Modal coherence in specialised discourse: a case study of persuasive oral presentations in business and academia

Julia Valeiras-Jurado
Ghent University (Belgium)
julia.valeirasjurado@ugent.be

Abstract

The multimodal character of academic and business discourse has long been acknowledged. Several studies have brought to the fore the multimodal nature of academic and business genres, and have shown that spoken or written words are only one among the many resources available to convey meaning (Carter-Thomas & Rowley-Jolivet, 2003; Kast, 2008; Querol-Julián & Fortanet-Gómez, 2012, 2014; Morell, 2014, 2015). In particular, in the case of persuasive oral genres within business and academia, semiotic modes such as gestures (Kendon, 2004) or intonation (Brazil, 1997) play a crucial role in the design of a persuasive message. This is the case of conference presentations, research dissemination talks and product pitches. In these genres speakers have been shown to resort to multimodal persuasive strategies to craft an effective presentation (Valeiras-Jurado & Ruiz-Madrid, 2015; Valeiras-Jurado, 2015; Valeiras-Jurado, Ruiz-Madrid & Jacobs, 2018).

This paper probes into what exactly makes these persuasive presentations effective. Specifically, the question addressed is whether it is the choice of persuasive strategies, the number of semiotic modes, or the consistency in the use of these modes (i.e. modal coherence) that affects persuasion more directly.

The case study presented in this paper compares a research dissemination talk and a product pitch. The study combines video-based, computer-aided multimodal discourse analysis (MDA) with ethnographic interviews. The results suggest that modal coherence is of crucial importance for the effectiveness of a persuasive presentation. Likewise, several practices detrimental to persuasion are identified. These findings have interesting

pedagogical implications, since they can contribute with a genre-based, multimodal methodology to the teaching of English for Specific Purposes (ESP).

Keywords: multimodality, oral genres, modal coherence, persuasion.

Resumen

La coherencia modal en el discurso especializado: un caso de estudio de presentaciones orales persuasivas en el sector empresarial y en el sector académico

La naturaleza multimodal del discurso académico y empresarial se ha abordado en diversos estudios que demuestran cómo el lenguaje escrito u oral es solo uno de los muchos recursos semióticos disponibles (Carter-Thomas & Rowley-Jolivet, 2003; Kast, 2008; Querol-Julián & Fortanet-Gómez, 2012, 2014; Morell, 2014, 2015). En el mundo académico y empresarial se utilizan ciertos géneros orales en los que recursos como la gestualidad (Kendon, 2004) o la entonación (Brazil, 1997) desempeñan un papel crucial en el diseño de mensajes persuasivos, como es el caso de las charlas de divulgación científica y de las presentaciones de productos, dos géneros en los que los ponentes recurren a estrategias persuasivas multimodales para diseñar una presentación efectiva (Valeiras-Jurado & Ruiz-Madrid, 2015; Valeiras-Jurado, 2015; Valeiras-Jurado, Ruiz-Madrid & Jacobs, 2018).

El objeto de este trabajo es contribuir a determinar qué hace que estas presentaciones sean eficaces y convincentes. En particular, se intenta dilucidar cuál de los tres factores siguientes consigue un efecto persuasivo mayor: la selección de estrategias persuasivas, el número de recursos semióticos empleados, o la consistencia en el empleo de dichos recursos (es decir, la coherencia modal).

El estudio combina dos métodos: el análisis del discurso multimodal y la etnografía. Se trata de un análisis comparativo de una charla de divulgación científica y una presentación de producto. Los resultados sugieren que la coherencia modal ejerce una enorme influencia en el efecto persuasivo de estas presentaciones, y permiten identificar prácticas que pueden afectar negativamente a la persuasión. Estos resultados tienen aplicaciones pedagógicas interesantes, ya que pueden contribuir a mejorar las metodologías para la enseñanza de inglés con fines específicos mediante la aportación de un enfoque multimodal al uso de géneros en el aula.

Palabras clave: multimodalidad, géneros orales, coherencia modal, persuasión.

1. Introduction

1.1. Multimodality in oral business and academic discourse

The multimodal nature of oral academic and business genres has been acknowledged in several studies that have shown how spoken or written words are only one among the many resources available to convey meaning. Carter-Thomas & Rowley-Jolivet (2003), for example, were among the first to approach the academic genre of conference presentations (CPs) from a multimodal perspective. They identify three main modes in this genre: language, visual communication and gesture. Rendle-Short (2005, 2006) has also shown how speakers indicate orientation to the audience through positioning, gaze and action, and mark the macrostructure of their discourse using kinesics, prosody (e.g. reduced pitch or volume) and actions (e.g. changing the slides). Along the same line, Querol-Julián and Fortanet-Gómez (2012, 2014) study the use of kinesic and paralinguistic features during the Q&A session of CPs. They are found to have pragmatic and modal functions, conveying the speaker's attitude or intensifying evaluative meaning. From a more pedagogical perspective, Morell (2014, 2015) uses a multimodal pedagogy in the teaching of CPs, and finds that a skilful combination of modes makes presentations easier to understand. She identifies four modes: verbal, written, non-verbal material and body language. The more aware presenters are of these modes, she claims, the more able they are to combine them effectively, foregrounding them at specific moments of their presentations and achieving effective redundancy. Busá (2010) also uses a multimodal approach to the teaching of English communication skills, and, like Morell, she finds that when students are made aware of the complex interplay of semiotic resources, their oral presentation skills improve.

Concerning research dissemination talks (DTs), the genre of TED talks (www.ted.com) has recently received increasing attention. Recent studies such as Caliendo (2014), Caliendo and Compagnone (2014), and Compagnone (2015) describe TED talks as a genre lying at the intersection of a number of other genres, including university lectures, newspaper articles, conference presentations and TV science programmes. Interestingly, Caliendo (2014) and Compagnone (2015) also acknowledge their multimodal nature, noting that it mixes different semiotic modes (i.e. spoken, written, video and audio). McGregor et al. (2016) focus on a specific multimodal aspect of this genre: intonation. This study finds that intonation strongly

contributes to conveying the speaker's attitude towards both the content and the audience, and proposes the use of TED talks as valid material to teach intonation. Similarly, Masi (2016) centres on the use of gestures, and shows the important contribution of deictic and metaphoric gestures to the meaning-making process in a selection of TED talks.

Regarding product pitches (PPs), Steve Jobs's presentation of the iPhone in 2007 is an example that has attracted considerable attention. Kast (2008) identifies a complex interplay of semiotic resources in this presentation. For example, appeals to emotions are achieved through the use of rhetorical figures of speech which are supported by non-verbal and paraverbal communication, as well as by visual aids. He also notes how argumentations are very often implicit or even replaced by multimedia aspects (e.g. an image in a slide can make explicit the conclusion the speaker wants to make, or prove a claim). With a focus on intonation, Niebuhr et al. (2016) analyse the use of prosodic features in Steve Jobs's presentations, to find out how intonation contributes to projecting a charismatic figure. They reveal important characteristics that make Jobs's public addresses stand out from the rest. For example, Jobs's pitch range is particularly high for a male voice, and he produces a particularly varied pitch contour throughout his presentations. He also varies loudness, and, in comparison with other public speeches, produces shorter tone units at a relatively fast and fluent speech rate. In addition, this use of prosody frequently co-occurs with emphatic words that are elongated and pronounced louder for emphasis. The importance of multimodality in product pitches has also been approached from a pedagogical perspective. Palmer-Silveira (2015) shows how students can improve communication skills in their presentations when they learn how to exploit proxemics, head movements, gestures and eye contact. For example, a good use of space and body posture can help students project confidence. In addition, gently nodding while smiling can help reinforce positive aspects of the product. In contrast, overusing hand gestures can be distracting.

1.2. Persuasion in conference presentations, research dissemination talks and product pitches

Another common feature of the oral genres discussed in this paper is that they are, in addition to multimodal, persuasive. Rowley-Jolivet (1999) and Rowley-Jolivet and Carter-Thomas (2003, 2005) claim that CPs are inherently persuasive. In the same line, Räisänen (1999, 2002) points out

that the primary aim of CPs is to publicise, critically discuss and ratify research, which turns it into a persuasive genre. Swales and Burke (2003), Wulff et al. (2009) or Querol-Julián (2011) classify the genre of CPs within the "contingent" repertoire of academic discourse, which is more personal and includes insights and beliefs from researchers, in contrast to the "empiricist" repertoire, which includes more objective and impersonal genres.

The role of persuasion in research dissemination has been dealt with in recent literature, with particular emphasis on TED talks. For example, Scotto di Carlo (2014) has performed a case study that shows how persuasion is very different in a TED talk and a CP dealing with the same topic and delivered by the same speaker. This author draws on Aristotle's three types of persuasive appeals to the audience: *pathos,* or appeal to emotions; *ethos,* or appeal to the credibility and reliability of the communicator; and *logos,* or appeal to rationality. Regarding the first persuasive appeal, the TED talk resorts much more to pathos through the use of stories, videos, pictures, graphs and other visual media (note how these resources include a variety of semiotic modes). The speaker also uses evaluative adjectives and engages the audience using direct addresses and inclusive pronouns. In addition, the speaker's position on stage and delivery style create proximity with the audience and enhance the sharing of emotions. In relation to the second persuasive appeal, the way the speaker applies ethos in the TED talk differs from the CP: in the TED talk "trustworthiness" (i.e. ability to be consistent with the message) and "similarity" (i.e. ability to make the audience identify with you) play a more important role than "authority" or "reputation". Third, concerning logos, in the TED talk the speaker uses examples, metaphors and visuals to enhance the comprehension of the message, while in the CP the use of specialised terminology limits the type of audience that will be able to follow the presentation.

Regarding persuasion in business presentations, Bamford (2007) also identifies Aristotle's three persuasive appeals in this genre by appeals to reason (complex argumentations), appeals to the credibility of the speaker (achieved linguistically by statements of belief, positive evaluations, emphatics and hyperboles) and emotional appeals (achieved linguistically by means of personal testimonies and direct addresses to the audience). Interestingly, Bamford also acknowledges that aspects such as body language and intonation play a role in persuasion and that they can reveal useful information about the communicative intention of the speaker. Other

interesting studies about persuasion in oral business presentations are provided by Bowker (2006, 2007). Bowker focuses on linguistic features, and notes how certain narrative techniques are employed to make oral presentations more persuasive. One of these techniques is the change from indirect to direct mode, or impersonating instead of referring to a third person. The information presented in this way is loaded with affective force, and perceived as particularly relevant and real. A second technique is the use of rhetorical questions to guide interpretation. Third, extension questions can also be used to prompt the audience to relate ideas to their own experience, and check their relevance. These narrative techniques are aimed basically at easing interpretation and presenting information as relevant for the addressee, which is expected to enhance persuasion, as evidenced in the following section.

1.3. Multimodal persuasive strategies

As pointed out by Valeiras-Jurado, Ruiz-Madrid and Jacobs (2018), in persuasive oral genres some strategies or communicative techniques are likely to be used for persuasive purposes and enacted multimodally. Seven multimodal persuasive strategies have been identified in the genres that concern this paper, which are briefly explained in the following paragraphs.

"Anticipation and control of responses" refers to the way speakers predict reactions and adapt their behaviour accordingly in order to obtain a desirable response. In other words, they anticipate responses from their audience and also trigger or prevent (i.e. control) certain responses to better suit their communicative intentions. Previous literature has shown how this can be achieved through words, intonation and gestures. Carter (1997) for example points out how the use of question tags, rhetorical questions and implicit comparisons can trigger a particular 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 can also be used to prevent potential counterclaims, especially when they co-occur with extreme evaluations that speakers anticipate as debatable (Kendon, 2004).

The strategy "attention getting" is used to raise and maintain the interest of the audience. It can take a variety of forms (i.e. realised in a variety of semiotic modes) including a particularly marked use of intonation (e.g. a high pitch), a marked use of gesture (e.g. wider amplitude or more repetitions) or

the use of stylistic devices such as narratives or direct addresses to the audience. Attention-getting techniques are especially relevant in presentation openings, but can also appear later on.

"Emphasis" consists in highlighting parts of the message so that they become more salient. 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).

"Evaluation" occurs when speakers assess something and are thus implicitly inviting the listener to accept their 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.

"Processing aids" are resources that help the message to be processed in real time and make it easy to understand, such as visual metaphors materialised through gestures that can help clarify aspects of the referential content of the message, or a use of intonation that clarifies the structure of a message (e.g. contrasting elements, or elements in a list).

"Projection of the context of interaction" is a term borrowed from Discourse Intonation. According to Brazil (1997), speakers use intonation to project their understanding of the context of interaction as they communicate to negotiate and classify information: what is considered new vs. given information; or contrastive (selection among a range of options) vs. non-contrastive (no other options available). In the context of oral presentations this means that speakers present some parts of the message as shared with the audience and agreed upon, as opposed to new and open to discussion.

"Rapport" in the context of presentations refers to a relationship of sympathy and mutual understanding with the audience. Tickle and Rosenthal (1990) identify three essential components of rapport, and show how they can be expressed non-verbally:

    i) "mutual attentiveness" that can be expressed through spatial configurations and bodily postures that signal accessibility;

ii) "positivity", which can be materialised through smiling and nodding that signal liking and approval;

iii) "coordination", which can take the form of postural mirroring and interactional synchrony.

Intonation can also achieve these abovementioned effects in a subtler 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; Hyland, 2005; Bamford, 2007, 2008). Gestures in CPs can also help create this sense of inclusivity, bringing the audience into the discussion and establishing common ground (Hood & Forey, 2005; Holler, 2010).

1.4. Modal coherence

A question that is raised when adopting a multimodal approach is whether modes have an inherent hierarchical importance in a message. Norris (2004, 2004b) states that language is not necessarily central in communicative interaction, as there are other semiotic modes that play an important role and are not ancillary to language. In addition, she claims that the number of modes in itself is not significant: what is relevant is their intensity and the complexity of the interrelationships among them. She refers to this as "modal density". Communicative events are complex, and participants engage in different (inter)actions at the same time with varying degrees of attention and awareness. These are called "higher level actions", and are in turn composed by a chain of different "lower level actions". An example of a higher level action can be the action of supporting a claim with an argument. In order to do this, speakers may resort to different lower level actions, such as pointing to visual material or raising their voice, each lower level action using one or more semiotic modes. The persuasive strategies presented in the previous section can be analysed as lower level actions that contribute to the higher level action of "persuading an audience to accept a particular idea", or in more general terms "communicating with an audience".

Modal density makes it possible to organise higher level actions in a continuum of foreground-midground-background according to the intensity and interplay of modes employed in their realisation. Foregrounded higher level actions are modally dense either because they are produced with more

modes that interrelate with each other (i.e. they are modally complex) or because the modes involved in them take a more intense role (i.e. these actions are modally intense).

The concept of "modal density" is highly complex, because it encapsulates number, intensity and interplay of modes. In the present analysis this concept is broken down into smaller aspects. For this reason, the term "modal density" is used to refer exclusively to the number of modes, while the way modes interact with each other is called "modal coherence". This concept refers to the way modes are "orchestrated" into coherent "ensembles", as Kress (2010) puts it. In other words, modal coherence means that the modes are being used in a consistent way and do not contradict each other. This distinction is useful to test the hypothesis that both aspects (i.e. number of modes and coherence among them) can have a distinct influence in the persuasive effect.

In fact, the importance of how modes are combined has been pointed out by Tang (2013), who shows how different combinations of modes do not always produce the same meaning even if they deal with the same content, and warns against the "injudicious use of multiple representations" (Tang, 2013: 34) just for the sake of motivating and getting the attention of the audience.

Therefore, the purpose of this paper is to help elucidate whether it is the choice of persuasive strategies, the number of semiotic modes (i.e. modal density), or the consistency in the use of these modes (i.e. modal coherence) that affects persuasion more directly in oral presentations within business and academia.

2. Methodology

For this case study two presentations (i.e. a research dissemination talk and a product pitch) were analysed using Multimodal Discourse Analysis (MDA) and ethnographic interviews. The DT is part of a TED contest organised by a university, during which researchers presented their research to the general public. The PP is the closing event of a one-year project in which bachelor university students in the Business degree collaborate with companies in the design and launch of new products into the market. During this closing event the resulting products are presented in a product pitch. In both cases the communicative event is framed as a contest, and the winners are

rewarded with a space in a wider-scope TED event and funding for their product, respectively.

The presentations were recorded after obtaining consent, using a static camera that pointed constantly to the speaker. At the end of each event, the researcher interviewed the speakers in a face-to-face semi-structured interview. The main goal of this interview was to gain insights into the speakers' previous knowledge and experience, their attitudes to public speaking and their satisfaction with their performance, since these aspects could help interpret the speakers' multimodal behaviour. More specifically, the researcher probed into the following aspects:

a) what the speakers knew about the event beforehand

b) the way they prepared for the presentation

c) their motivation to participate

d) what they considered to be the main goal with their presentation (their take-home message)

e) whether they were satisfied with the results

Using the valuable input of this first round of interviews, the researcher selected two extracts in each presentation that stood out for their intense persuasive efforts from the speakers (measured in relative terms and compared to other moments in the presentation). The selection was crosschecked by other three researchers.

The next step in the process was a computer-aided MDA analysis of these extracts. The focus of this multimodal analysis lies on three modes: words, intonation and gestures. The term "words", as used in this study, includes speakers' choices in terms of lexis, grammar, style and register. Intonation is frequently defined as variations in pitch that make up a pitch contour. The approach adopted for intonation is "Discourse Intonation" (DI), which sees intonation as a tool facilitating the flow of discourse and helping the negotiation process involved in every communicative interaction (Brazil, 1997). Gestures are understood as "visible action when it is used as an utterance or as part of an utterance" (Kendon, 2004: 7).

Two programs were used for the multimodal analysis: the phonetic analysis tool PRAAT (http://www.fon.hum.uva.nl/praat) and the annotation tool ELAN (http://tla.mpi.nl/tools/tla-tools/elan/). PRAAT provided accurate

measurements of pitch and intensity relevant for the analysis of intonation. ELAN made it possible to annotate information regarding the modes included in the analysis and organise this information in time-aligned tiers. This is of crucial importance for MDA, because it enables the researcher to visualize how the speaker is using each mode at a specific moment.

The annotation of intonation follows DI conventions (Brazil, 1997). The annotation of gestures is based on taxonomies proposed by McNeill (1992), Bavelas et al. (1995), Kendon (2004), and Querol-Julián (2011). Appendix 1 provides more details of these annotation systems. The annotation of words, however, is entirely corpus-driven and did not follow a pre-defined set of categories. There are two reasons for this methodological decision. First, a predefined set of categories for the analysis of words would most likely incline the researcher towards an overestimation of words as a mode, and limit the analysis of the other modes to instances in which they help co-express a particular meaning simultaneously with words (i.e. words would be the primary mode, the others would be ancillary). Second, both the researcher and the speakers participating in the study are more self-conscious of their use of words and need no support to discuss it, but find more difficulties discussing the use of gestures or intonation without the help of a systematic categorization.

The last step of the analysis was a second round of interviews with the speakers, in the case of the PP face to face, in the case of the DT a video call. In this second interview a playback methodology (Norris, 2004) was used. The speakers visualised the video of the excerpts of their presentation that had been analysed, but without annotations or interpretations. Then the researcher drew attention to particular uses of words, intonation and gestures that the MDA had revealed as potentially relevant for persuasion. The speakers' interpretation was prompted, followed by a discussion of the intended effect of their multimodal behaviour. Most speakers remembered word choices as conscious efforts to be persuasive, but the use of intonation and gestures required more reflection. Finally, the views of the speakers were contrasted with the results and interpretations of the researcher. In many cases they concurred. In others, the alternative views of the speaker enriched the analysis and made it more solid. This exchange of interpretations also contributed to a greater objectivity in the relevance attached to each mode. It counteracted the inherent tendency in video-based analysis to minute detail, and helped assign the correct importance to each modal aspect.

A more detailed account of the combined methodology employed in this study can be found in Valeiras-Jurado et al. (2018).

3. Results and discussion

In the present study two similar excerpts of oral presentations (i.e. a DT and a PP) are analysed and compared for effectiveness. The term "effectiveness" here refers to the extent to which speakers manage to fulfil their communicative goals, i.e. to convince the audience of the value of a scientific methodology (in the DT) and a product (in the PP). The results of the combined ethnographic and MDA analysis suggest that the DT is more effective than the PP.

First of all, it is worth noticing that the speaker of the DT won the first prize in the contest. As reported by the speaker, the jury particularly valued the use of metaphors and analogies to make the topic accessible, and the enthusiasm that pervaded the presentation. On the other hand, the PP was less successful in the contest. The speaker also reported being less satisfied with his performance. When asked about his use of modes, the speaker in the PP explained that his performance was influenced by his efforts (not always successful) to remember a script he had prepared and repeatedly rehearsed.

Secondly, the MDA analysis shows that both extracts present strong similarities in their use of multimodal persuasive strategies, but also noticeable differences. Table 1 presents the orthographic transcription of the DT, in which the speaker presents his topic using an analogy between chemistry and cooking.

Orthographic transcription DT
Now which experience am I talking about? Well, you all know a bit of cooking, and in fact there's not much of a difference between chemistry and cooking. In both cases you need a recipe which tells you which ingredients you are going to use. Now in my case this was hafnium chloride and benzyl alcohol. I just needed to mix these (oops) I just needed to mix these and put them in an oven at 220 degrees. And then I had to wait. For three days. And this was a bit of a personal crisis, because I really can't wait. I like fast reactions. Who doesn't like fast? So think about your kitchen. If you want to heat stuff really fast you are not going to use an ordinary oven. You are going to use the microwave.

Table 1. Orthographic transcription DT

The speaker directly addresses the audience through two rhetorical questions and an imperative ("think"), getting them engaged. He also implicitly evaluates a situation (slow is bad, fast is better) using intense language

("personal crisis", "really fast"). In addition, he uses parallelisms with everyday experiences that help understanding and bring his topic close to the audience.

Intonation is skilfully used in this excerpt to guide the audience towards a desirable response. Table 2 presents the DI transcription (see Appendix 1 for details of transcription conventions).

DI transcription of DT
1.   now WHICH experience am i TALking about
2.   well you ALL know a bit of COOking
3.   and in FACT there's not much of a DIfference
4.   between CHEmistry and COOking
5.   in BOTH cases you need a REcipe
6.   which TELLS you which inGREdients you are going to use
7.   now in my case this was HAFnium chloride and BENzyl alcohol
8.   i just needed to MIX these oops
9.   i just needed to MIX these and PUT them in an oven at 220 degrees
10.   and THEN i had to WAIT
11.   for THREE DAYS
12.   and this was a bit of a PERsonal CRIsis
13.   because i really CAN'T WAIT
14.   i like FAST reactions
15.   who DOESN'T like FAST
16.   so THINK about your KITchen
17.   if you WANT to HEAT stuff
18.   REAlly FAST
19.   you are NOT going to use an ORdinary oven
20.   you are GOING to use the MIcrowave

Table 2. DI transcription of DT

For example, in units 3 and 20 the speaker uses a low key that has an equating effect and presents his statement as taken for granted. This prevents any possible discussion as to whether chemistry and cooking are really similar or whether a microwave oven is faster than an ordinary one. During the interview the speaker confirmed that this was actually the effect intended in the case of unit 20, but not in unit 3 (he was not consciously presenting the similarity between chemistry and cooking as taken for granted). He also uses a high key in unit 10 and unit 15. In unit 10 this has the effect of presenting this need to wait as something surprising and not what should be expected. As a consequence, this waiting time is interpreted negatively, as something that should not be there. In unit 15 the high key also adds a nuance of surprise, but in this case the speaker conveys surprise at a possible response

to this question (=I don't like fast). This is a way of implicitly evaluating fast as good and slow as bad (nobody likes waiting, and the opposite would be surprising). This effect is reinforced by the high termination in unit 11, which prompts an active role from the audience and requests a judgment regarding the need to wait for three days (which the speaker expects to be evaluated as negative).

Gestures are also used extensively in this excerpt. "Chemistry and cooking" are synchronous with alternating beats to both sides with an open hand supine position (see Appendix 1 for details about the gesture annotation conventions). They visually represent the parallelism conveyed between these two concepts (see Figure 1).

Figure 1. Beats to both sides.

They fulfil a cohesive function, bringing together the two elements of a comparison. The same gesture occurs again in units 19 and 20, this time to clarify a contrast between the ordinary vs. microwave oven. The speaker actually confirmed this was a way of "giving structure" for the audience.

Later, both hands moving in circles represent the action of "mixing" (see Figure 2).

Figure 2. Hands moving in circles.

This is a metaphoric, referential gesture that works as a processing aid.

Then an open hand supine gesture, as if yielding the turn and inviting an answer from the audience, is performed in synchrony with the rhetorical question "Who doesn't like fast" (Figure 3).

Figure 3. Yielding the turn.

This is a metaphoric gesture that fulfils an interpersonal function. In this case the response of the audience is not verbalised, but it is likely to happen as a mental process. The speaker reported that this gesture was planned and intended to show honesty and create rapport, and agrees with the interpersonal function of inviting an answer. There is a second example of

an interpersonal gesture in the excerpt: a gesture with parallel open hands starting at his head and moving outwards representing the action of thinking (Figure 4).

Figure 4. Thinking

This is a metaphorical, interpersonal gesture that again prompts a response from the audience, and is synchronous with the imperative "think".

Finally, a closed fist beat (Figure 5) intensifies the extreme evaluation in "really fast".

Figure 5. Intensifying beats

This gesture fulfils a pragmatic function. The speaker also interprets it as a way of showing emotional involvement, which supports his contrast fast = good vs. slow = bad. In fact, the same gesture occurs previously, emphasizing the prominences in "I really can't wait" (unit 13).

To sum up, the speaker uses a variety of persuasive strategies realised through words, intonation and gestures:

- Evaluates through words and intonation (i.e. the need to wait as something negative).

- Projects a particular context of interaction through intonation presenting statements as obvious or surprising.

- Directs a response towards a shared positive evaluation (fast is better), using intonation and gestures.

- Uses processing aids through cohesive gestures, visual representations and parallelisms with common experiences.

- Emphasises through gestures that intensify evaluations.

- Builds rapport through gestures that suggest honesty.

- Uses attention-getting techniques that get the audience engaged.

The speaker's use of multimodal persuasive strategies can be graphically represented as the multimodal ensemble shown in Figure 6:
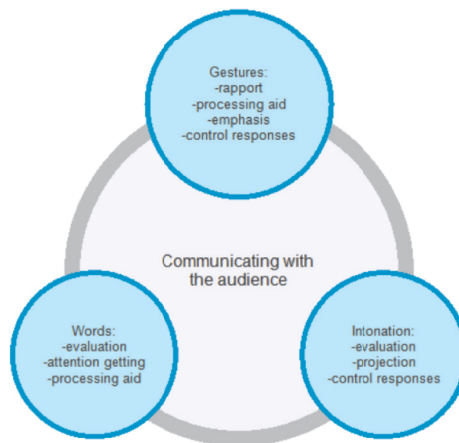

Figure 6. Multimodal ensemble in DT.

With this multimodal ensemble the speaker tries to engage the audience, making his message relevant for them and also easy to understand. This behaviour is consistent with what the speaker reported in the interview, especially in his conscious attempts to make his message accessible and keep the audience engaged.

In the excerpt corresponding to the PP the speaker presents his product (a tablet for doctors) and the need it covers (doctors occasionally need to draw parts of the body to explain their diagnosis). Table 3 shows the orthographic transcription.

Orthographic transcription PP
Now meet my sister. Her name is Ariane, she is a gastroenterologist. She takes a lot of time to take care of her patients and to talk with them. Like the other doctors, she draws badly, but her patients understand her. Why? Because she owns Anapad. Anapad is a small plastic tablet, with parts of the body printed on it, and you can write down and erase, very easily.

Table 3. Orthographic transcription PP

In this example the speaker shares a personal experience that is meant to create rapport. The speaker uses this experience to compare the general situation of doctors ("like the other doctors") and the specific case of his sister, which establishes a contrast between what we have vs. what we need, and therefore highlights the market gap for his product. The "why?" question anticipates the potential reaction of the audience (surprise, reversal of expectations) and introduces a response to it. The speaker confirmed that his intention with this question was to engage people in reflection. As a matter of fact, he remembers he had planned to make a longer pause right after "why" for effect, but finally missed it because he was nervous during the presentation and spoke too fast. The name of the product is repeated for memorability. The speaker reported that the name had been chosen as a marketing technique, precisely because it was easy to remember and would help the audience perceive this product as real.

Only half of the tone units in this excerpt carry a tone movement, which makes intonation rather flat throughout. This flat intonation makes the story about the speaker's sister sound recited and not spontaneous, which in turn makes it less credible and less effective. Still, there are some interesting uses of intonation, as shown in Table 4.

DI transcription of PP
1.   NOW meet my SISter →
2.   her NAME is ariANE ↘
3.   she is a GAStroenteROlogist →
4.   she TAKES a lot of TIME →
5.   to to take CARE of her PAtients ↘
6.   and to TALK with them ↘
7.   likw the other DOCtors she draws BADly ↖
8.   BUT →
9.   her Patients undersTAND her →
10.   WHY →
11.   beCAUSE she OWNS anapad ↖
12.   anapadis a SMALL plastic TAblet →
13.   with PARTS of the body PRINted on it ↘
14.   and you can write DOWN and Erase ↖
15.   VEry easily ↖

Table 4. DI transcription PP

In units 5 and 6 ("to take care of her patients and to talk with them") the rise tone employed suggests this is treated as shared knowledge: everybody knows that doctors need to communicate effectively with their patients. This sets the ground to establish a need for the product he is trying to sell: a tablet to ease communication. Similarly, in unit 7 ("like the other doctors she draws badly") the speaker uses a fall tone that matches the new information he is providing about his sister, but interestingly the mid termination prompts concurrence with this situation and invites to accept it as something widespread. The added meaning is: generally, doctors' ability to draw is not enough to communicate with their patients in an effective way. Again, the speaker is reinforcing the need or market gap. The next tone unit, number 8 ("but"), has a high pitch that adds the meaning increment of "contrary to what you might expect". The same effect is achieved in unit 10 (why?) which anticipates this reversal of expectations experienced by the audience. Unit 11 ("because she owns Anapad") works in a similar way to unit 7: it features a fall tone that indicates this is new information, but ends with mid termination that prompts concurrence. The speaker uses this to present his product as a feasible solution to the need previously established: the tablet helps doctors communicate effectively.

This speaker uses some referential gestures that make the meaning easier to process for the audience, and also beats that can be considered cohesive. For example, as the speaker says "parts of the body" he first brings his hands together and then moves his left hand outwards (Figure 7).

Figure 7. Parts of the body.

This metaphoric gesture can be interpreted as representing the range of body parts available.

Immediately following this gesture, as the speaker pronounces "printed on it" he moves his hands from a resting position close to his body to an open hand vertical position in a semi-circular movement outwards (Figure 8).
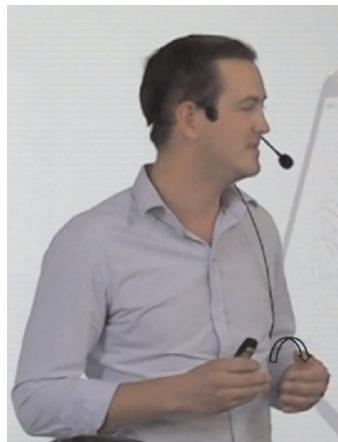

Figure 8. Printed on it

This metaphoric gesture can be interpreted as symbolising the availability of the printed information which visually "comes out".

Then, as the speaker pronounces "write down and erase" a beat is repeated twice (one time at each side) to reinforce the contrast write/erase (Figure 9).

Figure 9. Write down and erase.

The speaker also pointed out that the first time he makes this beat he actually mimics the action of writing briefly, as if holding a pen with one hand and writing on his open palm.

Finally, as he says "very easily" both hands with palms facing up move circularly in a metaphoric gesture that suggests the smooth flow conveyed by "very easily" (Figure 10).

Figure 10. Very easily.

Persuasion in this rich point is built around two key aspects: creating rapport with the audience and showing a market gap. To that aim the use of a personal experience is especially appropriate because it creates rapport and shared ground with the audience and provides a specific example that is more convincing than abstract figures.

Interestingly, the speaker reported during the interview that for him the most important thing in a pitch is to "get the attention of the audience and mention something they are used to". This can be related to his use of an anecdote to create rapport. As he put it himself, his intention was to "impersonate the problem" and make himself likable. Altogether, this should make the message credible and more memorable. However, the speaker uses a rather flat intonation that makes the text sound recited at points, which can be considered "oblique orientation" (Brazil, 1997). This is particularly noticeable as he introduces the narrative with "now meet my sister". It takes away vividness and spontaneity to this narrative because it makes it sound rehearsed, and in turn makes it less credible. As a matter of fact, the speaker admitted during the interview that he "rehearsed a lot", almost memorising the text.

It is also notable that the gestures employed are all concentrated in the very last part of the extract, as the speaker describes the product. This is probably because he needs more visual support to convey the way users can interact with the product and show its potential use. The speaker confirmed this interpretation. During the interview he reported having "a visual mind", and a tendency to explain things through drawings. The multimodal ensemble orchestrated by this speaker can be represented as Figure 11 shows:
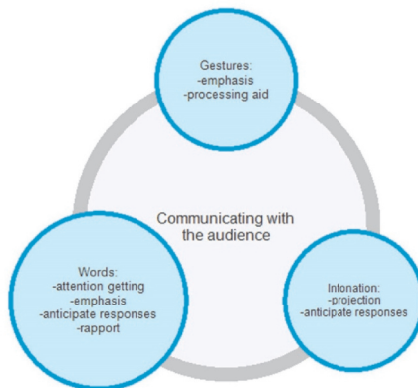

Figure 11. Multimodal ensemble in PP

Table 5 presents a comparative view of the persuasive strategies and their modal realisations in both excerpts.

DTJO1		PPTO1	
Intonation	evaluation	Intonation	
	projection		projection
	control responses		control responses
Gestures	control responses	Gestures	
	emphasis		emphasis
	processing aid		processing aid
	rapport		
Words		Words	
	Attention-getting		Attention-getting
	processing aid		control responses
	evaluation		emphasis
			rapport
7 strategies (3 through several modes)		6 strategies (2 through several modes)	

Table 5. Multimodal persuasive strategies in both excerpts

As the analysis shows, both excerpts use a similar number and choice of strategies. The DT uses a total of seven strategies and the PP a total of six. The only strategy that differs in both excerpts is evaluation: it is present in the DT through words and intonation, but not so evident in the PP. It can be argued that the adjectives "badly" and "easily" introduce some degree of evaluation through words, but the evaluative effect is so subtle that it is outweighed by the speakers' control of a response: presenting a problematic situation that his product will solve.

The examples are also similar in the number of modes used: words, intonation and gestures. In the DT three strategies are realised through different modes (evaluation, control of responses and processing aids) and in the PP two of them are realised through several modes (control of responses and emphasis).

However, the two excerpts present some remarkable differences that seem to have an influence on their final effectiveness. First of all, in the PP there is an overreliance on words: the mode intonation is used less than words (i.e. to realise fewer persuasive strategies), and gestures are only used at the end of the excerpt and again to fulfil fewer persuasive strategies than words.

Second, in the PP intonation is at some points flat and reveals disengagement from the communicative situation. In Brazil's (1997) terms, the speaker is using an oblique orientation. In this case, intonation is not used to realise any of the persuasive strategies (i.e. a lower level action that

would contribute to the higher level action of communicating with the audience). Instead, it is used to fulfil a new higher level action: remembering a script. In addition, the speaker seems to fall into a contradiction in his use of modes: he uses words carefully to craft a vivid representation of how his product can help doctors, but at the same time projects disengagement through a flat intonation that takes away this vividness. Consequently, the use of modes becomes less coherent.

In contrast, the multimodal ensemble orchestrated in the DT shows more coherence. The modes that form part of it realise a series of lower level actions (i.e. a series of persuasive strategies) that consistently contribute to a unique higher level action: communicating with the audience in order to convince them of the validity of his scientific method.

4. Conclusions

The analysis presented in the previous section shows two examples of persuasive oral presentations that vary in their efficiency. Two aspects have been identified as potential causes for this difference:

- Overreliance on words, and consequently neglecting other modes, seems to be detrimental to the persuasive effect.

- Disengaging one mode from the higher level action of communicating with the audience jeopardises the persuasive effect significantly, because the use of modes becomes incoherent.

These results suggest that effective multimodal communication in these presentations is not only a question of quantity but also of quality. In other words: a varied and balanced use of modes seems to be beneficial, but increasing the number of modes will not make communication more effective unless these modes are used coherently. Crucial to this modal coherence is a focus on the higher level action of communicating with the audience, to which all modes should ideally contribute.

These results have direct pedagogical implications. On the grounds of these findings, courses on Academic English and Business English should devote enough attention to teaching students how to master a variety of modes. Students should also be warned against the risk of over-relying on words (or any other mode). They should also be made aware of the detrimental effects of using modes contradictorily. This is in line with previous studies such as Busá (2010), Tang (2013) and Morell (2014, 2015).

In this sense, further research involving more complex multimodal ensembles is needed to expand our knowledge of how persuasion works in oral persuasive genres. Similarly, larger scale studies are needed to reveal patterns of good practices that can be later fed into teaching materials. Such studies can also shed light on distinctive generic traits and help determine if some persuasive strategies have different effects and need to be adapted depending on the genre.

If, as the present case study suggests, persuasive communication in academic and business settings is highly dependent on carefully orchestrated multimodal ensembles, speakers should be trained to become competent conductors of the orchestra of semiotic modes available to them.

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. Rome: Aracne Editrice.

Bavelas, J., N. Chovil, L. Coates, & L. Roe (1995). "Gestures specialized for dialogue". *Personality and Social Psychology Bulletin* 21,4: 394-405.

Bowker, J. (2006). "Referential and affective force in oral business presentations: The role of narration" in J. Bamford & M. Bondi (eds.), *Managing Interaction in Professional Discourse: Intercultural and Interdiscoursal Perspectives*, 58-71. Rome: Officina.

Bowker, J. (2007). "Uncovering Intentionality and the Process of Persuasion in Oral Business Narratives" in J. Bamford, R. Salvi & J. Boker (eds.), *Business Discourse: Language at work*, 47-87. Rome: Aracne Editrice.

Brazil, D. (1997). *The Communicative Value of Intonation in English*, 2nd ed. Cambridge: Cambridge University Press.

Busà, M.G. (2010). "Sounding Natural : Improving Oral Presentation Skills". *Language Value* 2: 51-67.

Caliendo, G. (2014). "The popularisation of science in web-based genres" in G. Caliendo & G. Bongo (eds.), *The Language of Popularisation: Theoretical and Descriptive Models*, 101-132. Bern: Peter Lang.

Caliendo, G., & A. Compagnone (2014). "Expressing epistemic stance in university lectures and Ted Talks : A contrastive corpus-based analysis". *Lingue e Linguaggi* 11: 105-122.

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: La Revue du GERAS* 40,39-40: 59-72.

Compagnone, A. (2015). "The reconceptualization of academic discourse as a professional practice in the digital age : A critical genre analysis of TED Talks". *HERMES-Journal of Language and Communication in Business* 27,54: 49-69.

Holler, J. (2010). "Speakers' use of interactive gestures as markers of common ground". *Gesture in Embodied Communication and Human-Computer Interaction* 5934: 11-22.

Hood, S., & G. Forey (2005). "Introducing a conference paper: Getting interpersonal with your audience". *Journal of English for Academic Purposes* 4(4): 291–306.

Hyland, K. (2005). "A convincing argument: Corpus analysis and academic persuasion" in U. Connor & T. Upton (eds.), *Discourse in the Professions: Perspectives from Corpus Linguistics*, 87-114. Amsterdam: Benjamins.

Kast, B. (2008). *iKeynote-Representation, Rhetoric and Visual Communication by Steve Jobs in His Keynote Ar Macworld 2007* (Unpublished PhD). Salzburg: University of Salzburg.

Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.

Kress, G. (2010). *Multimodality: A Social Semiotic Approach to Contemporary Communication,* 2nd ed. London: Taylor and Francis.

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.

Masi, S. (2016). "Gestures in motion in TED Talks: Towards a multimodal literacy" in V. Bongsinori (ed.), *Multimodality across Communicative Settings, Discourse, Domains and Genres*, 146-165. Newcastle Upon Tyne, UK: Cambridge Scholars Publishing.

McGregor, A., B. Zielinski, C. Meyers, & M. Reed (2016). "An exploration of teaching intonation using a TED Talk" in *Proceedings of the 7th Pronunciation in Second Language Learning and Teaching Conference*, 143-149. Iowa City: Iowa State University.

McNeill, D. (1992). *Hand and Mind*. Chicago: University of Chicago Press.

Morell, T. (2014). "Communicating research at international conferences: A multimodal analysis of an intercultural or a disciplinary specific genre?" in A. Lyda & K. Warchal (eds.), *Occupying Niches: Interculturality, Cross-culturality and Aculturality in Academic Research*, 153-162. Heidelberg: Springer.

Morell, T. (2015). "International conference paper presentations: A multimodal analysis to determine effectiveness". *English for Specific Purposes* 37: 137-150. URL: http://dx.doi.org/10.1016/j.esp.2014.10.002 [03/02/2017]

Niebuhr, O., J. Voße, & A. Brem (2016). "What makes a charismatic speaker? A computer-based acoustic-prosodic analysis of Steve Jobs tone of voice". *Computers in Human Behavior* 64: 366-382.

Norris, S. (2004a). *Analysing Multimodal Interaction: A Methodological Framework*. London & New York: Continuum.

Norris, S. (2004b). "Multimodal Discourse Analysis: A conceptual framework". in P. Levine & R. Scollon (eds.) *Discourse and Technology. Multimodal Discourse Analysis*, 101-115. Washington D.C.: Georgetown University Press.

Palmer-Silveira, J.C. (2015). "Multimodality in business communication: Body language as a visual aid in student presentations" in B. Crawford Camiciottoli & I. Fortanet-Gómez (eds.), *Multimodal Analysis in Academic Settings: From Research to Teaching*, 171-192. London: Routledge.

Querol-Julián, M. (2011). "Evaluation in discussion sessions of conference paper presentations: A multimodal approach". Saarbrücken, Germany: LAP Lambert Academic Publishing GmbH & Co. K.G. Google Scholar.

Querol-Julián, M., & I. Fortanet-Gómez (2012). "Multimodal evaluation in academic discussion sessions: How do presenters act and react?" *English for Specific Purposes* 31,4: 271-283.

Querol-Julián, M., & I. Fortanet-Gómez (2014). "Evaluation in discussion sessions of conference presentations: Theoretical foundations for a multimodal analysis". *Kalbotyra* 66: 77-97.

Räisänen, C. (1999). "The conference forum as a system of genres . A socio-cultural study of academic conference practices in automotive crash-safety engineering". *Acta Universitatis Gothoburgensis. Gothenburg Studies in English*. Göteborg, Sweden: Acta Universitatis Gothoburgensis.

Räisänen, C. (2002). "The conference forum: A system of interrelated genres and discursive practices" in E. Ventola, C. Shalom, & S. Thompson (eds.), *The Language of Conferencing*, 69-93. Frankfurt a.M.: Peter Lang.

Rendle-Short, J. (2005). "Managing the transitions between talk and silence in the academic monologue". *Research on Language and Social Interaction* 38,2: 179-218.

Rendle-Short, J. (2006). *The Academic Presentation: Situated Talk in Action*. Farnham, UK: Ashgate.

Rowley-Jolivet, E. (1999). "The pivotal role of conference papers in the network of scientific communication" *ASp* 26: 179-196.

Rowley-Jolivet, E., & S. Carter-Thomas (2005).

"The rhetoric of conference presentation introductions: Context, argument and interaction". *International Journal of Applied Linguistics* 15: 45-70.

Scotto di Carlo, G. (2014). "Ethos in Ted Talks : The role of credibility in popularised texts". *Facta Universitatis* 12: 81-91.

Swales, J., & A. Burke (2003). "'It's really fascinating work': Differences in evaluative adjectives across academic registers" in P. Leistyna & C.F. Meyer (eds.), *Corpus Analysis: Language Structure and Language Use*, 1-18. Amsterdam and Atlanta: Rodopi.

Tang, K. (2013). "Instantiation of multimodal semiotic systems in science classroom discourse". *Language Sciences* 37: 22-35. URL: http://dx.doi.org/10.1016/j.langsci.2012.08.003) [03/02/2017]

Tickle-Degnen, L., & R. Rosenthal (1990). "The nature of rapport and its nonverbal correlates". *Psychological Inquiry* 1,4: 285-293.URL: http://www.jstor.org/stable/1449345. [03/02/2017]

Valeiras-Jurado, J. & Ruiz-Madrid, N. (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. Fortanet-Gómez (eds.), *Multimodal Analysis in Academic Settings. From Research to Teaching*, 108-130. London: Routledge.

Valeiras-Jurado, J., Ruiz-Madrid, N. & Jacobs, G. (2018) "Revisiting persuasion in oral academic and professional genres. Towards a methodological framework for multimodal discourse analysis of research dissemination talks". *Iberica* 35: 93-118.

Wulff, S., J. Swales, & K. Keller (2009). "'We have about seven minutes for questions': The discussion sessions from a specialized conference". *English for Specific Purposes* 28,2: 79-92.

Julia Valeiras-Jurado is a teaching assistant for Business and Scientific English at Ghent University in Belgium. She holds a joint PhD awarded by Universitat Jaume I and Ghent University. Her research interests are Multimodality, academic and business discourse, oral genres, persuasive language and ESP. She has published on these areas in the Routledge volume *Multimodal Analysis in Academic Settings. From Research to Teaching* (2015) as well as in a recent issue of *Ibérica*.

Appendix 1

Tone unit
Tone unit= onset (first prominent syllable) + tonic syllable (last prominent syllable).

Onset and tonic syllables are capitalised.

Tone: pitch movement in tone unit	Key: relative pitch of onset syllable	Termination: relative pitch of tonic syllable
Fall	High key: superscript	High termination: superscript
Level	Mid key: normal font	Mid termination: normal font
Rise	Low key: subscript	Low termination: subscript

Example:
which TELLS you which in^{GRE}dients you are going to use ◀

This is a tone unit with mid key, high termination and fall tone.

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		