

# Revisiting persuasion in oral academic and professional genres: Towards a methodological framework for Multimodal Discourse Analysis of research dissemination talks

**Julia Valeiras-Jurado, Noelia Ruiz-Madrid and Geert Jacobs**

Ghent University (Belgium) and Universitat Jaume I (Spain)

julia.valeirasjurado@ugent.be & madrid@ang.uji.es & geert.jacobs@ugent.be

## Abstract

Previous work on oral genres (Kress & Van Leeuwen, 2001; Kress, 2010; Bateman, 2011) as well as on persuasion (O’Keefe, 2002; Perloff, 2003; Poggi & Pelachaud, 2008) has indicated that effective persuasive oral communication depends heavily on the use of a wide range of different semiotic modes including words, gestures and intonation. However, little attention has been paid so far to how speakers convey their communicative intentions orchestrating different modes into a coherent multimodal ensemble (Kress, 2010). In this paper we propose a methodological framework for Multimodal Discourse Analysis (MDA) of persuasion in oral academic and professional genres. Drawing on previous studies on persuasion (Fuertes-Olivera et al., 2001; O’Keefe, 2002; Perloff, 2003; Virtanen & Halmari, 2005; Dafouz-Milne, 2008), our framework combines earlier proposals for MDA (Querol-Julián, 2011; Querol-Julián & Fortanet-Gómez, 2014) with an ethnographic perspective (Rubin & Rubin, 1995). Our study focuses specifically on the analysis of persuasive strategies used in dissemination talks. The proposed MDA caters for the following modes: words, intonation, head movements and gestures. Preliminary findings hint at a relation between persuasion and so-called modal density (Norris, 2004). Finally, we propose a tentative taxonomy of persuasive strategies and how they are realised multimodally.

**Keywords:** academic and professional discourse, multimodality, MDA, ethnography, persuasion.

## Resumen

### *Persuasión en géneros académicos y profesionales. Propuesta de un marco metodológico para el análisis multimodal de charlas divulgativas*

Tal y como demuestran investigaciones previas sobre géneros orales (Kress y Van Leeuwen, 2001; Kress, 2010; Bateman, 2011) y persuasión (O'Keefe, 2002; Perloff, 2003; Poggi y Pelachaud, 2008), el lenguaje oral persuasivo requiere del uso de varios recursos semióticos para ser efectivo, entre los que se encuentran las palabras, los gestos y la entonación. Sin embargo hasta ahora no se ha dedicado demasiada atención al estudio de cómo los hablantes orquestan esta variedad de recursos semióticos inherentes en el lenguaje oral para formar un mensaje coherente (Kress, 2010). En este artículo proponemos un marco metodológico específico para el análisis multimodal de la persuasión en géneros orales académicos y profesionales. Dicho marco metodológico se basa en estudios previos sobre persuasión, (Fuertes-Olivera et al., 2001; O'Keefe, 2002; Perloff, 2003; Virtanen y Halmari, 2005; Dafouz-Milne, 2008), propuestas metodológicas para el estudio multimodal de género (Querol-Julián, 2011; Querol-Julián y Fortanet-Gómez, 2014) y métodos etnográficos (Rubin y Rubin, 1995). El estudio que aquí presentamos se centra en el análisis de las estrategias persuasivas utilizadas en charlas divulgativas. El análisis multimodal implica los siguientes recursos semióticos: palabras, entonación, movimientos de cabeza y gestos. Los resultados preliminares apuntan a una relación entre persuasión y densidad modal o combinación de recursos semióticos (Norris, 2004). Por último, proponemos una taxonomía provisional de estrategias persuasivas llevadas a cabo de forma multimodal.

**Palabras clave:** discurso académico y profesional, multimodalidad, MDA, etnografía, persuasión.

## 1. Introduction

Persuasion is a defining trait of many oral genres that share a broad communicative purpose of convincing the audience of the value of a product, service or idea. This is the case of oral presentations used in both academic and business contexts (Carter-Thomas & Rowley-Jolivet, 2003; Rowley-Jolivet & Carter-Thomas, 2005; Bamford, 2007, 2008). On the other hand, the literature suggests that in these oral genres (and in fact in most communicative situations) speakers use a variety of semiotic modes to express meaning (Kress & Van Leeuwen, 2001; Kress, 2010, 2012, 2014; Bateman, 2011; Jewitt, 2013; O'Halloran & Smith, 2011). Research on multimodality pays particular attention to how each mode interacts with

others and how they are orchestrated in a specific context to produce meaning. When more than one mode is used to produce a specific meaning it is said these modes “ensemble”, and this meaning is inferred based on the interrelation between them. The communicator orchestrates multimodal ensembles, based on designs, where each mode has a function (Kress, 2010) and “each mode is partial in relation to the whole meaning” (Jewitt & Kress, 2003: 3). In other words, in many communicative situations words constitute just one among many modes which are of equal importance in the process of meaning making.

Returning to the centrality of persuasion, it can be hypothesized that all the modes that speakers use in academic and business presentations can have an effect on the broad communicative purpose of these genres, i.e. persuasion. Indeed, previous research on persuasion has highlighted how different semiotic modes can contribute to it (Sparks et al., 1998; O’Keefe, 2002; Perloff, 2003; Poggi & Pelachaud, 2008). This suggests that the orchestration of different modes can be especially decisive in achieving effective persuasive oral communication.

For this reason Multimodal Discourse Analysis (MDA) (Querol-Julián, 2011; Querol-Julián & Fortanet-Gómez, 2014) is particularly suitable to the study of persuasion in these genres, due to its potential to explain the overall communicative effect of the multimodal ensemble rather than the individual contribution of each mode. Likewise, ethnographic methods such as observation and interviews can prove to be very valuable tools to interpret and complement the data obtained from the MDA analysis.

According to authors such as Fuertes-Olivera et al. (2001), O’Keefe (2002), Perloff (2003), Virtanen & Halmari (2005) and Dafouz-Milne (2008), persuasive messages tend to be more effective when:

- (i) Speakers have credibility and the audience can identify with them: addressees are more prone to be persuaded by speakers they trust.
- (ii) The message is easy to understand: when the effort to understand the message is too big, addressees will tend to stop trying and eventually they will stop paying attention.
- (iii) The content is relevant for the audience: addressees are less likely to be persuaded by messages that do not appeal to them or they feel as not applicable to them.

- (iv) The message is made memorable: addressees are more likely to be convinced by a message that they can easily remember without making a conscious effort.
- (v) The message is innovative and surprising: persuasion becomes more effective the more unexpected it is. Forewarning is a persuasion killer, because it activates the addressee's mind and stirs potential counterarguments in advance.
- (vi) The message is perceived as not imposed, but inferred: addressees should feel they have reached their own conclusions, even if they have been guided towards them.

As pointed out by Valeiras-Jurado (2015) and Valeiras-Jurado & Ruiz-Madrid (2015), in persuasive oral genres some strategies or communicative techniques such as emphasis, evaluation (Martin & White, 2005; Querol-Julián, 2011), projection of context of interaction (Brazil, 1997), anticipation and control of responses, etc. are likely to be used for persuasive purposes, because they help provide the message with the aforementioned characteristics. In addition, they are prone to be performed through a variety of semiotic modes, which makes them particularly interesting for an MDA approach to persuasion. Some examples are the use of intonation to present parts of the message as already agreed upon as opposed to open to discussion (Brazil, 1997) and gestures used to discourage potential counterclaims (Kendon, 2004).

The objective of this paper is twofold. Firstly, we present a methodological framework that can facilitate a MDA approach to persuasion in oral genres. Then we test the methodology by applying it to the genre of dissemination talks. This testing of the methodological framework has yielded a tentative taxonomy of persuasive strategies that are realised as multimodal ensembles.

The MDA approach to the analysis of persuasion proposed here tries to elucidate how speakers convey their communicative intentions orchestrating different modes into a coherent multimodal ensemble. As already mentioned, we consider this process to be something more than the mere accumulation of each separate mode. For this reason, we do not prioritize any of the modes, considering the potential contribution of each of them as equally important and frequently mutually interdependent.

This paper is structured in six sections. First, the framework for our analysis of persuasion is presented. Second, we discuss the added value of using combined methodologies (i.e. video-based multimodal analysis and

ethnography). Then, the specialised software used for the purpose of the study and the data gathering process is described. After that we present the persuasive strategies identified in our analysis and the final section offers our discussion of results and conclusions.

## 2. A framework for a multimodal analysis of persuasion

The analysis of persuasion proposed in this paper is framed as a data-driven and cyclical process that draws from three sources.

- (a) input from previous literature, including:
  - previous studies on persuasion (Fuertes-Olivera et al., 2001; O’Keefe, 2002; Perloff, 2003; Virtanen & Halmari, 2005; Dafouz-Milne, 2008)
  - methodological proposals for a MDA approach to genre analysis (Querol-Julián, 2011; Querol-Julián & Fortanet-Gómez, 2014)
  - work on kinesics (McNeill, 1992; McClave, 2000; Kendon, 2002; Kendon, 2004)
  - work on intonation (Brazil, 1997; Chen, 2002; Pickering et al., 2012)
  - previous research on professional and academic discourse (Lakoff, 1982; Carter, 1997; Hyland, 1998, 2009)
- (b) multimodal data (results from the multimodal analysis), and
- (c) ethnographic data (results from interviews and observation sheets)

In other words, we triangulate what we observe in the videos (i.e. the results of multimodal analysis) with what we learn from previous literature, and with the speakers’ own observations and reported intentions. In this way, in an iterative process of literature reviewing, multimodal analysis and ethnographic analysis, we obtain data that enables us to propose a list of persuasive strategies realised multimodally (see section 6). The data retrieved from the interviews with the speakers help us to support our tentative conclusions or reject them, refining our list of persuasive strategies, which becomes the basis for our analysis of persuasion. Figure 1 represents the framework for the analysis.

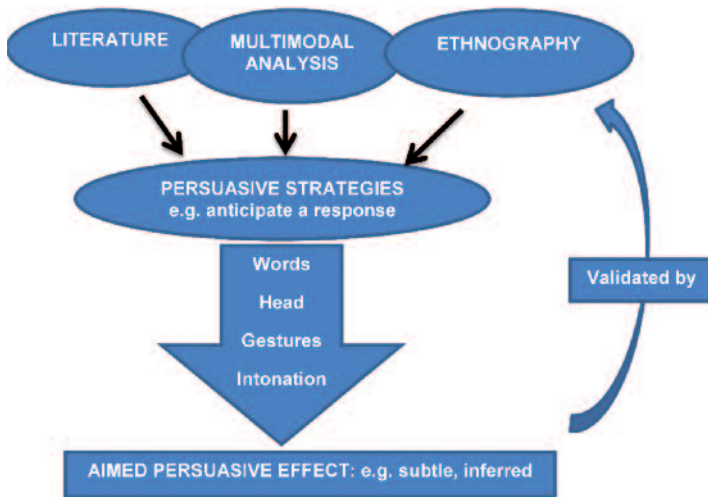


Figure 1. Multimodal analysis framework.

A specific example may help clarify this cyclic process. Input from literature informs us that anticipating responses of interlocutors can contribute to persuasion (O’Keefe, 2002; Perloff, 2003). The multimodal analysis further reveals that speakers seem to anticipate their audience’s reactions combining words, intonation, gestures and/or head movements. With this information we hypothesize that “anticipation of responses” can be a multimodal persuasive strategy. The interview with the speakers helps us confirm that their intention was to prevent a potential counterclaim to their messages. Further discussion with speakers leads to the joint conclusion that the aimed effect is to be subtle in order to prevent pushback and maximize chances for the message to be accepted by the audience.

We believe that one of the main advantages of the analytical framework we suggest is that it enables us to study how a particular aspect of the message (in this case persuasion) is jointly expressed through all the semiotic modes included in the analysis, instead of starting from a particular modal realisation and then looking for instances of co-expression in other modes. In this way, we avoid prioritising any mode, as mentioned in the previous section. Our MDA analysis focuses on four modes in particular that are almost always present in any instance of oral discourse: words, intonation, gestures and head movements. However, at some points in the analysis it has been necessary to refer to other modes not initially included (e.g. gaze) in order to interpret the four modes that were the focus of our analysis.

Therefore, occasional references to gaze or facial expression (sometimes provided by the speakers during the interviews) can be found in the discussion of the results to support the interpretation of the modes being analysed.

### **3. The use of combined methods in the multimodal analysis of persuasion**

As pointed out by Jewitt (2012) video-based analysis provides a highly detailed material record of the communicative process, but it is limited in certain aspects, since the final result is always affected by decisions such as what and how to record, and the potential interference of the camera in naturally occurring data. Video-based research can also lead to a focus on minute detail that makes it difficult to contextualise results within the global picture (Jewitt, 2012). Therefore, video data needs to be cautiously interpreted (Schindler, 2009) and the results obtained complemented with data from other methodological approaches (e.g. ethnographic analyses).

Ethnography as a methodological approach presents a remarkable advantage for the study of persuasion: it makes it possible to gather first-hand information about the communicative situation and about the communicative intentions of the speakers, which is crucial to interpret the data in a more objective way and cross-check initial hypotheses. It also presents some limitations, such as access and reliability of data, obtrusiveness of observations and the high dependence on the quality of the interview design (Rubin & Rubin, 1995).

Yet, we consider that the use of both methods (i.e., video-based multimodal analysis and ethnography) in combination can help to overcome the limitations of each methodological approach used separately (Jewitt, 2012) and eventually help us to understand how the speaker is shaping the genre in real time, using a variety of semiotic modes.

### **4. The use of specialised software in the multimodal analysis of persuasion**

An MDA approach requires the use of different specialised software packages to look into the data. For the present study, we have employed

PRAAT<sup>1</sup> and ELAN<sup>2</sup>. The former helped us in the analysis of intonation as a mode. It is a tool for phonetic analysis that allows obtaining accurate measuring of pitch and intensity. The later, ELAN, allows us to analyse gestures, head movements and speech by means of transcribing and annotating audio and video files. Transcriptions and annotations are organised on layers (or tiers in the program nomenclature). Since these tiers are time-aligned, it is possible to determine what is happening in each mode at a specific moment in time in the presentation. In order to encode our information we have followed Kipp's (2014) recommendations. He insists that our encoding schemes must be consistent (i.e. data is set with same structure, for example tier names), valid (i.e. annotations should correspond to acceptable “variables”), efficient (i.e. annotations should be easy to use and understand); and reliable (i.e. annotations must clearly refer to one specific aspect). For the present study we have used the tier structure shown in Figure 2 below:

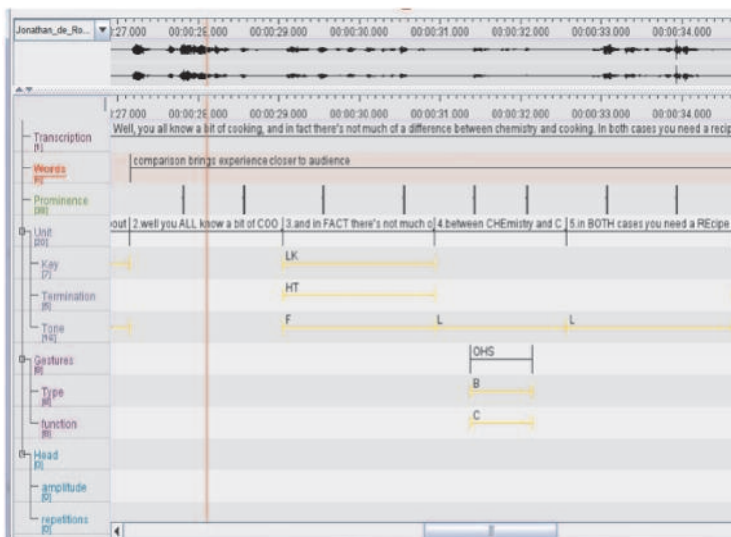


Figure 2. ELAN tiers encoding multimodal transcriptions and annotations.

As shown in Figure 2, tiers can encode different transcriptions and annotations from the modes involved in an MDA approach. For the present study we have used the following tiers.

- (a) “Transcription” tier, including the orthographic transcription.



- (b) “Words” tier, with annotations about grammatical, lexical and stylistic devices such as evaluative language, three-part lists, symmetrical syntactic structures, inclusive pronouns, examples, comparisons, narratives, among others (Lakoff, 1982; Carter, 1997; Hyland, 1998, 2009) that can have an effect on persuasion, because they help provide the message with characteristics that according to literature make the text persuasive (see Introduction).
- (c) Prosodic transcription tiers (“Prominence”, “Unit”, and three dependent tiers: “Tone”, “Key” and “Termination”), including a Discourse Intonation (DI) transcription (Brazil, 1997). This approach to intonation was adopted for its potential to account for speakers’ moment-to moment decisions about their communicative intentions. The tier labelled “Prominence” is a tier imported from PRAAT of the type “point tier” according to the programme’s nomenclature, which means that a vertical line marks the time where the prominence occurs. In the tier named “Unit” the following standard DI conventions are used: upper case letters for prominent syllables and lower-case letters for non-prominent syllables. Tone, key and termination are indicated in three dependent tiers according to the transcription conventions shown in Table 1:

Tone	Key	Termination
F: fall	HK: high key	HT: high termination
L: level	MK: mid key	MT: mid termination
R: rise	LK: low key	LT: low termination

Table 1. Transcription conventions for tone, key and termination.

In the present study combined tones (fall-rise and rise-fall) have not been distinguished. The reason for this research decision is that they mainly convey information about power relationships and in particular dominance, which in the case of the presentations analysed tend to remain rather stable throughout. It was considered that the potential added value of this distinction did not justify the analysis cost at this stage

- (d) “Gestures” tier, indicating gesture family (Kendon, 2004). This tier hosts two other dependent tiers (i.e. their division into slots and their input is linked with and determined by the tier they depend on, in this case “Gestures”), with data concerning the type and

function of the gesture. The classification system draws from McNeill (1992), Bavelas et al. (1995), Kendon (2004) and Querol-Julián (2011), and it is shown in Table 2:

Gesture family:	Gesture type:	Gesture function:
R: ring (Kendon, 2004)	I: iconic (represent concrete objects and events) (McNeill, 1992)	R: referential (represent some aspect of the content) (Kendon, 2004)
G: grappollo (finger bunch) (Kendon, 2004)	M: metaphoric (represent abstract ideas) (McNeill, 1992)	I: interpersonal (regulate interaction) (Bavelas et al., 1995)
OHP: open hand prone (palms down) (Kendon, 2004)	B: beat (repetitive gestures that usually mark the discourse flow) (Kendon, 2004; Bavelas et al., 1995)	P: pragmatic (show attitude or perlocutionary meaning) (Kendon, 2004)
OHS: open hand supine (palms up) (Kendon, 2004)	D: deictic (point to something) (Kendon, 2004)	C: cohesive (link parts of discourse ) (McNeill, 1992; Querol, 2011)
CF: close fist		

Table 2. Transcription and annotation conventions for gestures.

We find the gesture families proposed by Kendon (2004) useful for the interpretation of the meaning intended by particular gestures found in our sample (e.g. OHP meaning stopping or negating something). These interpretations were nonetheless contrasted and complemented with the speakers’ opinions. In addition, in order to account for all the gestures included in the analysis, the addition of an extra category, i.e. the close fist (CF) was needed.

- (e) “Head” tier, indicating the kinesic characteristics (shake, nod, lateral movement), and two dependent tiers with information about amplitude and repetitions (McClave, 2000; Kendon, 2002).

The tiers “words” and “gestures-function” contain annotations, since they provide additional information about the mode and its potential contribution to persuasion, while the rest of the tiers describe the modes and can be regarded transcriptions.

It will be noted that the transcription and annotation of gestures and intonation is highly systematic and follows pre-defined conventions that have been incorporated in the software through the use of controlled vocabularies, which limit the input possibilities in specific tiers through a predefined list of options from which the analyst selects. The annotation of

words, on the other hand, is corpus-driven rather than pre-defined. The rationale behind this methodological decision is a need to keep the focus on the multimodal ensemble. The authors took into account the input from previous literature regarding persuasive strategies realised through words (i.e. lexical, grammatical and stylistic devices), and some of these strategies were indeed identified in the sample (e.g. use of direct speech or evaluative lexicon). However, we avoided having a pre-defined list as the starting point of our analysis, because this would very likely lead to prioritizing words as the main mode and would limit the analysis of the other modes to finding instances of co-expression. On the other hand, the interviews prove that speakers are generally very capable of accounting for their use of words, but find it harder to explain their use of intonation and body movements because they are less conscious of them. Thus, the use of a systematic classification system was necessary to facilitate the identification of particular uses of intonation, gesture and head movements and discuss them during the interviews, but it was not necessary in the case of words.

## **5. Data gathering: Observation sheets recordings and interviews**

The methodology described in this paper has been applied to study persuasion in five dissemination talks, with a specific focus on the following modes: words (Lakoff, 1982; Carter, 1997; Hyland, 1998, 2009), intonation (Brazil, 1997), head movements (McClave, 2000; Kendon, 2002) and gestures (Kendon, 2004). The aim is to demonstrate the validity and reliability of such an approach to study the nature of persuasion in oral presentations in the academic and professional fields. The rest of this section will be devoted to describe the data gathering process.

The event where the presentations were recorded was an independent university TED contest in which scientists in different fields disseminate their activity to a lay audience. We obtained the specific permission of the speakers for using both their video recordings and interviews for the purpose of the present study. Table 3 below summarizes the information of each presentation and speaker in the corpus:

<b>1. With adaptive systems, we can make the best of our differences</b>					
Gender	English proficiency	Expertise	Delivery	Support/ devices	Duration
Female	High	High	Explain	Lapel mic	00:11:07
<b>2. Snow White's smart textiles twist</b>					
Gender	English proficiency	Expertise	Delivery	Support/ devices	Duration
Female	High	High	Explain	PPT Lapel mic Pointer	00:08:59
<b>3. Why does mathematics count?</b>					
Gender	English proficiency	Expertise	Delivery	Support/ devices	Duration
Male	High	High	Explain	Prezzi Lapel mic Pointer	00:07:56
<b>4. Open transport data</b>					
Gender	English proficiency	Expertise	Delivery	Support/ devices	Duration
Male	High	High	Explain	PPT Lapel mic Pointer	00:10:17
<b>5. Microwave chemistry: time is money</b>					
Gender	English proficiency	Expertise	Delivery	Support/ devices	Duration
Male	High	High	Explain	PPT Lapel mic Pointer	00:07:15

Table 3. The corpus of dissemination talks.

In each of the five presentations, two excerpts were further selected for fine-grained multimodal analysis. These excerpts correspond to moments in the presentation which are particularly rich in terms of persuasive efforts from the speakers. Subsequent analysis also proved that they are also particularly rich in modal density (Norris, 2004) in relation to the rest of the presentation, therefore they are likely to illustrate persuasion realised through different modes simultaneously. This approach was adopted to avoid prioritizing any semiotic mode in particular, and has proved useful to keep the focus on the multimodal ensemble as a whole and the way different modes interact to encode a persuasive message. The identification and selection of these excerpts was what Goldman et al. (2007) would call an

inductive approach to video data. The selection was also aided by ethnographic interviews, as it will be detailed at the end of this section. Table 4 provides an overview of the excerpts selected for multimodal analysis, which add up to 4.9 minutes:

<b>1. With adaptive systems, we can make the best of our differences</b>			
	<b>Begin</b>	<b>End</b>	<b>Duration</b>
1.1	00:03:13	00:04:04	00:00:51
1.2	00:04:04	00:04:52	00:00:47
<b>2. Snow White's smart textiles twist</b>			
	<b>Begin</b>	<b>End</b>	<b>Duration</b>
2.1	00:00:00	00:00:24	00:00:24
2.2	00:04:40	00:04:50	00:00:10
<b>3. Why does mathematics count?</b>			
	<b>Begin</b>	<b>End</b>	<b>Duration</b>
3.1	00:01:45	00:1:52	00:00:07
3.2	00:05:03	00:05:08	00:00:05
<b>4. Open transport data</b>			
	<b>Begin</b>	<b>End</b>	<b>Duration</b>
4.1	00:00:33	00:01:03	00:00:30
4.2	00:03:13	00:03:53	00:00:40
<b>5. Microwave chemistry: time is money</b>			
	<b>Begin</b>	<b>End</b>	<b>Duration</b>
5.1	00:00:00	00:00:24	00:00:24
5.2	00:00:24	00:01:17	00:00:53

Table 4. Excerpts selected for MDA.

The events were recorded using only one camera focused on the speaker. The camera was mainly static and only moved if the speaker moved. No close ups of the head, body or the room (e.g. screen behind speaker) were used. These decisions were made on the basis of the research questions that drive the analysis: our interest is mainly in the persuasive activity of the speakers, rather than on the uptake by the audience.

During the recording, we used observation sheets to gather contextual information about the event, the presentation and the speaker. These sheets included a pre-defined open list of items including the event, the genre (type of presentation), the speaker (bio-data), the room (its physical configuration), the audience (size and type), resources (e.g. microphone,

pointer, etc.), and speakers' performance (posture and style of presentation such as "reading from a script" or "improvising").

Shortly after each presentation (at the end of the event, and on the same day) face to face semi-structured interviews were carried out with the speakers in order to obtain crucial information about their previous knowledge and assumptions about the event and their preparation for the presentation, which in turn helps us interpret their use of persuasion. These interviews probed into aspects such as:

- (i) the speakers' motivation to participate in the event,
- (ii) what they knew about the event beforehand (e.g. size and type of audience, topics expected, types of presentations expected, etc.)
- (iii) the way they prepared for the presentation,
- (iv) their main goal in their presentation,
- (v) their satisfaction with their performance,
- (vi) their opinion of other presentations as members of the audience.

We considered that the reason why speakers had decided to participate in the event (i.e. their motivations) would ultimately determine their main goal in their presentations and shape their main message, something crucial to fully understand the communicative intentions lying behind their use of modes. Items ii) and iii) were also important to fully understand how speakers were using the modes, and in particular to determine how much of this was unconscious or spontaneous and how much the result of practice or explicit training. Item v) was a prompt for self-assessment, and was intended to put our own interpretations in perspective and help us focus on particularly persuasive moments. Finally, the last item tries to include the audience uptake in the analysis, although admittedly in a very limited way.

The ethnographic information provided by the observation sheets and the interviews proved crucial, first in the selection of the data to be analysed (i.e. information about what speakers considered their main goal helped select the excerpts for fine-grained multimodal analysis), and later on in its interpretation, since the data provided by the interviews (particularly what speakers knew about the event in advance and how they prepared accordingly) helped interpret the speakers' intended communicative effect and their use of persuasion.

Ethnographic interviews were also used at a later stage to discuss results with speakers and cross-check interpretations. These interviews were open and did not include a predefined battery of questions, but roughly followed the same structure:

- (a) explanation of the content and purpose of the interview,
- (b) visualisation of the excerpts analysed with the speakers (raw video without transcripts or annotations),
- (c) discussion of aspects the analysis had revealed as relevant for the persuasive effect (e.g. a particular use of intonation), including speakers opinion about intent and potential effect on audience,
- (d) exchange of interpretations regarding the intent and potential effect of these aspects (the speakers' interpretation was prompted before we offered them ours).

For the sake of consistency, the same researcher carried out all the interviews. All interviews were audio recorded and written notes were taken by the researcher.

## 6. An MDA approach to persuasive strategies

Our MDA approach to persuasive strategies hinges on integrating previous literature, multimodal analysis and ethnographic data. Triangulating input from literature with our results from multimodal and ethnographic analysis we have identified a series of persuasive strategies that are realised through different modes (words, intonation gestures and head movements) in a corpus of 5 dissemination talks, as Figure 3 below illustrates. The rest of this section is devoted to describing and illustrating these strategies. Although the triangulation of results was carried out for the four modes included in the analysis, for the sake of clarity and brevity the speaker's comments are only included in the discussion of the examples when they are divergent from ours or add any additional aspects. They are omitted when they simply confirm our interpretations.

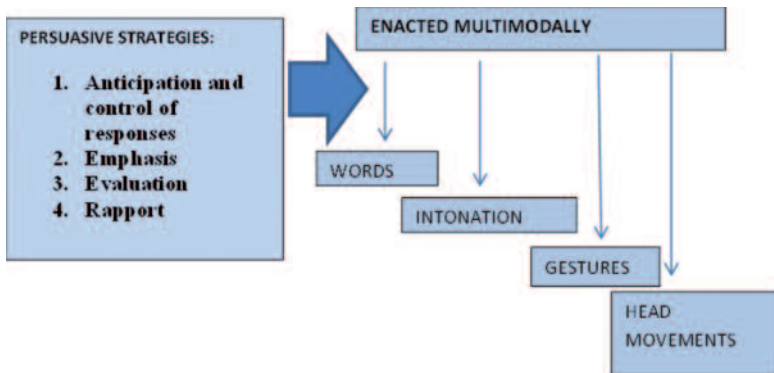


Figure 3. Persuasive strategies.

The first strategy, that is, “Anticipation and control of responses” takes place when speakers predict reactions and adapt their behaviour accordingly in order to obtain a desirable response. Previous literature has shown how this can be achieved through words, intonation, gestures and head movements. Carter (1997), for example, points out how the use of implicit comparisons, question tags and rhetorical questions all trigger a presupposed response from the audience. Brazil (1997) argues that a final high pitch (high termination) can be a cue to show that the speaker expects the listener to be surprised. Gestures and head movements can also be used to prevent potential counterclaims, especially when they co-occur with extreme evaluations that speakers anticipate as debatable (Mac Clave, 2000; Kendon, 2002; Kendon, 2004).

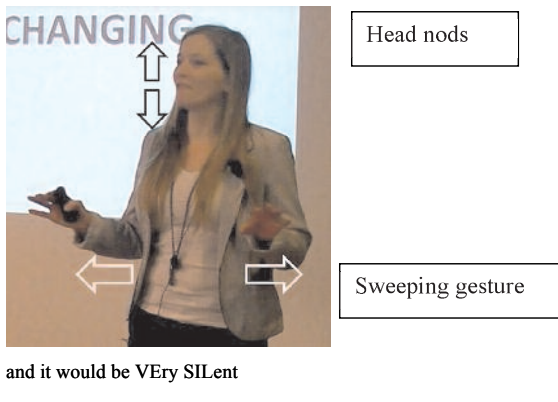
The following example (Example 1) shows how the speaker is anticipating answers using words, intonation, gestures and head movements. The excerpt corresponds to the opening lines of the presentation.

Example 1: Anticipation of responses in excerpt 2.1



**DI (discourse intonation) transcription<sup>i</sup>**

1. ARE YOU SMART
2. **i KNEW that this would come**
3. but LET me rePHRASE
4. is your PHONE smart
5. **you would say of COURSE it is a smart PHONE phone**
6. so WHY do you ASK
7. well then I SAY
8. is the CLOTHES you WEAR smart
9. the JEANS
10. the JACKET
11. the CLOTHES that you have
12. **and it would be VErY SILent**

**Gestures and head movements**

The words marked in bold in this transcript show how the speaker explicitly claims to be anticipating answers. Interestingly, this is also noticeable in the gestures she uses, in particular the sweeping gesture with both hands palm down (open hand prone, OHP) accompanying “very silent”. During the interview the speaker explained that she had chosen a question as a hook to open her presentation because other types of hook (e.g. an image) can distract the attention from the speaker, which is not desirable, as she was informed in specialized courses on presenting skills. She also confirmed that

her communicative intention in this case was to anticipate responses from the audience, and rather than being interested in the information she wanted to get them “activated”. Regarding the gesture, the speaker accounted for it as a metaphoric representation of silence. On the other hand, we interpret it as a metaphoric, pragmatic gesture having a dual meaning: a) it prevents potential counterclaims to this hypothetical situation that she is describing and b) it represents the scope implied (the whole audience). The speaker found both interpretations plausible and not contradictory with her initial explanation. Interestingly, the gesture is split in two parts to follow the rhythm created by the two prominences in “very silent”, which is also reinforced by subtle head nods synchronous with the prominences. Apart from this reinforcing effect, the speaker interpreted these nods as confirming that she was indeed anticipating the right answers. It is particularly interesting to note that, although words begin to play a role in anticipating answers earlier in the excerpt, towards the end the four modes are carefully orchestrated so that they work simultaneously to achieve the speakers aim: anticipate a response from the audience.

Strategy 2, “Emphasis”, refers to highlighting parts of the message so that they receive more attention. Intonation can contribute through the use of prominent syllables (Brazil, 1997). It can also be used in combination with rhetorical devices that make the text more memorable, such as lexical creativity, three-part lists, parallel structures, and words related to the semantic field of novelty (Lakoff, 1982; Carter, 1997; Bamford, 2007, 2008). Emphasis can further be aided by beats (McNeill, 1992; Kendon, 2004).

The following example (Example 2) shows how the speaker uses words, intonation and gestures to emphasize a part of his message.

Example 2: Emphasis in excerpt 5.2

DI transcription
<ol style="list-style-type: none"> <li>1. so THINK about your KITchen</li> <li>2. if you WANT to HEAT stuff</li> <li>3. <b>REAlly FAST</b></li> <li>4. you are NOT going to use an ORdinary oven</li> <li>5. you are GOING to use the Mlcrowave</li> </ol>

### Gestures



REALLY FAST



It is remarkable how a closed fist beat synchronous with “really” intensifies the adverb sequence in “really fast”. In the interview the speaker also interpreted this gesture as a way of showing emotional involvement. The prominence in the words “really fast” also contributes to the emphasis.

“Evaluation”, strategy 3, occurs when speakers assess something and are thus implicitly inviting the listener to accept this opinion (Bamford, 2007; Querol-Julián, 2011). Hood and Forey (2005) point out how speakers can include multimodal expressions of attitude in their introductions that subtly evaluate their presentations in positive terms and seek alignment with their audience. Along this line, Pomerantz (1986) notes how extreme case evaluations are frequently used to legitimize claims when speakers expect possible counterclaims. Interestingly, these claims are commonly accompanied by head shakes or gestures, which seem to deny in advance a potential counter-argument (Kendon, 2002). This was the case in Example 1 shown above. Another example of multimodal evaluation is provided in the following excerpt (Example 3).

Example 3: Evaluation in excerpt 3.2

DI transcription	
1.	<b>COOL</b>
2.	i <sup>THINK</sup>
3.	WELL
4.	if you're <sup>GEE</sup> ky eNOUGH
5.	YOU SHOULD find it <b>cool</b>
6.	eehh

Gestures and head movements	
	Head nods
if you're <sup>GEE</sup> ky eNOUGH	
	Head nods
YOU SHOULD find it <b>cool</b>	

In this excerpt we find two head nods synchronous with “geeky” and “find”. The speaker had not initially noticed them until we pointed them out to him when watching the video during the interview, which shows that they were done unconsciously. The speaker interprets them as “emphasizing my agreement, I think” (meaning agreement with the evaluation he had just

expressed). We believe they also prompt the audience to share this positive evaluation and concur with “cool”. It is also interesting to note that the second nod precedes speech, being synchronous with “find” and not with the evaluation itself (“cool”). This is probably a result of the greater encoding effort implied by using words as opposed using gestures or head movements, which are easier and faster to produce (McNeill, 1992). Regarding intonation, unit 4 has a high key (+ 124 Hz) that adds the meaning of reversal of expectations. This is consistent with the innovative use of slang (“geeky”, “cool”) in this situation. The speaker commented in the interview that he was consciously using unexpected vocabulary and agreed with us that his intention with this was to set a humorous tone. Interestingly, unit 5 begins with a low key (-37 Hz) that has an equating effect (being geeky= find cool) and ends with a mid termination that prompts concurrence. In sum, this example shows the orchestration of words, intonation and head nods working together to convey a positive evaluation and prompt concurrence with it.

“Rapport” (strategy 4), in the context of presentations, refers to a relationship of sympathy and mutual understanding with the audience. Intonation, for example, can achieve these effects in a more subtle way than words. It can be used in combination with inclusive pronouns that help create rapport with the audience and enhance the sense of shared knowledge (Rowley-Jolivet & Carter-Thomas, 2005; Bamford, 2007, 2008). Gestures in conference presentations can also help create this sense of inclusivity, bringing the audience into the discussion and establishing common ground (Hood & Forey, 2005; Holler 2010).

The following example (Example 4) shows the use of words and gesture to build rapport.

Example 4: Rapport in excerpt 4.1

DI transcription	
1.	<b>i'm <sup>JUST</sup> trying to GET you from a to b</b>
2.	<b>that's ALL i want to do</b>



In this example the speaker is using words that are overtly humble, making himself likable and therefore paving the ground for rapport. The open hand supine (OHS) metaphoric gesture synchronous with “that’s all I want to do” suggests openness, fulfilling a pragmatic function. According to the speaker he wanted to transmit “there’s nothing more to it”. The meaning of the words and the gesture reinforce each other to convey “honest humbleness”, and the whole ensemble is consequently prone to create rapport.

## 7. Discussion and conclusions

As already stated in the introduction section, the objective of this paper is twofold, that is, describing a methodological framework that can facilitate a MDA approach to persuasion in oral genres and applying it to the genre of dissemination talks in order to explore how persuasion is multimodally ensembled in this specific genre.

Concerning the first objective, we have presented a framework for an MDA approach to persuasion in academic and professional genres that combines video-based multimodal analysis with ethnography. We have then tested this framework on a corpus of 5 dissemination talks, which has provided us with the response to our second objective, that is, a tentative taxonomy of 4 persuasive strategies that are realised as multimodal ensembles and that could be considered as a generic trait of dissemination talks.

The examples discussed in section 6 show how speakers orchestrate different semiotic modes into multimodal ensembles to make their presentations

persuasive. Research on persuasion (Fuertes-Olivera et al., 2001; O’Keefe, 2002; Perloff, 2003; Dafouz-Milne, 2007) points out that addressees are more likely to be persuaded by messages that feature certain characteristics (i.e. are credible, easily understood, relevant, memorable, surprising, and they are perceived as inferred rather than imposed). Therefore, speakers employ different combinations of semiotic modes to provide their messages with these characteristics. As the examples analysed in this paper suggest, it is actually the multimodal ensemble taken as a whole that becomes persuasive, to the point that it is difficult to delimit the contribution of each mode and it is not always possible to establish a one to one relationships between modes and persuasive strategies. This finding, in turn, seems to indicate a relation between a skilful integrated use of modes and effective oral persuasive communication, which is in line with results obtained in previous studies of oral presentations (Rendle-Short, 2006; Busá, 2010; Morell, 2015).

A video-based multimodal study of oral discourse such as the one proposed here presents remarkable challenges (partiality of video, time-consuming analysis or tendency to micro-analysis). However, ethnographic methods such as observation and interviews can greatly contribute to overcoming these challenges, providing information that can help select data, analyse them and understand them in the context of the wider and highly complex communicative phenomenon of a presentation. The use of combined methods also provides a unique opportunity to triangulate results and cross-check interpretations, which is particularly valuable in a study involving communicative intent.

It is necessary to point out that the results from this study should be considered tentative, since there are clear limitations, such as the limited number of presentations analysed, or the limited scope of the multimodal analysis. The study presented here is still work in progress. Further analysis is expected to enlarge and refine the taxonomy of multimodal persuasive strategies, especially as the analysis is extended to include other types of presentations and other modes.

Nonetheless, the analysis framework presented in this paper has the added value of being flexible, easily adaptable and data-driven. These characteristics make it particularly appropriate to keep the focus on the multimodal ensemble and avoid prioritizing modes. The taxonomy can serve as a basis for future MDA research on persuasion, and can be further developed to include more modes and persuasive strategies realised

multimodally to answer future research questions in fields such as discourse analysis, genre analysis and language training.

Article history:

Received 01 April 2016

Received in revised form 23 September 2016

Accepted 25 September 2016

## References

- Bamford, J. (2007). "Accentuating the positive. Evaluation and persuasive discourse in business presentations" in J. Bamford & R. Salvi (eds.), *Business Discourse: Language at Work*, 135-155. Rome: Aracne Editrice.
- Bamford, J. (2008). "Telling a convincing story: A corpus assisted analysis of business presentations" in A. Martelli & V. Pulcini (eds.), *Investigating English with Corpora. Studies in Honour of Maria Teresa Prat*, 277-291. Monza: Polimetrica International Scientific Publisher.
- Bateman, J. (2011). *Multimodality and Genre: A Foundation for the Systematic Analysis of Multimodal Documents*. London: Palgrave MacMillan.
- Bavelas, J., Chovil, N., Coates, L. & L. Roe (1995). "Gestures specialized for dialogue". *Personality and Social Psychology Bulletin*, 21,4: 394-405.
- Brazil, D. (1997). *The Communicative Value of Intonation in English*. Cambridge: Cambridge University Press.
- Busà, M.G. (2010). "Sounding natural: Improving oral presentation skills". *Language Value*, 2,1: 51-67.
- Carter, R. (1997). *Working with Texts: A Core Book for Language Analysis*. London: Routledge.
- Carter-Thomas, S. & E. Rowley-Jolivet (2003). "Analysing the scientific conference presentation (CP). A methodological overview of a multimodal genre". *ASp* 39/40: 59-72
- Chun, D. 2002. *Discourse Intonation in L2: From Theory and Research to Practice*. Amsterdam: John Benjamins.
- Dafouz-Milne, E. (2008). "The pragmatic role of textual and interpersonal metadiscourse markers in the construction and attainment of persuasion: A cross-linguistic study of newspaper discourse". *Journal of Pragmatics* 40: 95-113.
- Fuertes-Olivera, P.A., M. Velasco-Sacristán, A. Arribas-Baño & E. Samaniego-Fernández (2001). "Persuasion and advertising in English: Metadiscourse in slogans and headlines". *Journal of Pragmatics* 33: 1291-1307.
- Goldman, R., Erickson, F., Lemke, J. & S. Derry (2007) "Selection in video" in S. Derry (ed.), *Guidelines for Video Research in Education: Recommendations from an Expert Panel*, Data Research and Development Center (NORC at the University of Chicago), 19 - 27. URL: <http://drdc.uchicago.edu/what/video-research.html>
- Holler, J. (2010). "Speakers' use of interactive gestures as markers of common ground" in S. Kopp & I. Wachsmuth (eds.), *Gesture in Embodied Communication and Human-Computer Interaction*, 11-22. Berlin: Springer Berlin Heidelberg.
- Hood, S. & G. Forey (2005). "Presenting a conference presentation. Getting interpersonal with your audience". *Journal of English for Academic Purposes* 4: 291-306.
- Hyland, K. (1998). "Persuasion and context: The pragmatics of academic metadiscourse". *Journal of Pragmatics* 30: 437-455.
- Hyland, K. (2009). *Academic Discourse*. London: Continuum.
- Jewitt, C. (2012). "An introduction to using video for research". National Centre for Research Methods (NCRM) unpublished working paper. URL: [http://eprints.ncrm.ac.uk/2259/4/NCRM\\_workingpaper\\_0312.pdf](http://eprints.ncrm.ac.uk/2259/4/NCRM_workingpaper_0312.pdf) [03/02/2015]
- Jewitt, C. (2013). *The Routledge Handbook of Multimodal Analysis*. London: Routledge.
- Kendon, A. (2002). "Some uses of the head shake". *Gesture*, 2, 147-183.
- Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.
- Kipp, M. (2014) "ANVIL: The Video Annotation Research Tool" in Durand, U. Gut & G. Kristoffersen (eds.), *Handbook of Corpus Phonology*, 420-436. Oxford: Oxford University Press.



- Kress, G. & T. Van Leeuwen (2001). *Multimodal Discourse: The Modes and Media of Contemporary Communication*. London: Arnold.
- Kress, G. (2010). *Multimodality: A Social Semiotic Approach to Contemporary Communication*. London: Taylor and Francis.
- Kress, G. (2012) *Multimodality. A Social Semiotic Approach to Contemporary Communication*. London: Routledge.
- Kress, G. (2014). "The rhetorical work of shaping the semiotic world" in A. Archer & D. Newfield (eds.), *Multimodal Approaches to Research and Pedagogy: Recognition, Resources, and Access*, 131-152. New York: Routledge.
- Lakoff, R. (1982). "Persuasive discourse and ordinary conversation, with examples from advertising" in D. Tannen (ed.), *Analysing Discourse: Text and Talk*, 25-42. Washington D.C.: Georgetown University Press.
- Martin, J. & P. White (2005). *The Language of Evaluation: Appraisal in English*. London: Palgrave Macmillan.
- McClave, E. (2000). "Linguistic functions of head movements in the context of speech". *Journal of Pragmatics* 101: 72-87.
- McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. Chicago & London: The University of Chicago Press.
- Morell, T. (2015). "International conference paper presentations: A multimodal analysis to determine effectiveness". *English for Specific Purposes* 37: 137-150.
- Norris, S. (2004). *Analysing Multimodal Interaction: A Methodological Framework*. London & New York: Continuum.
- O'Keefe, D. (2002). *Persuasion: Theory and Research*. London: Sage Publications.
- O'Halloran, K. & B. Smith (eds.) (2011). *Multimodal Studies. Exploring Issues and Domains*. London: Routledge.
- Perloff, R. (2003). *The Dynamics of Persuasion*. New Jersey, London: Lawrence Erlbaum Associates.
- Poggi, I. & C. Pelachaud (2008). "Persuasion and the expressivity of gestures in humans and machines" in I. Wachsmuth, M. Lenzen & G. Knoblich (eds.), *Embodied Communication in Humans and Machines*, 391-424. Oxford: Oxford University Press.
- Pickering, L., Hu, G.G. & A. Baker (2012). "The pragmatic function of intonation: Cueing agreement and disagreement in spoken English discourse and implications for ELT" in J. Romero-Trillo (ed.), *Pragmatics and Prosody in English Language Teaching*, 199-218. Springer: Netherlands.
- Pomerantz, A. (1986). "Extreme case formulations: A way of legitimizing claims". *Human Studies* 9: 219-229.
- Querol-Julián, M. (2011). *Evaluation in Discussion Sessions of Conference Paper Presentations: A Multimodal Approach*. Saarbrücken: LAP Lambert Academic Publishing GmbH & Co. K.G.
- Querol-Julián, M. & I. Fortanet-Gómez (2014). "Evaluation in discussion sessions of conference presentations: Theoretical foundations for a multimodal analysis". *Kalbotyra* 66: 77-97.
- Rendle-Short, J. (2006). *The Academic Presentation: Situated talk in Action*. Hampshire: Ashgate Publishing.
- Rowley-Jolivet, E. & S. Carter-Thomas (2005). "The rhetoric of conference presentation introductions: Context, argument and interaction". *International Journal of Applied Linguistics* 15,1: 45-70.
- Rubin, H.J. & I.S. Rubin (1995). *Qualitative Interviewing. The Art of Hearing Data*. London: Sage Publications.
- Schindler, L. (2009). "The production of 'vis-ability': An ethnographic video analysis of a martial arts class" in U. Kissmann (ed.), *Video Interaction Analysis*, 135-154. Frankfurt: Peter Lang.
- Sparks, J., C. Areni & K. Cox (1998). "An investigation of the effects of language style and communication modality on persuasion". *Communication Monographs* 65,2: 108-125.
- Valeiras-Jurado, J. & N. Ruiz-Madrid (2015). "A multimodal approach to product presentations". *Procedia. Social and Behavioral Sciences* 173: 252-258.
- Valeiras-Jurado, J. (2015). "A multimodal approach to persuasion in conference presentations" in B. Crawford-Camiciottoli & I. Fortanez-Gómez (eds.), *Multimodal Analysis in Academic Settings. From Research to Teaching*, 108-130. London: Routledge.
- Virtanen, T. & H. Halmari (2005). "Persuasion across genres: Emerging perspectives" in H. Halmari & T. Virtanen (eds.), *Persuasion across genres*, 2-24. Amsterdam, Philadelphia: John Benjamins.

**Julia Valeiras-Jurado** is a teaching assistant for Business and Scientific English at Ghent University. Her research interests are oral discourse and genres, intonation and gestures as semiotic modes, and persuasive language. She has recently contributed to the Routledge volume *Multimodal Analysis in Academic Settings. From Research to Teaching* with the chapter “A multimodal Approach to Persuasion in Conference Presentations”.

**M<sup>a</sup> Noelia Ruiz-Madrid** is Senior lecturer in the Department of English Studies at Universitat Jaume I (Castellón, Spain), where she teaches both undergraduate and postgraduate courses in the areas of EAP, ESP and EFL methodology. Her research focuses on discourse analysis, genre and ICTs and EFL. Her more recent publications have appeared in *Ibérica* and *Routledge*.

**Geert Jacobs** is a Full Professor of Language for Specific Purposes at Ghent University. He is the co-founder and head of the international NewsTalk&Text research group. He teaches Business Communication and English for Specific Purposes in various Bachelor's and post-graduate programmes. He has published in international journals in the fields of business communication, journalism and linguistic pragmatics.

## NOTES

<sup>1</sup> <http://www.fon.hum.uva.nl/praat/>.

<sup>2</sup> <http://tla.mpi.nl/tools/tla-tools/elan/>

<sup>3</sup> Tone units are sequentially listed and numbered. Prominent syllables are capitalized. Syllables below or above the line represent a lower or higher pitch (key and termination). Words marked in bold play a role in the persuasive strategy.