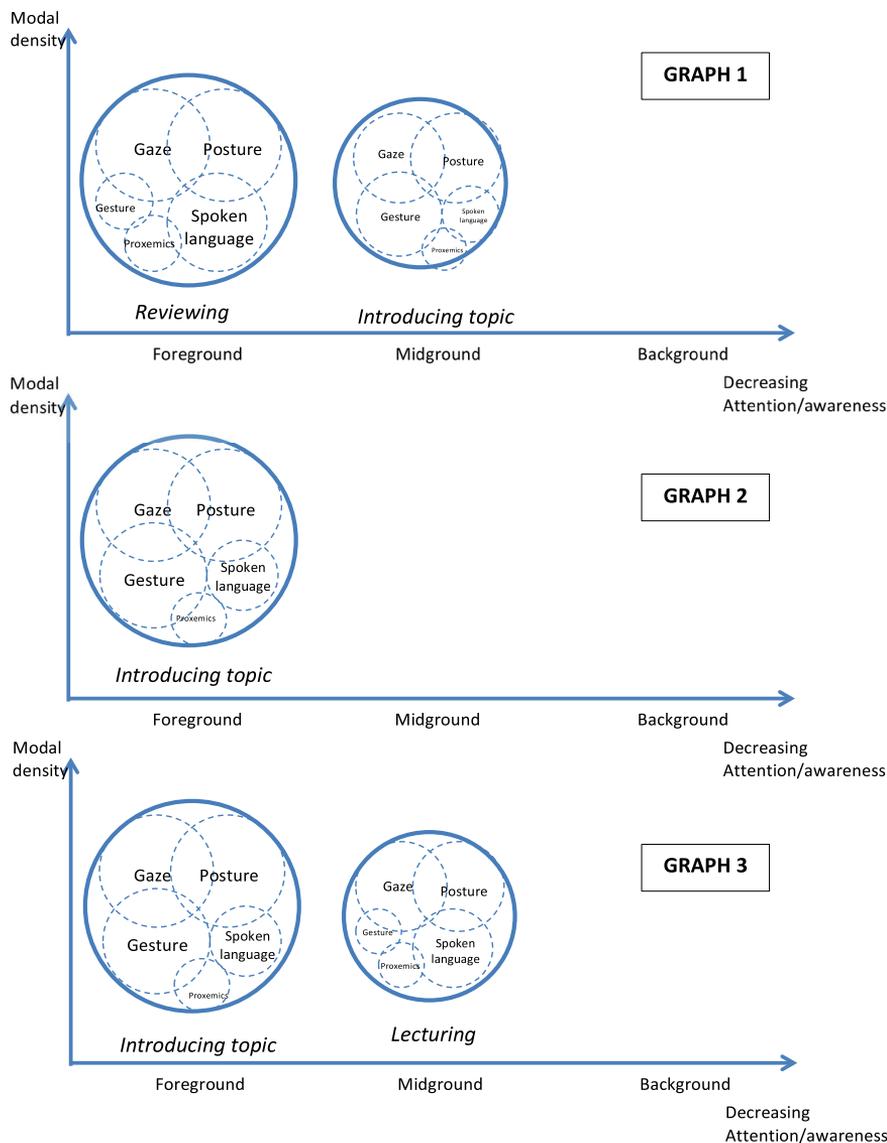
The second higher-level action, introducing the topic, is mainly realized through the lecturer's gaze—directed at the audience—an upright posture and a more spontaneous use of gestures (image 2). In line with the results presented by myself and Fortanet-Gómez (2019), the particular combination of modes employed in this section can be said to aim at engaging the audience. Finally, the higher-level action of lecturing is only realized at the end of the fragment (image 3). This action is accomplished at a low modal density as the lecturer turns his gaze and posture towards the notes in preparation for the next part of the lecture, while devoting gestures and verbal language to the action of introducing a topic.

The modal density and, consequently, the attention paid by the lecturer to each of these actions varies across time. This variation is represented in the foreground-background continuums in figure 4. At the beginning of the action, the lecturer is completely focused on the performance of reviewing, as shown by the modal density. At the end of the enumeration of previous topics, the lecturer pauses for a short period of time. Furthermore, he changes his posture, moving from leaning on the lectern to an upright position. At this point, the higher-level action of introducing a topic is present in the midground while the action of reviewing simultaneously stays in the foreground (graph 1).

This situation changes slightly in the next seconds. While the lecturer utters the sentence “I was trying to cram a lot of things into the lecture”—belonging to the action of reviewing—his nonverbal modes are already all focused on the audience, indicating that the action of introducing a topic is gaining ground. After this sentence, when the

lecturer begins the introduction of the topic verbally—“Give me two minutes to talk about the FEPC very quickly”—he is completely focused on the action of introducing a topic and the action of reviewing is no longer in play (graph 2).

FIGURE 4. Transitions between actions:  
analysis of modal density in the foreground-background continuum



Finally, when the lecturer is finishing the introduction of the topic—“and move us into the 1940s”—he uses a semantic/pragmatic means: his left hand goes back to the lectern. This is a pronounced lower-level action that is not related to the higher-level action of introducing a topic. This semantic/pragmatic means is accompanied by a quick glance at his notes. These elements indicate that a new higher-level action is about to be brought to the fore. Specifically, the lecturer is preparing for the action of lecturing about the FEPC. At this point, the higher-level action of lecturing is in the midground while the higher-level action of introducing a topic remains in the foreground (graph 3).

This type of analysis throws light on how lectures develop as sequences of actions. It also provides an insight into the levels of attention/awareness that a lecturer devotes to each of the actions. From the analysis of attention, the analyst can reflect, for example, on the importance given to any one action, on whether the lecturer is focusing on the audience and trying to engage them in the lecture, etc. Furthermore, comparisons between lecturers and the ways in which they structure higher-level actions can also be carried out.

### 3.4. Combining Methodologies for the Study of Lectures

As I have shown above, the capabilities of each one of the multimodal approaches discussed in this article are abundant and the choice of one approach over another will ultimately depend on the research questions that need to be answered. However, I argue that an integration of these methods is possible and that such a combination provides a more complete view of the elements that make up communication in lectures.

By combining these approaches broader and deeper conclusions may be reached. The process of meaning-making is seen as the result of the choices made by lecturers, which, in turn, may reveal underlying layers of meaning such as the creation of emphasis, the relationship between the lecturer and their notes or their foci of attention. The modes are also seen as entities with the capacity to convey meaning, and such capacity can be explored, delimited and described. From a quantitative point of view, the system of choices can be studied as a whole so as to discern which elements play a more important role in the creation of meaning and which combinations are being employed at any given moment. Finally, lectures can also be examined as sequences of actions, disclosing information about the levels of attention given to each of the actions in the communicative process. Consequently, a combination of methodologies would provide a more comprehensive analysis of the intricacies of multimodal communication in lectures.

## 4. CONCLUSIONS

The main purpose of this article was to offer a methodological reflection on the affordances of three approaches to multimodality—MSS, MDA and MIA. Through a series of test analyses on a short fragment from a lecture on African American history, I have shown how these approaches can be used for the analysis of academic lectures.

MSS offers contextual information about the communicative event and explores how meaning is created as a selection of choices by the speaker from the affordances of each mode. MDA, in contrast, provides information about what semiotic resources are used and how they combine to create meaning through a specific system of choices. Moreover, quantitative results on the relevance of each of the modes can be obtained. Finally, MIA zooms in on the social actor and how lecturers structure their sessions in a series of actions to which they may devote more or less attention. The combination of these analyses contributes to a holistic understanding of lectures that yields reflections on the contexts, the systems of choices and the social actors—which constitutes the main novelty put forward in this article and opens a door to future research in the field.

Specifically, the application of these methodologies to the fragment analyzed in this article has revealed information about the relationship between the lecturer and his audience and between the lecturer and his notes. Moreover, this study has also reflected upon the affordances of the modes employed by the lecturer and how they are utilized in the meaning-making process for various purposes—emphasis, engagement, etc. Furthermore, the quantitative data and the analysis of the system of choices suggest that the lecturer being analyzed adopts a conversational style. Finally, the analysis has shown that the action of introducing a topic is fully focused on the audience and therefore receives more attention, and how this action is intertwined with the previous and the following one.

Although the applicability of these approaches is undoubtedly useful for the multimodal analysis of lectures, there are, indeed, certain limitations to be considered. First and foremost, multimodal analyses are laborious and time consuming. Therefore, only small fragments can realistically be analyzed. These analyses, however, are fine-grained and produce complex qualitative results that open up new possibilities for expanding research. In this sense, although quantitative analyses are possible, they are still based on short fragments of video. Consequently, careful preplanning is needed before carrying out multimodal analyses. The analyst needs to make sure that the examples that are being analyzed are representative of the data.

To conclude, the possibilities for further research that arise from this methodological proposal are worth mentioning, even if briefly. Although the results shown here were only meant to serve as examples of how the three approaches to multimodality can be used, the study of lectures can easily be expanded. For example, the use of semiotic resources could be compared between lecturers with different lecturing styles or across different sections within the same lecture. Similarly, the particular structures of higher-level actions might also be compared in search of patterns and similarities. Finally, from a methodological perspective, the specific combination of approaches proposed in this article could be reformulated so as to create a series of steps that could be followed for a complete analysis of lectures.

## WORKS CITED

- ADAMI, Elisabetta. 2017. "Multimodality." In García, Flores and Spotti 2017, 451-72.
- ARCHER, Arlene and Esther Odilia Breuer, eds. 2016. *Multimodality in Higher Education*. Leiden: Brill.
- BERNAD-MECHÓ, Edgar. 2017. "Metadiscourse and Topic Introductions in an Academic Lecture: A Multimodal Insight." *Multimodal Communication* 6 (1): 39-60.
- . 2018. "A Multimodal Approach to Metadiscourse as an Organizational Tool in Lectures." PhD diss., University Jaume I.
- BERNAD-MECHÓ, Edgar and Inmaculada Fortanet-Gómez. 2019. "Organizational Metadiscourse across Lecturing Styles: Engagement beyond Language." In Sancho-Guinda 2019, 321-40.
- BIBER, Douglas, Susan Conrad and Viviana Cortés. 2004. "If You Look at... : Lexical Bundles in University Teaching and Textbooks." *Applied Linguistics* 25 (3): 371-405.
- BONSIGNORI, Veronica and Belinda Crawford Camiciottoli, eds. 2016. *Multimodality across Communicative Settings, Discourse Domains and Genres*. Newcastle upon Tyne: Cambridge Scholars.
- BUCK, Gary. 2001. *Assessing Listening*. Cambridge: Cambridge UP.
- CALZOLARI, Nicoletta et al., eds. 2006. *Proceedings of LRforthEC 2006, Fifth International Conference on Language Resources and Evaluation*. Genoa: European Language Resources Association.
- CRAWFORD CAMICIOTTOLI, Belinda. 2015. "Elaborating Explanations during OpenCourseWare Humanities Lectures: The Interplay of Verbal and Nonverbal Strategies." In Crawford Camiciottoli and Fortanet-Gómez 2015, 144-70.
- . 2016. "A Multimodal Analysis of Interaction in Academic Lectures: A Case Study." In Bonsignori and Crawford Camiciottoli 2016, 65-92.
- CRAWFORD CAMICIOTTOLI, Belinda and Inmaculada Fortanet-Gómez, eds. 2015. *Multimodal Analysis in Academic Settings: From Research to Teaching*. London and New York: Routledge.
- DEPPERMAN, Arnulf, ed. 2013. "Conversation Analytic Studies of Multimodal Interaction." Special issue, *Journal of Pragmatics* 46 (1).
- DUDLEY-EVANS, Tony. 1994. "Variations in the Discourse Patterns Favoured by Different Disciplines and Their Pedagogical Implications." In Flowerdew 1994, 146-58.
- FLOWERDEW, John, ed. 1994. *Academic Listening: Research Perspectives*. Cambridge: Cambridge UP.
- FLOWERDEW, John and Lindsay Miller. 1992. "Students' Perceptions, Problems and Strategies in L2 Lectures." *RELC Journal* 23 (12): 60-80.
- FORCEVILLE, Charles J. and Eduardo Urios-Aparisi, eds. 2009. *Multimodal Metaphor*. Berlin: De Gruyter.
- FORTANET-GÓMEZ, Inmaculada. 2004. "The Use of 'We' in University Lectures: Reference and Function." *Journal of English for Specific Purposes* 23: 45-66.

- FORTANET-GÓMEZ, Inmaculada and María Noelia Ruiz-Madrid. 2014. "Multimodality for Comprehensive Communication in the Classroom: Questions in Guest Lectures." *Ibérica* 28: 203-24.
- FOX, Janna and Natasha Artemeva. 2013. "The Cinematic Art of Teaching University Mathematics: Chalk Talk as Embodied Practice." *Multimodal Communication* 1 (1): 83-103.
- GARCÍA, Ofelia, Nelson Flores and Massimiliano Spotti, eds. 2017. *The Oxford Handbook of Language and Society*. Oxford: Oxford UP.
- GEE, Paul and Michael Handford, eds. 2012. *The Routledge Handbook of Discourse Analysis*. London and New York: Routledge.
- GUNEL, Murat, Brian Hand and Sevket Gunduz. 2006. "Comparing Student Understanding of Quantum Physics When Embedding Multimodal Representations into Two Different Writing Formats: Presentation Format versus Summary Report Format." *Science Education* 90 (6): 1092-112.
- HALLIDAY, M. A. K. 1978. *Language as a Social Semiotic: The Social Interpretation of Language and Meaning*. London: Edward Arnold.
- . 1985. *An Introduction to Functional Grammar*. London: Edward Arnold.
- HYLAND, Ken and Brian Paltridge, eds. 2011. *Companion to Discourse*. London: Continuum.
- IEDEMA, Rick. 2003. "Multimodality, Resemiotization: Extending the Analysis of Discourse as Multi-Semiotic Practice." *Visual Communication* 2 (1): 29-57.
- JEWITT, Carey. 2014a. "Different Approaches to Multimodality." In Jewitt 2014b, 28-39.
- , ed. 2014b. *The Routledge Handbook of Multimodal Analysis*. London and New York: Routledge.
- JUST, Marcel A. and Patricia A Carpenter. 1976. "Eye Fixations and Cognitive Processes." *Cognitive Psychology* 8 (4): 441-80.
- KRESS, Gunther. 2010. *Multimodality: A Social Semiotic Approach to Contemporary Communication*. London and New York: Routledge.
- . 2012. "Multimodal Discourse Analysis." In Gee and Handford 2012, 35-50.
- KRESS, Gunther and Theo van Leeuwen. 1996. *Reading Images: The Grammar of Visual Design*. London and New York: Routledge.
- . 2001. *Multimodal Discourse: The Modes and Media of Contemporary Communication*. London: Arnold.
- MANUSOV, Valerie L. and Miles L. Patterson. 2006. *The SAGE Handbook of Nonverbal Communication*. London: Sage.
- MCNEILL, David. 1992. *Hand and Mind: What Gestures Reveal about Thought*. Chicago, IL: U of Chicago P.
- MORELL, Teresa. 2018. "Multimodal Competence and Effective Interactive Lecturing." *System* 77: 70-79.

- NORRIS, Sigrid. 2004. *Analyzing Multimodal Interaction. A Methodological Framework*. London and New York: Routledge.
- O'HALLORAN, Kay, ed. 2004. *Multimodal Discourse Analysis: Systemic Functional Perspectives*. London: Continuum.
- . 2011. "Multimodal Discourse Analysis." In Hyland and Paltridge 2011, 120-37.
- O'HALLORAN, Kay et al. 2012. "Interactive Software for Multimodal Analysis." *Visual Communication* 11 (3): 363-81.
- PAK-HIN KONG, Anthony et al. 2015. "A Coding System with Independent Annotations of Gesture Forms and Functions during Verbal Communication: Development of a Database of Speech and GESTure (DoSaGE)." *Journal of Nonverbal Behavior* 39 (1): 93-111.
- QUEROL-JULIÁN, Mercedes and Inmaculada Fortanet-Gómez. 2012. "Multimodal Evaluation in Academic Discussion Sessions: How Do Presenters Act and React?" *English for Specific Purposes* 31: 271-83.
- ROUNDS, Patricia L. 1987. "Characterizing Successful Classroom Discourse for NNS Teaching Assistant Training." *Tesol Quarterly* 21 (4): 643-69.
- SANCHO-GUINDA, Carmen, ed. 2019. *Engagement in Professional Genres: Deference and Disclosure*. Amsterdam and Philadelphia: John Benjamins.
- SCOLLON, Ron. 2001. *Mediated Discourse: The Nexus of Practice*. London: Longman.
- TAN, Sabine, Kay O'Halloran and Peter Wignell. 2016. "Multimodal Research: Addressing the Complexity of Multimodal Environments and the Challenges for CALL." *ReCALL* 28 (3): 253-73.
- THESEN, Lucia. 2016. "The Past in the Present: Modes, Gaze and Changing Communicative Practices in Lectures." In Archer and Breuer 2016, 31-52.
- VAN LEEUWEN, Theo. 2005. *Introducing Social Semiotics*. London and New York: Routledge.
- . 2008. *Discourse and Practice: New Tools for Critical Discourse Analysis*. Oxford: Oxford UP.
- WITTENBURG, Peter et al. 2006. "ELAN: A Professional Framework for Multimodality Research." In Calzolari et al. 2006, 1556-59.
- YALE UNIVERSITY. 2021. Yale University's OpenCourseWare. [Accessed February 9, 2021].

Received 1 May 2020

Revised version accepted 7 July 2020

Edgar Bernad-Mechó holds a PhD in Applied Linguistics. He is a member of the "Group for Research on Academic and Professional Discourse (GRAPE)" at Universitat Jaume I. His research interests focus on the study of academic English in lectures and multimodality studies and methodologies. He is currently working on the organization of lectures from the perspective of metadiscourse and the ensemble of semiotic resources by lecturers.