

Research Article

Julia Valeiras-Jurado* and Teresa Morell

How do presenters engage with their audience? Speakers' multimodal interpersonal behaviour in research dissemination talks

<https://doi.org/10.1515/mc-2019-0014>

Abstract: Speakers in research dissemination talks are challenged with the need to connect with an audience that does not necessarily share their knowledge and expertise. This communicative situation can be particularly challenging for speakers using English both as a foreign language and for academic purposes. This study combines multimodal and ethnographic methods to explore how speakers of dissemination talks engage with their public. It focuses on four presenters' use and combination of language, paralanguage, kinesics, proxemics and gaze during intensive moments of engagement. The results show that these interpersonal rich points consist of dense multimodal ensembles that serve to shorten the distance between presenters and their audiences. The findings suggest that a skilful orchestration of modes can be greatly beneficial to achieve the desired level of audience engagement. Therefore, developing speakers' multimodal communicative competence should be a priority in English for specific and academic purposes (ESP/EAP) training.

Keywords: dissemination talks, interpersonal meaning, engagement, multimodal communicative competence, multimodal ensembles

1 Introduction

Making the connection with one's audience has been a major concern of speakers for ages. As far back as the 5th Century BC, classical rhetoricians already sought to formulate principles and techniques to create effective, credible and convincing speeches (Alcaraz Varó 2000, p. 156). In the past century conversational analysts, such as Atkinson (1984) began to watch and listen to politicians' language and body language to learn about persuasive communication techniques. However, until recently little attention has been given to speakers' actions and (inter)actions (Norris 2019, 2020) or presenters' combined use of other modes together with words. Fortunately, Social Semiotics derived from Halliday's Systemic Functional Linguistics, which studies the ways of meaning-making in specific social and cultural contexts, has acknowledged communication to be inherently multimodal (O'Halloran 2011, p. 123). Speakers naturally communicate by means of spoken language in combination with, at least, non-verbal modulations of voice, positions in space, gaze and gestures. Consequently, exploring how presenters engage with their audiences should take into account the combinations of semiotic resources which serve to shorten the distance between speakers and their public.

Ever since Kress and van Leeuwen (1996) provided the foundations for multimodal research there has been an increased interest among linguists to explore the integration of language with other modes in diverse contexts and for specific purposes. Work has been done on spoken discourse together with, for example, intonation (Halliday

*Corresponding author: Julia Valeiras-Jurado, Department of Linguistics, Faculty of Arts and Philosophy, Ghent University, Ghent, Belgium, E-mail: julia.valeirasjurado@ugent.be

Teresa Morell, Universidad de Alicante, Facultad de Filosofía y Letras, Alicante, Spain

and Greaves 2008), proxemics (Lim et al. 2012), or gestures (Kendon 2004). In addition, others have analysed speakers' effective use and combination of semiotic resources (multimodal ensembles) to support comprehension (Norte Fernández-Pacheco 2018), interaction (Morell 2004), amusement (Ruiz-Madrid and Fortanet-Gómez 2015), persuasion (Valeiras-Jurado et al. 2018), solidarity (Hood and Forey 2005) and engagement (Forey and Feng 2016). These Multimodal Discourse Analysis (MDA) studies, informed by Systemic Functional Linguistics (SFL), focus on the overall communicative effect of the multimodal ensembles rather than on the individual modes. They attempt to demonstrate how a skilful orchestration of modes helps presenters convey ideational (experiential and logical), textual (organisational) and interpersonal (engagement/involvement) meaning.

Conveying interpersonal meaning seems to constitute the most important and challenging metafunction in dissemination talks, especially when English is not the mother tongue. Speakers can have a good control of English and be able to organise their presentations logically, but this does not imply that they will be able to engage their layman audience. According to Forey and Feng (2016, p. 419), "Engagement is a key resource the speaker adopts to develop a positive relationship with the audience throughout the talk in order to strengthen his/her argument." Speakers who engage their audience manage to shorten the distance between themselves, the public and the topic being mediated. This is done through a complex process of multimodal communication that entails spoken and written words, intonation and gestures (Hargie 2011). However, understanding speakers' combined use of modes for audience engagement calls for a multimodal analysis of the moments in which interpersonal meaning-making is foregrounded.

The main objective of this study is to explore how speakers of dissemination talks using English as a foreign language engage with their audience, by analysing the multimodal ensembles that are instantiated in their rich points or intensive moments of engagement. These rich points of engagement can be compared to Norris' (2020, p. 70) "higher level mediated actions" in which speakers make use of different (inter)actions at the same time, realized through different semiotic modes, and with varying degrees of attention and awareness. For the purpose of this study, we make use of MDA and ethnographic tools to compare presenters' use and combination of embodied modes such as language, paralinguistics, kinesics, proxemics and gaze to convey interpersonal meaning. The rationale behind this choice is that they tend to receive less attention by researchers, language teachers, and speakers themselves. Our ultimate goal is not to develop a grammar of embodied modes, but rather to develop critical understanding among ESP and EAP teachers and students (Ledin and Machin 2019) of the effect of multimodal ensembles in the specific context of dissemination talks.

2 Methodology

In this study we have analysed the multimodal behaviour of speakers in research dissemination talks, combining Multimodal Discourse Analysis (MDA) and ethnographic methods. The presentations were part of a university contest in which six researchers talked about their studies to a layman audience. After obtaining consent from the speakers, the event was recorded using one camera that focused on the speakers and followed their movements. The presenters were doctoral and post-doctoral students. All of them were non-native fluent speakers of English and were experienced in presenting their research. The event took place in a large room equipped with a podium, a computer, a projector, a screen, a lectern and a desk. The speakers did not stand still on the podium, but came down and moved around the front part of the room to be closer to the audience. They all wore lapel microphones. Four of them controlled their presentations with a pointer, and one used no visual aids at all. From the six presentations, one was discarded due to poor quality of the image of the recording. The total size of this dataset is 45 min.

During the event, observation sheets were used to take note of relevant information for the subsequent analysis of the speakers' multimodal behaviour. Among the annotations were included contextual details regarding the event (e.g. physical location, the order of each presentation within the event, the setting, props available, size and type of audience, etc.), and the speakers (name, age, gender, mother tongue, presenting style, position on stage, props used, etc.). For the sake of consistency, all observation sheets were filled in by the same researcher.

After the event the speakers were interviewed in face to face, semi-structured interviews. These interviews probed into aspects that could potentially influence the speakers' strategies to connect with their audience, such as their experience presenting, their motivation to participate in the event, previous knowledge about the event, their prior preparation, their satisfaction with their performance, and their opinion about other presentations in the event. All interviews were carried out by the same researcher.

Since the focus of this study lies on interpersonal meaning and how it is conveyed multimodally in oral presentations, the videos were visualized to select a number of rich points in which the speakers were perceived to be making an effort to connect with their audience. This was an inductive approach to video-based analysis (Goldman et al. 2007), aided by the data gathered from the observations and interviews. The two researchers involved in the study made an initial, individual selection which was later compared. This comparison revealed clear agreement regarding four rich points. These rich points, which will be discussed in Section 3, stood out in interpersonal meaning, and at the same time illustrated different strategies adopted by speakers to establish a relationship with their audiences. For these reasons they were selected for fine-grained analysis, regardless of the presentation in which they occur. Although this selection process resulted in four rich points distributed in four different presentations, it must be noted that this distribution is not intentional but purely coincidental. This selection process did not involve any external researcher, but we believe this does not invalidate the scientific reliability of the study for several reasons. First, the selection is mainly informed by ethnographic data gathered through interviews and provided by the speakers rather than the researchers' opinion. Second, the results are shared and contrasted with the speakers, so once more they are not just the researchers' interpretations. Thirdly, the inductive approach to video data applied in this study is not only common practise, but in fact recommended when working with

'raw' video data sets that have been collected with broad questions in mind but without a strong orienting theory. The process is usually to view all of the video data repeatedly and in increasing depth where the research team agree on major events, themes and identify key moments of importance and to describe the structure of the event (Jewitt 2012).

The repeated view of the videos made it apparent that there were more interactive verbal and non-verbal markers in the rich points than in the surroundings. Therefore, much like is explained by Norris (2019: 194), we narrowed the site of engagement and selected the data, i.e. the rich points, for micro analysis. The use of rich points was adopted to avoid prioritizing any semiotic mode in particular, so as not to use one mode as the driver of the analysis. In this sense, the approach has proved useful to keep the focus on the multimodal ensemble and the way different modes interact to convey interpersonal meaning. The rich points in the dissemination talks add up to 4.9 min. The size of this dataset does not allow for quantitatively-based generalisations, but it is valid for qualitative analysis. It is also in line with previous multimodal studies, which due to their minute level of detail and the lack of automating tools cannot afford the use of larger corpora (Beltrán-Planques and Querol-Julián 2018; Forey and Feng 2016; Morell 2015; Querol-Julián 2011; Valeiras-Jurado et al. 2018). These studies, though modest in size, are nonetheless extremely useful because they provide preliminary results that can pave the way for further research.

The multimodal analysis presented here caters for the following modes: language, paralinguistic, kinesics, proxemics and gaze. Language refers to speakers' choices in terms of lexis, grammar, style and register. These choices include evaluative language, intensifying language, pronouns, examples, questions, imperatives, humour, informal language and hedging among others (Carter 1997; Hyland 2009; Lakoff 1982; Morell 2004), all of which, according to the literature, can have a significant effect on the relationship established with the audience. In terms of language, the unit of analysis used in this study is the proposition, which can be defined as the meaning of an utterance that remains constant despite changes in voice qualities (i.e. paralinguistic aspects).

According to Poyatos (1983) language, paralinguistic and kinesics are closely intertwined. Poyatos (1983, p. 129) defines paralinguistic as "a series of vocal/narial voice modifications and independent sounds and meaningful silences", which together with kinesics is what "truly gives the spoken words their total meaning." Kinesics is in turn defined as "the systematic study of psychomuscularly-based body movements (gestures and manners)

and/or their resulting positions (postures) (...) that, whether isolated or combined with the linguistic-paralinguistic structures and with the situational context, possess communicative value (...). For the purposes of this study, interest is given to two kinesic aspects: gestures and head movements; and two paralinguistic aspects: intonation and pauses.

Intonation is frequently defined as a pitch contour that, according to Discourse Intonation (DI), reveals speakers' assumptions about the interaction in the communicative process (Brazil 1997). As explained in Valeiras-Jurado et al. (2018), discourse intonation divides speech into tone units, each one containing one or two prominent syllables. These syllables are louder and longer (in the transcripts they are capitalized). The number of prominences and the distance between them can also provide an indication of the pace of delivery or rhythm. The first prominent syllable in a tone unit is called onset, the second is called tonic syllable. Tone is the pitch movement that begins in the tonic syllable and continues until the end of the tone unit. In the transcripts tone is represented with arrows. Key and termination refer to the relative pitch of onset syllables and tonic syllables. A high key and termination is represented as superscript, and low key and termination as subscript. DI does not consider pauses as boundary markers of tone units; however, they are meaningful in that they affect the pace of delivery, the way the audience perceives the message, and the way they are expected to react accordingly. The location and duration of these pauses is indicated in the transcripts within parenthesis. For the purpose of this study, the tone unit includes the proposition together with illocutionary meaning added by speakers' intonational choices. The following example is a proposition and a tone unit pronounced with fall tone and high termination, and with a 2 s pause following:

is your ^{PHONE} smart ↘ (2)

Kendon (2004, p. 7) defines gesture as “visible action when it is used as an utterance or as part of an utterance.” In this study manipulations of objects (i.e. a phone) are included within gestures. The focus is on the type, function and family of the gesture (Bavelas et al. 1995; Kendon 2004; McNeill 1992; Querol-Julián 2011). The following gesture types are considered: (a) iconic (represent concrete objects and events), (b) metaphoric (represent abstract ideas), (c) beats (repetitive gestures that usually mark the discourse flow) and (d) deictic (point to something). Likewise, the following functions are relevant for the present study: (a) referential (they are part of the referential content), (b) pragmatic (they show the attitude of the speaker towards the content and indicate how content is to be interpreted), (c) interpersonal (they regulate interaction) and (d) cohesive gestures (they connect thematically related but temporally separated parts of discourse). Kendon's gesture families are also applied in the analysis of gestures. They refer to gestures with similar kinesic characteristics that seem to share a common semantic theme. For example the open hand supine (OHS) family includes gestures made with the palm of the hand facing up, and share the semantic theme of offering or willingness to receive something.

Regarding head movements, nods, shakes and lateral movements are taken into account, in particular the amplitude of the movement and the number of repetitions that provide insights into pragmatic meaning (Hadar et al. 1983). In terms of proxemics, in line with Lim et al. (2012), social distance created by speakers' position on stage is focused on. Finally, in the case of gaze, the degree of audience engagement is determined by examining whether the speaker is looking at the audience, the screen, or the floor (Forey and Feng 2016).

The multimodal analysis presented in this article entailed the use of specialised software. The analysis of intonation was aided by the tool for phonetic analysis PRAAT,¹ which provided accurate measurements of pitch and loudness. The analysis of gestures and head movements was assisted by the annotation tool ELAN,² which allowed classifying and time-aligning the occurrence of these gestures and head movements with the other modes. In the case of language, the presentations were transcribed orthographically, and this transcript was time-aligned with the rest of the modes. Position on stage and gaze were mapped, time aligned with the other modes and represented in diagrams.

¹ <http://www.fon.hum.uva.nl/praat>.

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

As a last step in the analysis, results were contrasted with speakers through a second round of interviews. In this case, a playback methodology was used (Norris 2011, p. 59). Three of these interviews took place face to face and two of them were online video calls. All of them were recorded. These interviews did not include a predefined battery of questions, but all of them included the following steps:

- i) Explanation of the content and purpose of the interview;
- ii) Visualisation of the excerpts analysed with the speakers;
- iii) Discussion of aspects the analysis had revealed as relevant for the interpersonal behaviour (e.g. a particular use of intonation). First the speakers' opinion was prompted, followed by an exchange of interpretations regarding the intent and potential effect of these aspects.

Occasionally, speakers offered alternative interpretations, which were always integrated in the discussion of results. This approach was useful not only to triangulate results, but also to take a step back from the micro-analysis of modes and focus on the multimodal ensemble as a whole.

In Section 3, the results of the multimodal and ethnographic analysis of the ensembles found in the four rich points included in the study are discussed and then compared.

3 Results and discussion

The results of the analysis show that the interpersonal rich points identified in the corpus of dissemination talks are all characterized by a shortened distance between the speakers and the audience, and an increased level of audience engagement. This finding agrees with previous literature on interpersonal communication in academic discourse. Forey and Feng (2016, p. 419) find that speakers in academic presentations use engagement as a key resource to develop a positive relationship with the audience and strengthen arguments. They also point out that a shortened social distance is frequently mirrored by a shortened physical distance between speaker and audience. Morell (2018) also shows how the use of multimodal ensembles (including among other modes the use of classroom space) can increase engagement in lectures. In line with these studies, our results show that the interpersonal behaviour of speakers in dissemination talks also involves a combination of different modes (i.e. is multimodal). However, speakers can resort to different interpersonal strategies in each case, as the four examples discussed in this section will illustrate.

3.1 Example 1: “kiss you” (establishing a close connection with the audience)

In this example the speaker is a chemist. The title of the presentation is “*Microwave chemistry: Time is money*”, and the topic is the production of nanoparticles and how it can be accelerated using a microwave oven. He begins his presentation playing with the expression “*keep it simple, but not stupid*” and its acronym KISS to engage with the audience.

Table 1 offers the orthographic and DI transcription, and Figure 1 shows the speaker's position on stage and gaze direction. A full account of the multimodal ensembles in this example is offered in Appendix A.

The speaker gradually comes closer to the audience while gazing and pointing at them with an extended finger (Figure 1. Positions 1 and 2). Using Lim's terms (Lim et al. 2012), he is occupying what in a classroom setting can be described as an “interactional space.” Then, keeping the same pointing gesture, he comes even closer (Figure 1. Position 3) and points at specific members of the audience (Figure 2).

Simultaneously, he also addresses them with the expression “*kiss you.*” The speaker explained during the interview that he wanted to create “a very personal connection” with the audience, and he is doing this using his position on stage, his gaze, his gestures and language.

As he keeps addressing the audience, this time collectively saying “*kiss all of you*”, he represents the scope of the whole room with an open hand prone (OHP) sweeping gesture of one of his hands (palm down moving from right to left and simultaneous beating, Figure 3). This gesture can also prevent potential counterclaims to the intensified quantification (“*all of you*”).

Table 1: Example 1 transcriptions.**Orthographic transcription**

Hi. I'm a chemist, and I really like kissing. I'd like to kiss you, I'd love to kiss you, I'm gonna kiss all of you. Because I really want to keep it simple, but not stupid. Tonight I'm going to show you how your own experience, with a very simple technique, can revolutionise the chemical industry, and finally your life.

DI transcription

1. HI
2. I'm a CHEmist
3. And i REally like Klssing
4. i'd LIKE to KISS you
5. i'd LOVE to KISS you
6. i'm GONna kiss ALL of you
7. beCAUSE (11)
8. i REALLY want to (1)
9. KEEP it SIMple (0.5)
10. but NOT stupid
11. toNIGHT
12. i'm going to SHOW you how your OWN experience
13. with a VEry simple technique
14. can revolUtionise the CHEmical industry
15. and Finally YOUR life

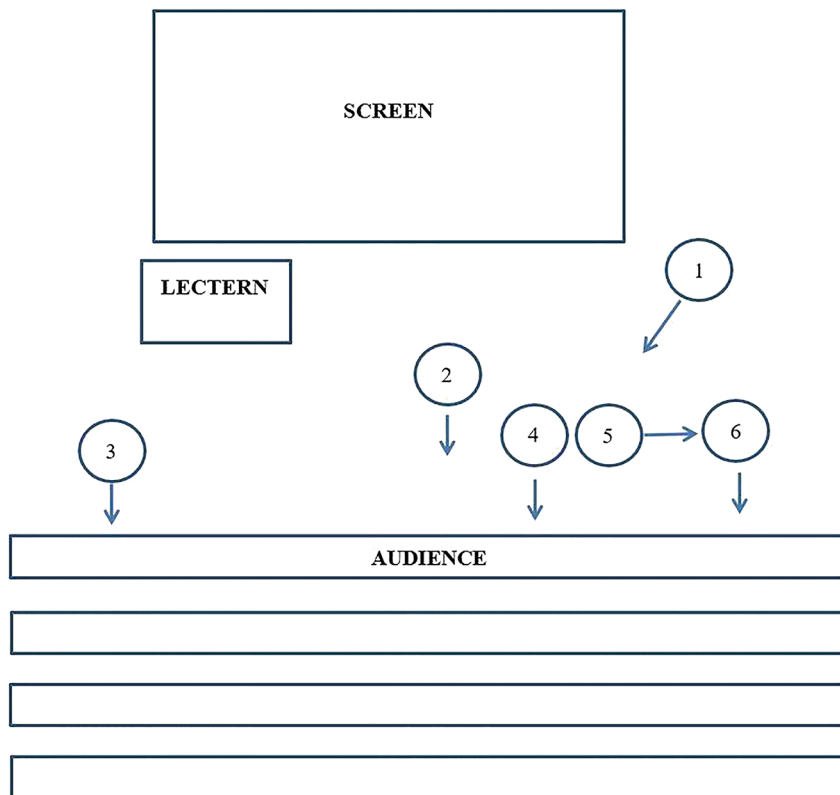


Figure 1: Position on stage and gaze direction in Example 1. Numbers represent speaker's position and arrows indicate gaze direction.



Figure 2: Pointing at the audience in Example 1.



Figure 3: Sweeping gesture with one hand in Example 1.



Figure 4: Closed-fist beats in Example 1.

Throughout his presentation the speaker makes repeated closed-fist beats that tend to be synchronous with prominences in words like “*really*” or “*revolutionise*” that carry an intense meaning. This is the case as he says “*I really like kissing*” (Figure 4). According to him, these gestures show emotional involvement, which can be interpreted as a sign of interpersonal engagement.

Once the speaker gets the attention of the audience using the word “*kiss*” and pointing at them, he moves on to explain the meaning of the acronym KISS. At this point he temporarily disengages from the audience to plan his message. This is shown proxemically by gradually moving away from the audience and orienting his body and gaze away from them (Figure 1, Positions 4 and 5, momentarily looking at the floor). His rhythm also slows down at this point (Table 1, units 7–10). As he utters unit 10 (“*but not stupid*”) he also makes a beating gesture with an extended finger (see Figure 5). This seems to indicate that the speaker has finished the encoding of his message (i.e. he has found the right words) and is orienting back to the audience.



Figure 5: Beating with extended finger in Example 1.

Full engagement with the audience takes place again as the speaker gazes and points at them using both open hands (Figure 1. Position 6. and Figure 6). He reported that with this way of pointing he was consciously conveying honesty. The gestures are synchronous with prominences in “*your own experience*” and “*your life*”, clarifying with language and paralinguistics who these pointing gestures refer to (=you, the audience).

A second sweeping OHP gesture synchronous with “*very simple*” serves to prevent a potential counter-claim to the intensified evaluation (Figure 7).

The speaker is anticipating reactions, a sign of being interpersonally engaged with the audience. Although he claimed in the interview that he tends to associate this gesture with the concept of “*simple*”, on second thoughts he became aware of the fact that he also uses this gesture with intensified evaluations that imply other qualities as well (as illustrated by his use of the same gesture earlier with “*all of you*”).



Figure 6: Pointing at the audience with open hands in Example 1.



Figure 7: Sweeping gesture with both hands in Example 1.

In short, language, paralanguage, gaze, kinesics and proxemics are used in this example to establish a connection with the audience and to anticipate their responses. Our findings suggest that this connection with the audience is made possible not thanks to each separate mode, but rather through a skilful orchestration of these modes into a series of multimodal ensembles. Because of this, a coherent use of modes is a skill that should occupy a central role in ESP and EAP courses teaching academic presentations.

3.2 Example 2: “dinosaur traps” (use of humour)

In this example the speaker is a mathematician, and his talk is entitled “*Why does mathematics count?*” In the rich point selected, he highlights the relevance and applicability of this area of research.

The transcripts and changes in position and gaze are shown in Table 2 and Figure 8, respectively. Further details about the multimodal ensembles in this rich point are shown in Appendix B.

Table 2: Example 2 transcriptions.

Orthographic transcription	
We use mathematics to build bridges. To build dinosaur traps, you know, and other things we use on a daily basis.	
DI transcription	
1.	we USE mathematics to build BRIDges →
2.	ah to BUILD eh DInosaur traps ↗
3.	you KNOW ↗
4.	and ^{0th*} things we use on a daily BAis ↗

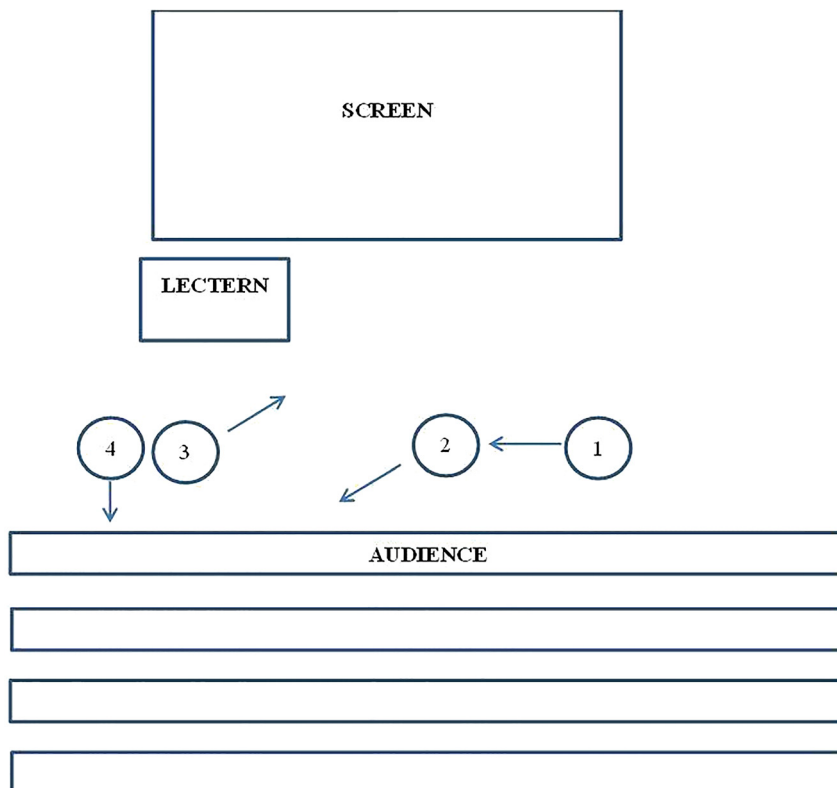


Figure 8: Position on stage and gaze direction in Example 2.

The excerpt is characterized by a prominent use of humour. This is noticeable in language, specifically in the example of “*dinosaur traps*” as a daily routine. This example gets the attention of the audience and builds rapport. The informal expression “*you know*” is also a way of building rapport and seeking concurrence. But the humour is not created solely with language, but in multimodal ensembles combining language, para-language, kinesics, proxemics and gaze, as we will illustrate in the following paragraphs.

As the excerpt begins, the speaker is located in a relatively central position on stage (Figure 8. Position 1) but immediately starts moving towards the left side. As he moves he gazes straight to the front, but alternates this gaze direction with some looks at the audience (Figure 8. Position 2).

The first tone unit that the speaker uses (“we USE mathematics to build BRIDges”) has a fall, proclaiming tone suggesting new information, but the rest of the units tend to a rise referring tone that matches the intention of giving familiar examples. This referring tone at the end attempts to assume common ground and creates rapport, and is consistent with the expression “*you know*.” While discussing his use of intonation in this extract during the second interview, the speaker interpreted his use of tone in a more cohesive way: the initial fall tone was his way of showing that this is the beginning of a list, and not a way of projecting new information. As the speaker utters this first unit, he is rotating his hands (Figure 9) in a metaphoric, referential gesture that seems to symbolise the process of physically using something (in this case mathematics).



Figure 9: Rotating hands in Example 2.

However, the speaker offered a different interpretation of this gesture during the interview, which is the cohesive function of showing that there is a longer list of things coming. This is probably a more plausible interpretation, taking into account the next gesture: enumerating with fingers as he actually begins the list (Figure 10).

The high key in unit 4 (“^{Other} things we do on a daily BASIS”) presents this information as surprising or denying expectations, and it prepares for the final humorous pun that is achieved by the mismatch between content and tone (an extravagant example presented as daily basis). Interestingly, the pun is preceded by a temporary change in body orientation: the speaker momentarily looks at the screen as if looking for a cue to continue his list (Figure 8. Position 3), which provides spontaneity and contributes to the humour.

The same effect is achieved by a series of beats with open hands palms up starting with “*dinosaur traps*” and continuing until the end of the extract (Figure 11).



Figure 10: Enumerating in Example 2.



Figure 11: Beats with open hands in Example 2.

These beats can be interpreted as an encoding aid for the speaker. They create the impression that the speaker is improvising his list of examples and provide the text with spontaneity.

In short, this speaker achieves humour multimodally through language, paralinguistics, gaze, kinesics and proxemics. Interestingly, this finding is very much in line with previous research on the use of multimodal humour in academia. Fortanet-Gómez and Ruiz-Madrid (2016) find similar multimodal ensembles in humorous episodes in plenary lectures, including paralinguistic aspects (words with exaggerated meanings, intonation, pauses and stress) and kinesic aspects (face expressions). This leads them to conclude that most of the humour in their corpus would be very difficult to understand without the non-verbal resources. This is also the case in

Table 3: Example 3 transcriptions.

Orthographic transcription
Are you smart? I knew that this would come, but let me rephrase: Is your phone smart? You would say of course it is a smart phone so why do you ask. Well then I say, is the clothes you wear smart? The jeans the jacket, the clothes that you have...and it would be very silent.
DI transcription
1. ARE YOU SMART (3) 2. i KNEW thatthis would come 3. but LET me rePHRASE 4. is your PHONE smart → (2) 5. you would say of COURSE it is a smartPHONE 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 thatyouhave 12. andit would be VERY SILENT →

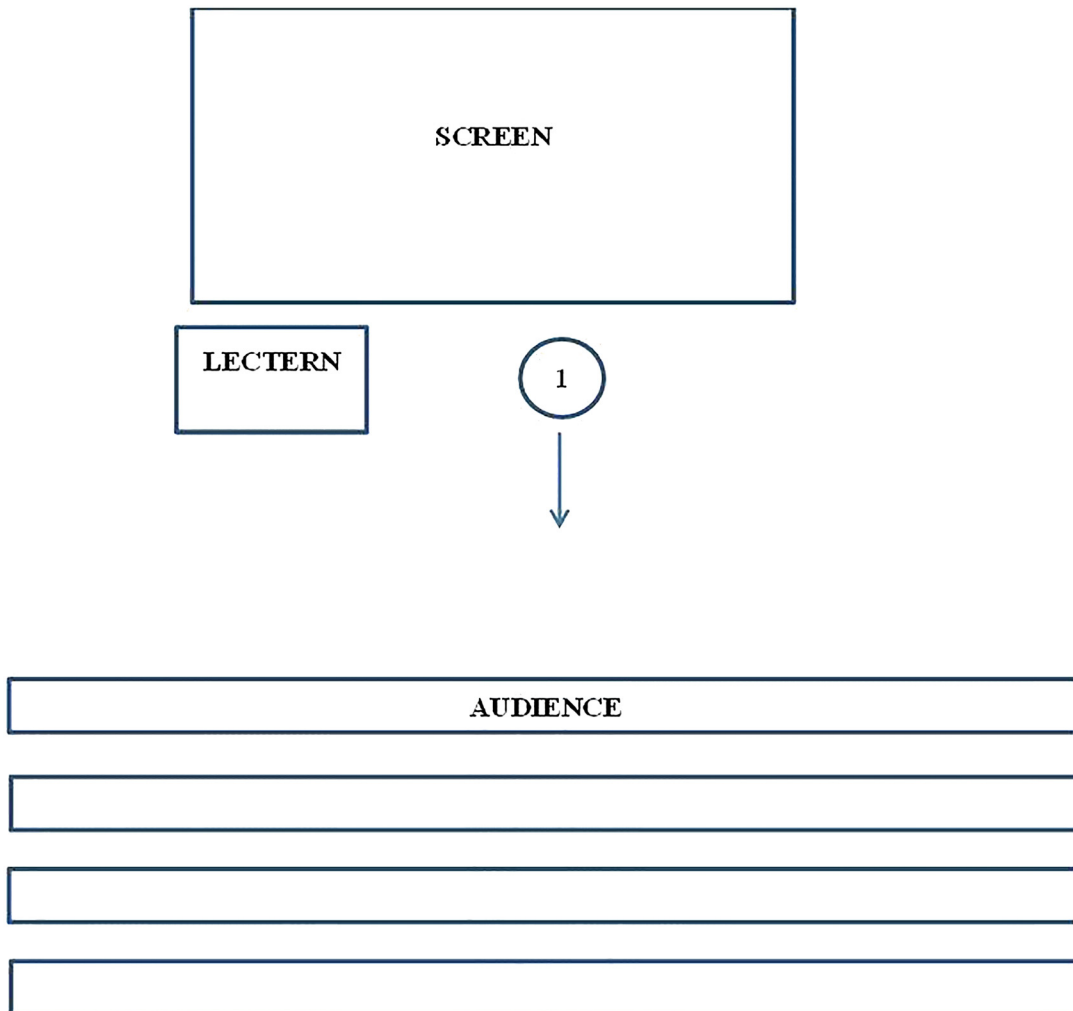


Figure 12: Position on stage and gaze direction in Example 3.



Figure 13: Open hand gesture inviting answer in Example 3.



Figure 14: Open hand gesture showing anticipation in Example 3.

our example, in which a carefully orchestrated multimodal ensemble is used to effectively convey interpersonal meaning.

3.3 Example 3: “I knew that this would come” (prompting and anticipating reactions)

In this example, we see the beginning of a presentation about smart textiles. The title is “*Snow White’s smart textile twist.*” In the rich point selected, the speaker introduces the concept of “*smart textiles*” by applying the adjective “*smart*” first to the people in the audience, then to their phones, and finally to their clothes. The transcripts can be seen in Table 3. Proxemics and gaze are represented in Figure 12, and a full list of multimodal ensembles is offered in Appendix C.

The speaker starts her presentation from a central position on stage (Figure 12. Position 1), gazing at the audience, but from a relatively distant position which could be described as the authoritative space (Lim et al. 2012). However, the speaker closes this distance with the audience by asking questions and inviting an answer

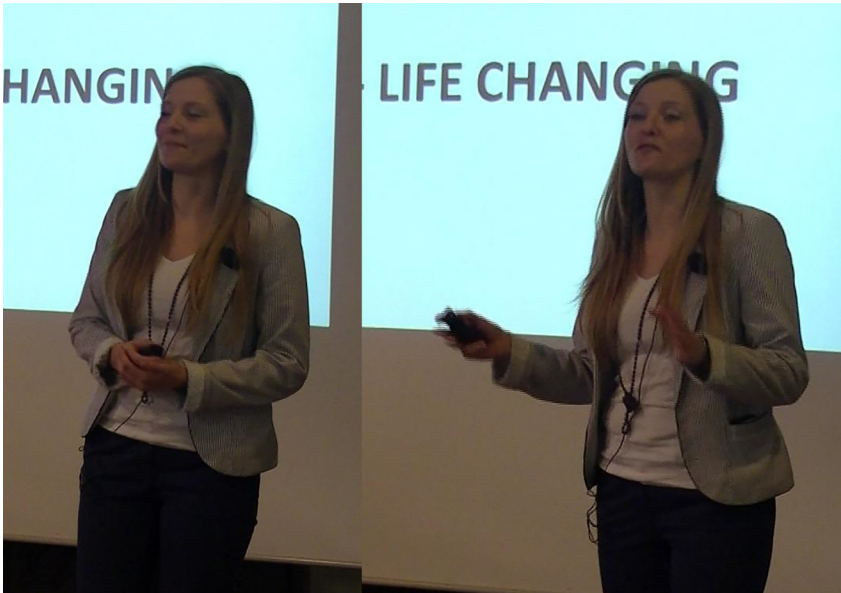


Figure 15: (a and b) Rephrasing with beats in Example 3.



Figure 16: (a and b) Taking out and holding the phone in Example 3.

with open hand gestures in combination with prominences (Figure 13). This is an interpretation that the speaker corroborated in the follow-up interview. She also makes two pauses (see DI transcript) to give the audience the opportunity to answer and makes a lateral movement of the head to shift her gaze and keep eye-contact with the whole audience while waiting for their answer. This, together with a fall tone in her next two questions, reinforces the effect of prompting an answer from them (the audience feels it is not just a rhetorical question, but a real one).

The speaker also anticipates these answers that she is inviting, which shows active engagement with the audience. This is noticeable in her words (“*I knew that this would come*”), her head nods and her gestures, for example an open hand supine (OHS) beat opening and closing arms twice (Figure 14). The head nods are a way of confirming that her anticipation was right (as the speaker explained in the interview), and the gesture implies a sense of obvious that reinforces the meaning expressed through language.

The speaker keeps prompting reactions from the audience as she says “*let me rephrase.*” She mirrors the meaning of these words kinesically through an OHP pragmatic beat holding hands together briefly and then opening arms (Figure 15 a and b).



Figure 17: Open hand beats holding phone in Example 3.



Figure 18: Putting back the phone in pocket in Example 3.

This rephrasing actually implies redirecting the concept “*smart*” to “*phones*”, as in the common expression “*smart phone*.” The speaker shows this visually taking out her phone from her pocket and holding it up for the audience to see (Figure 16 a and b).

The speaker even formulates the answer that she expects from the audience, momentarily assuming their role, and even enacting their reaction at her question by using a tone that implies boredom or annoyance (because they consider the question stupid). She does this explicitly in words (“*of course ... so why do you ask*”), in combination with OHS beats still holding the phone that are highly synchronous with prominent syllables, and with a low pitch (low key and termination) that presents what is being said as obvious (Figure 17).

Before the speaker shifts topic to clothes, she puts back the phone in her pocket (Figure 18).



Figure 19: (a and b) Pointing at clothes in Example 3.



Figure 20: Sweeping gesture in Example 3.

The focus on clothes is once more visually represented with a series of deictic gestures to her trousers and jacket (Figure 19 a and b).

Towards the end of the excerpt, the speaker once more anticipates reactions from the audience and uses an OHP sweeping gesture to counteract any potential rebuttals of her intensified evaluation “*very silent.*” Interestingly, the gesture is split in two parts to follow the rhythm created by the two prominences in “*very silent.*” This rhythm is also reinforced by subtle head nods synchronous with the prominences (Figure 20). This constitutes a coordinated use of words, gestures, head movements and prominence previously noted in Valeiras-Jurado et al. (2018).

Table 4: Example 4 transcriptions.

I must say all of a sudden I feel really, really humble. I've, just, These these five presentations that came before me they all want to change the world and they are doing it in an incredible way, they are doing superconductors in microwaves, there are things I've never heard of textiles, it's, it's, I'm just trying to get you from A to B, that's all I want to do. It's, it's, it's not rocket science, it's not superconductors, it's just trying to inform you, and this is really difficult.	
DI transcription	
1.	i must SAY ALL of a sudden
2.	i feel REALLY REALLY HUMble ↘
3.	i've just
4.	these these FIVE presentations that came beFORE me
5.	they all want to CHANGE the WORLD
6.	and they are DOING it in an in an inCREDible way ↘
7.	they are doing SUpErconductors in MICrowaves
8.	there are THINGS i've never HEARD of textiles
9.	it's it's
10.	i'm JUST trying to GET you from a to b ↘
11.	that's ALL i want to do ↘
12.	it's it's
13.	it's NOT ROcket science ↗
14.	it's not SUpErconductors ↗
15.	it's JUST trying to inFORM you ↘
16.	and THIS is really DIfFicult

In short, this example shows how the speaker prompts and anticipates answers through language, paralinguage, gaze, gestures, head movements and proxemics, once more showing the crucial importance of a coherent use of modes to establish the desired connection with the audience in academic presentations.

3.4 Example 4: “I feel really, really humble” (likeability through humbleness)

This example shows the beginning of a presentation about transport data systems. The title is “*Open transport data*”, and the talk discusses the advantages and technological challenges of sharing this type of data. In the rich point selected the speaker refers to the previous presentations in the event (his talk was the last one) and compares his own research to the previous ones in humble terms. The transcripts can be found in Table 4. Position and gaze are shown in Figure 21, and the list of multimodal ensembles can be found in Appendix D.

This rich point is characterized by a sustained use of different beats. The speaker uses beats with open hands (that have an emphasizing, pragmatic effect), beats with finger rings (that have a specifying effect, as in Figure 22) and lateral beats that show movement, as in Figure 23). The speaker begins with full orientation towards the audience from an authoritative space in the centre (Figure 21. Position 1), but closes distance moving to one side of the stage and closer to the audience (Figure 21. Position 2). At the same time, he introduces the idea of humbleness through language, paralinguage (prominences and fall tone that convey sincerity) and gestures (e.g. beats). For example, he explicitly uses the adjective “*humble*” and modifies it twice with the adverb “*really*.” This repetition in “*really, really humble*” is paralleled with three finger ring beats that

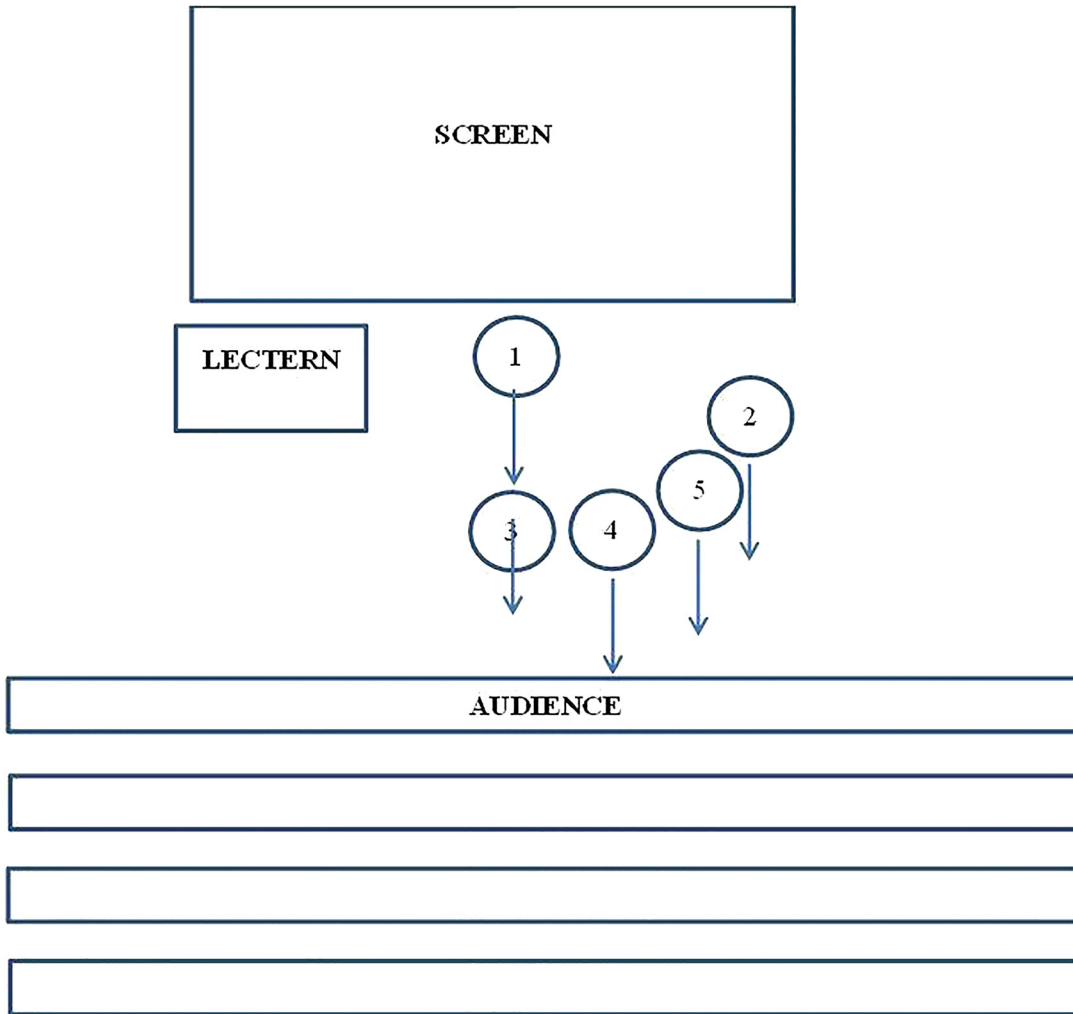


Figure 21: Position on stage and gaze direction in Example 4.

mark the rhythm created by the prominences (Figure 22). The function is therefore partly cohesive (connecting) and partly pragmatic (emphasizing, with the effect of adding sincerity; and specifying, as if saying ‘I really mean humble’).

The speaker then moves towards a more centred position on stage (Figure 21. Position 3) as he reflects on the presentations that came before him, and evaluates them in positive terms with an intensified evaluation: “*an incredible way.*” Interestingly, this is accompanied by subtle head shakes that seem to anticipate and rebate counterclaims to this evaluation (as if saying ‘believe me, it is really incredible’). Immediately after there is momentary disengagement from the audience as the speaker provides more details about the previous presentations, which is noticeable in the way he occasionally gazes away and looks at the floor. This is probably due to the effort of encoding his message, since this is a part of the presentation that he could not have prepared in advance and has to improvise (he did not know about the other presentations in advance). He then resumes full engagement with the audience as he moves to the right and closer to them (Figure 21. Position 4) and finally moves back but still keeping eye contact (Figure 21. Position 5). These two last positions coincide with the introduction of his research topic.

As he introduces his research topic he uses the adverb of degree “*just*” and he equates his research objective (“*all I want to do*”) with something apparently unimportant (“*take you from A to B*”). Interestingly, the



Figure 22: Finger ring beats in Example 4.



Figure 23: Lateral beats in Example 4.

high key in “i’m^{JUST} trying to GET you from a to b” conveys a sense of surprise and unexpectedness, and adds a humorous tone which is reinforced by lateral open hand beats that symbolize this movement from “a” to “b” (Figure 23).

An 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” (Figure 24) (Valeiras-Jurado et al. 2018).

Again, subtle head shakes simultaneous with “*from A to B*” and “*that’s all I want to do*” have the pragmatic function of emphasizing this idea of ‘there’s nothing else’.



Figure 24: Open hand palms up gesture in Example 4.

By humbly comparing himself to the other speakers in the event, the speaker is making himself likable and building rapport. According to the speaker, this is a technique he often resorts to when he is not well prepared for a talk: it is a “*captatio benevolentiae*” resource (in the speaker’s words) that shows interest for other presenters and their presentations and at the same time serves as an attention getting technique. What the analysis of this rich point reveals is that the technique depends on a series of tightly constructed multimodal ensembles for its effectiveness.

3.5 Comparative analysis of the four rich points

In sections 3.1–3.4, the four interpersonal rich points and their multimodal ensembles have been analysed separately. Here, we will proceed to compare their characteristics in terms of temporal and spatial intensity (see Table 5). In order to facilitate this comparison, a more quantitative approach is adopted in this part of the study. It must be noted that the aim is not to generalise from this quantitative data, but to compare the use of modes in the four rich points using objective, comparable criteria.

In terms of time, the rich points range from 7 to 30 s and on average each proposition lasts 0.58 s. In other words, every six-tenths of a second the speakers use a distinct multimodal ensemble to express a specific meaning. In all cases, the ensembles consist of spoken language, paralanguage, spatial positions and gaze. In most cases gestures also form part of this ensemble and head movements are occasionally added. It is not possible to establish trends or generic traits from this limited dataset, but we can safely conclude that these data hint at a prominent multimodal nature of dissemination talks, with a clearly dynamic multimodal behaviour of the speakers.

The linguistic analysis of the four rich points reveals a high concentration of varying interpersonal features, such as intensifying adverbs, humour, informal language, questions and hedging. However, the presence of personal pronouns in the majority of the propositions stands out as the most predominant tool implemented by the four speakers to connect with the audience. As indicated in Table 5, “*you*” is the most common one, something which was previously noted in Morell (2004) study on interactive discourse in lectures.

Figure 25 offers a visual representation of the use of interpersonal connections through language in the four presentations.

In terms of space, all the speakers move about in various positions all within what Lim et al. (2012) refers to as the authoritative space, where they can be seen and listened to by the entire audience. Nevertheless, the movements in Examples 1, 2 and 4 (lateral and/or closer to the audience), suggest that the speakers wish to

Table 5: Comparative analysis of the four rich points.

Duration (seconds)	Propositions	Interpersonal connections through language	Spatial positions	Gaze (audience vs. other)	Gestures (types function families)	Amount of gesturing/head movements
1 24"	15 0.63"/proposition	- Pronouns ('I' x 7, 'you' x 4, 'your' x 2) - Intensifications: 'really' x 2, 'revolutionise' - Humour: 'Kiss(ing)' x 4	6	21" = audience 3" = floor	- Deictics, beats and metaphoric - Pragmatic, referential, and interpersonal - Open hand prone (OHP), closed-fist, extended finger	23/24"
2 7"	4 0.57"/proposition	- Pronouns ('we' x 2, 'you' x 1) - Humour: 'dinosaur traps' - Informal language: 'you know'	4	3" = straight, screen 4" = audience	- Beats - Pragmatic, cohesive and referential - Open hand supine (OHS) rotating hands and enumerating with extended fingers	7/7"
3 24"	14 0.58"/proposition	- Pronouns ('you' x 5, 'I' x 2, 'your' x 1) - 3 questions to the audience: 'are you smart?' - Intensification: 'very' - Pronouns ('I' x 6, 'they' x 3, 'you' x 2, 'me' x 1)	1	24" = audience	- Beats, object manipulations, metaphoric - Pragmatic, referential, interpersonal - OHS, OHP	23/24"
4 30"	16 0.53"/proposition	- Intensification: 'really' x 2 - 'Incredible' - Hedging: 'just'	5	28" = audience 2" = floor	- Beats - Pragmatic, metaphoric - OHP, OHS, finger ring	27/30"

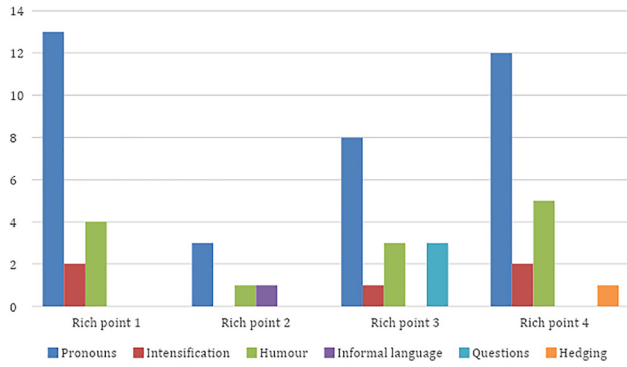


Figure 25: Interpersonal connections through language (number of occurrences).

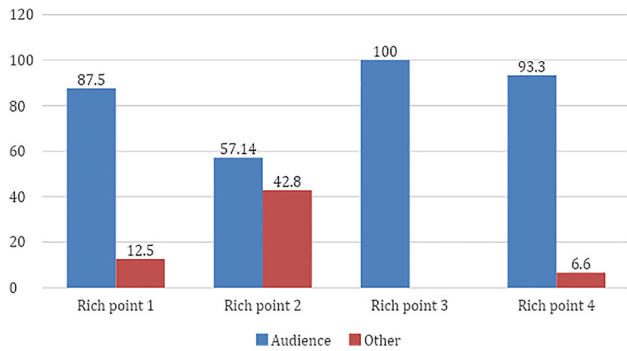


Figure 26: Gaze direction (percentage).

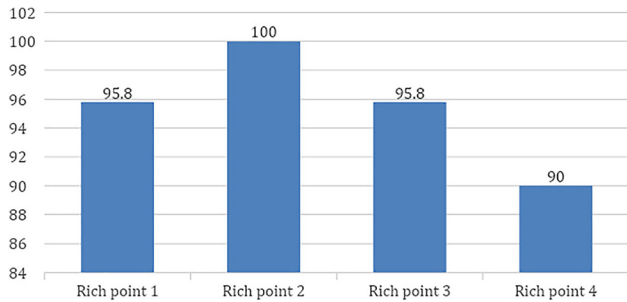


Figure 27: Gesturing (percentage).

shorten the distance with the public by going beyond the authoritative to the interactive space. This is an interesting preliminary finding worth confirming through larger scale studies.

Regarding intonation, Example 3 shows how pauses, fall tone, low key and termination can be used to engage the audience and to anticipate and prompt reactions from them. In Examples 2 and 4 the use of high key and referring tone creates a surprising and humorous effect which builds rapport.

As indicated in Figure 26, the amount of gaze addressed to the audience in all examples shows that it is a crucial tool to build and maintain an interpersonal relationship with the audience. In fact, the brief interruptions in eye contact are mainly caused by additional efforts in encoding the message (Examples 1, 2 and 4). They reveal temporary disengagement from the audience and focus on the message instead.

Figure 27 shows that the four speakers are gesturing most of the time. They make use of nearly a gesture per second (80 out of the total 85 s), which proves the relevance of gestures in establishing a connection with the audience.

Concerning gesture types, as indicated in Table 5, beats are present in all the examples. In addition to their usual cohesive function, beats in our sample frequently adopt a pragmatic function projecting the speakers' stance and contributing to the interpersonal function. Deictic gestures are only present in Example 1, but they seem to be a powerful tool to engage the audience. The same can be said about object manipulations, which

only occur in Example 3, but seem to play a crucial role in maintaining the level of attention of the audience. It is also remarkable that OHS gestures, which according to Kendon (2004) convey willingness to offer something, are considerably frequent, and the OHP gestures are mostly used with a pragmatic and interpersonal function, anticipating and preventing counterclaims from the audience and engaging them in a fictional dialogue.

The results discussed in the previous paragraphs highlight the important role that paralinguistic and kinesic features play in creating interpersonal connections, and suggests that they should be given a more prominent role in Academic English training.

Finally, it is worth noticing that the use of a skilfully orchestrated multimodal ensemble ensures a relatively constant focus on the interpersonal relationship with the audience, with one mode compensating for others when needed. This is illustrated in Example 3, where the static position which conveys distance is compensated by constant eye contact and direct addresses to the audience. Also the relatively less frequent eye contact found in Example 2 seems to be compensated by constant gesturing. This interplay of modes, together with the swift orchestration of distinct multimodal ensembles, proves that dissemination talks are multimodally dense (Norris 2004), and suggests that a skilful orchestration of different modes is key to achieving a desirable interpersonal relationship with the audience.

4 Conclusions

In this article we have presented a multimodal analysis of the interpersonal behaviour of speakers in a selection of dissemination talks. This genre is particularly interesting in terms of interpersonal meaning, because it differs from other academic presentations in the degree of specialization and shared knowledge assumed from the audience. This greater gap in terms of knowledge and interest between speaker and audience makes these presentations especially challenging for audience engagement. Speakers in specialized conference presentations can focus on the content of their presentations, which we might call the *what*. On the other hand, in order to keep their audience interested and engaged, speakers in dissemination talks need to focus on the relevance of their content, that is the *so what*. As we have shown in this article, this translates into a variety of strategies for audience engagement, such as establishing a close one-to-one relationship with them, anticipating their reactions or resorting to humour and humbleness.

The study presented in this article combines multimodal analysis with ethnographic methods such as observations and interviews. Although in line with other studies that approach interpersonal meaning from a multimodal perspective and draw on SFL (e.g. Forey and Feng 2016), our main objective is not to develop a descriptive grammar that can apply to different semiotic resources (Ledin and Machin 2019). Instead, we try to shed light on certain uses of multimodal ensembles within a specific context and discursive situation (i.e. dissemination talks) to ultimately improve our knowledge of this genre. This combined methodology provides a powerful tool to focus on multimodal ensembles and validate results with speakers.

Our results show that the speakers in the talks we have analysed shorten the distance with the audience by orchestrating multimodal ensembles that include language, paralanguage, kinesics, proxemics and gaze, among other modes. They also seem to adapt to the type of audience and select the most appropriate techniques to highlight the *so what* of their research and make it appealing to a layman audience. Interestingly, the moments that are richer in interpersonal meaning also feature a seamless orchestration of modes that results in something bigger than the sum of its parts. This allows speakers to maintain the degree of interpersonal connection with the audience at a constant high level and use modes to eventually compensate for each other when needed.

Our study presents limitations in terms of size and scope. The use of larger corpora would definitely allow for the generalizability of results and contribute to more comprehensive studies. As recent developments in software allow for an increasing automatization of multimodal analysis, quantitative MDA studies are becoming a promising niche for further research. Such advances will also allow for the study of more complex multimodal ensembles. Until then, exploratory, qualitative-based studies like the one we present here can hint

at useful venues for further quantitative research. Likewise, despite the limited generalisability, results already present pedagogical applications.

In fact, the qualitative analysis discussed in this article provides insights into what type of strategies speakers use to engage with their audiences in dissemination talks, which in turn present some interesting pedagogical implications. We agree with Forey and Feng (2016) that the multimodal nature of academic presentations is not yet fully reflected in the ESP and EAP curriculum, although it can contribute positively to develop students' communicative competence. Furthermore, our results also bring to the fore the importance of interpersonal meaning in these presentations. In view of our results and those of others (Morell 2015, 2018; Valeiras-Jurado et al. 2018), we suggest that a skilful orchestration of modes can be greatly beneficial to achieve the desired level of audience engagement. The role of multimodal competence in specialized discourse and across different spoken genres has been highlighted in recent studies, such as Morell and Pastor-Cesteros (2018) for academic presentations, Beltrán-Planques and Querol-Julián (2018) for business complaints, Querol-Julián and Fortanet-Gómez (2019) for discussion sessions or Jimenez-Muñoz (2019) for business pitches. Therefore, developing speakers' multimodal communicative competence should be a priority in L2 specific and academic training. Our findings, together with those of the cited studies above, on how speakers use and combine modes to engage their audience, allow us to recommend that EAP and ESP learners be encouraged to observe skilled speakers and to take note of, at least, their use of pronouns, intensifying adjectives, humour, spatial positions, gaze and gestures so as to develop their multimodal competence. This is particularly relevant in the case of dissemination talks for two main reasons. First, because non-verbal modes can help students to compensate for lack of expertise when using a foreign language. Second, because speakers in dissemination talks are challenged with the need to connect with an audience that does not necessarily share their specialized knowledge. They need to show the *so what* of their research. We believe that a mastery of different semiotic modes can provide them with the right tools to achieve this goal.

Appendix A: Multimodal ensembles in Example 1

Proposition	Tone unit	Spatial position	Gaze	Gesture
Hi (0.00–0.01)	Hi	1–2	Gaze towards centre of audience	
I'm a chemist. (0.01–0.02)	I'm a CHEmist	1–2	Gaze towards centre of audience	Deictic, pragmatic gesture, pointing with index finger of left hand.
And I really like kissing. (0.02–0.04)	and i REally like Klissing	2	Shifts towards left side of audience	Three pragmatic, closed-fist beats with left hand.
I'd like to kiss you. (0.04–0.06)	i'd LIKE to KISS you	2–3	Side view towards right side of audience	Deictic, referential-interpersonal gesture, pointing with index finger of left hand to person in centre.
I'd love to kiss you. (0.06–0.08)	i'd LOVE to KISS you	2–3	Continues side view further towards right side	Deictic, referential-interpersonal gesture, pointing with index finger of left hand to person on right side.
I'm gonna kiss all of you (0.08–0.10)	i'm GONna kiss ALL of you	3	Shifts from right to left side of audience	Open hand prone (OHP), metaphoric, pragmatic, gesture of left hand moving from right to left and simultaneous beating
because (0.10–0.11)	beCAUSE (11)	4 Shifts body orientation towards left	Straight (not to audience)	Closed-fist pragmatic beat with left hand
I really want to (0.11–0.12)	i REally want to (1)	5–6	Towards centre of audience	Closed-fist pragmatic beats with both hands

(continued)

Proposition	Tone unit	Spatial position	Gaze	Gesture
Keep it simple, (0.12–0.13)	KEEP it SIMple (0.5)	5–6	Towards centre of audience	Closed-fist pragmatic beats with both hands
But not stupid. (0.13–0.15)	but NOT stupid	5–6	Shifts towards floor	Closed-fist pragmatic beat with left hand and extended index finger
Tonight (0.15–16)	toNIGHT	6	Towards centre of audience	Closed-fist pragmatic beat with left hand
I'm going to show you how your own experience (0.16–0.18)	i'm going to SHOW you how your OWN experience	6	Shifts from left to right	-Closed-fist pragmatic beat with left hand -Deictic, referential, interpersonal gesture, pointing at audience with both hands and open palms.
with a very simple technique (0.18–0.20)	with a VEry simple technique	6	Shifts from right to left	OHP, metaphoric, pragmatic, gesture of both hands moving apart (sweeping) and simultaneous beating
can revolutionise the chemical industry (0.20–0.22)	can revolUTIONise the CHEmical industry	6	Shifts from left to right	Closed-fist pragmatic beats with both hands
And finally your life. (0.22–0.24)	and FInally YOUR life	6	Shifts towards centre	-Closed-fist pragmatic beats with both hands -Deictic, referential, interpersonal gesture, pointing at audience with both hands and open palms.

Appendix B: Multimodal ensembles in Example 2

Proposition	Tone unit	Spatial position	Gaze	Gesture
We use mathematics to build bridges. (0.00–0.02)	we USE mathematics to build BRIdges ↘	1–2	Shifts from right stage towards audience	– OHS metaphoric, referential-cohesive gesture with rotating hands – Referential-cohesive gesture (enumerating) with right hand holding left extended finger.
To build dinosaur traps, (0.02–0.04)	ah to BUILD ah DInosaur traps ↗	2–3	– Side view towards audience – Side view towards screen	Pragmatic Open hand supine (OHS) beats with both hands
you know, (0.04–0.05)	you KNOW ↗	4	Straight towards audience	Pragmatic OHS beats with both hands
and other things you use on a daily basis. (0.05–0.07)	and other things you use on a daily BASis ↗	4	Straight towards audience	Pragmatic OHS beats with both hands

Appendix C: Multimodal ensembles in Example 3

Proposition	Tone unit	Spatial position	Gaze	Gestures and head movements
Are you smart? (0.00–0.01)	ARE YOU SMART ↘	1	Gaze on audience	Three open hand supine (OHS) pragmatic and interpersonal beats
(3 s pause) (Audience- Yes) (0.01–0.04)	–	1	Shifts left, right left, right	Lateral head movements
I knew that this would come (0.04–0.05)	i KNEW that this would come	1	Shifts right, left	OHS pragmatic beat: Opens and closes arms twice + head nods
but let me rephrase (0.05–0.07)	but LET me rePHRASE	1	Shifts right, left, right	Open hand prone (OHP) pragmatic beat: Holds hands together, then opens arms
Is your phone smart? (0.07–0.09)	is your PHONE smart ↘	1	Shifts left, right, left	Takes phone out of left pocket with left hand
(Audience – Yes) (0.09–0.10)		1	Shifts right, left	Holds phone up
You would say of course (0.10–0.12)	You would say of COURSE	1	Shifts right, left	OHS pragmatic beats
It is a smart phone (0.12–0.13)	it is a smart PHONE phone	1	Shifts right, left	OHS pragmatic beats
So, why do you ask (0.13–0.15)	so, WHY do you ASK	1 Shifts slightly towards right	Shifts right, left	Puts phone back in pocket with left hand
Well, then I say, (0.15–0.17)	well then I SAY	1	Shifts right, left	Puts phone back in pocket with left hand
Is the clothes you wear smart? (0.17–19)	is the CLOTHES you WEAR smart ↘	1	Shifts left, right	Deictic referential gesture with both hands OHP moving downwards along the body
The jeans, (0.19–0.20)	the JEANS	1	Straight	OHS deictic referential gesture pointing at jeans
the jacket (0.20–0.22)	the JACKET	1	Straight	OHS deictic referential gesture: Lifting and crossing arms over the jacket
The clothes that you have (0.22–0.23)	the CLOTHES that you have	1	Right, left	
And it would be very silent (0.23–0.24)	and it would be VERY SILENT ↘	1	Right, left	OHP metaphoric, pragmatic gesture, with hands moving apart (sweeping) + head nods

Appendix D: Multimodal ensembles in Example 4

Proposition	Tone unit	Spatial position	Gaze	Gesture and head movements
I must say all of a sudden (0.00–0.02)	i must SAY ALL of a sudden	1–2	Side view towards right	Open hand prone (OHP) pragmatic beats
I feel really really humble. (0.02–0.04)	I feel REALLy REALLy HUMble ↘	2	On audience centre	Finger ring pragmatic beats
I've just (0.04–0.06)	I've just	2	Back towards right	
These these five presentations that came before me (0.06–0.08)	These FIVE presentations that came beFORE me	2–3	Further right	OHP pragmatic beats
They all want to change the world (0.08–0.10)	they all want to CHANGE the WORLD	3	Audience centre	OHP pragmatic beats
And they are doing it in an incredible way (0.10–0.12)	and they are DOING it in an inCREDible way ↘	3	Left	OHP pragmatic beats with right arm opening towards right + subtle head shakes
They are doing superconductors in microwaves (0.12–0.13)	They are doing SUperconductors in MICrowaves	3	Centre, then right	Open hand supine (OHS) pragmatic beats: Towards left, then opening arms wider + subtle head shakes
There are things I've never heard of textiles (0.13–0.14)	there are THINGS I've never HEARD of textiles	3	Left, down	OHS pragmatic beats: opening arms and hands wider and then moving arms towards right
It's it's, (0.14–0.15)	It's it's	3–4	Down	OHS pragmatic beats
I'm just trying to get you from a to b (0.15–0.16)	I'm JUST trying to GET you from a to b ↘	4	Centre	OHS pragmatic beats from left to right + subtle head shakes
That's all I want to do (0.16–17)	That's ALL I want to do ↘	4	Centre	OHS pragmatic beats + metaphoric OHS gesture opening arms and hands wider + subtle head shakes
It's it's (0.17–0.18)	It's it's	4–5	Centre	
It's not rocket science (0.18–0.20)	It's NOT ROcket ↗ science	5	Centre	OHS pragmatic beats opening arms and hands wider + subtle head shakes
It's not superconductors (0.20–0.22)	It's not SUperconductors ↗	5	Left	OHS pragmatic beats towards left

(continued)

Proposition	Tone unit	Spatial position	Gaze	Gesture and head movements
It's just trying to inform you (0.22–0.25)	It's JUST trying to inform you	5	Right	Finger ring pragmatic beats
And this is really difficult (0.25–30)	and THIS is really Difficult	5	Centre	Finger ring pragmatic beats

References

- Alcaraz Varó, E. (2000). *El inglés profesional y académico* [Professional and Academic English]. Alianza, Madrid.
- Atkinson, M. (1984). *Our master's voices: the language and body language of politicians*. Methuen & Co, NY.
- Bavelas, J., Chovil, N., Coates, L., and Roe, L. (1995). Gestures specialized for dialogue. *Pers. Soc. Psychol. Bull.* 21: 394–405.
- Beltrán-Planques, V. and Querol-Julián, M. (2018). 'English language learners' spoken interaction: what a multimodal perspective reveals about pragmatic competence. *System* 77: 80–90.
- Brazil, D. (1997). *The communicative value of intonation in English*, 2nd ed. Cambridge: Cambridge University Press.
- Carter, R. (1997). *Working with texts: a core book for language analysis*. Routledge, London.
- Forey, G. and Feng, D. (2016). Interpersonal meaning and audience engagement in academic presentations: a multimodal discourse analysis perspective. In: Hyland, K. and Shaw, P. (Eds.). *The Routledge handbook of English for academic purposes*. Routledge, London, pp. 416–430.
- Fortanet-Gómez, I. and Ruiz-Madrid, N. (2016). Multimodal Humour in plenary lectures in English and in Spanish. *Multimodal Commun.* 5: 246–251.
- Goldman, R., Erickson, F., Lemke, J., and Derry, S. (2007). Selection in video. In: Derry, S. (Ed.), *Guidelines for video research in education: recommendations from an expert panel*. Data Research and Development Center (NORC at the University of Chicago), Chicago, pp. 19–27.
- Hadar, U., Steiner, T., Grant, E.C., and Clifford Rose, F. (1983). Kinematics of head movements accompanying speech during conversation. *Hum. Mov. Sci.* 2: 35–46.
- Halliday, M. and Greaves, W. (2008). *Intonation in the grammar of English*. Equinox, London.
- Hargie, O. (2011). *Skilled interpersonal communication: research, theory and practice*, 5th ed. London: Routledge.
- Hood, S. and Forey, G. (2005). Introducing a conference paper: getting interpersonal with your audience. *J. Engl. Acad. Purp.* 4: 291–306.
- Hyland, K. (2009). *Academic discourse*. Continuum, London.
- Jewitt, C. (2012). *An introduction to using video for research*. Working paper. NCRM Working Paper Series.
- Jiménez-Muñoz, A. (2019). Beyond language: a multimodal analysis of business in non-native business English pitches. *Iberica* 37: 65–86.
- Kendon, A. (2004). *Gesture: visible action as utterance*. Cambridge University Press, Cambridge.
- Kress, G. and Van Leeuwen, T. (1996). *Reading images—a grammar of visual design*, 1st ed. London: Routledge.
- Lakoff, R. (1982). Persuasive discourse and ordinary conversation, with examples from advertising. In: Tannen, D. (Ed.). *Analysing discourse: text and talk*. Georgetown University Press, Washington, pp. 25–42.
- Ledin, P., and Machin, D. (2019). Doing critical discourse studies with multimodality: from metafunctions to materiality. *Crit. Discourse Stud.* 16: 497–513.
- Lim, F. V., O'Halloran, K., and Podlasov, A. (2012). Spatial pedagogy: mapping meanings in the use of classroom space. *Camb. J. Educ.* 42: 235–251.
- McNeill, D. (1992). *Hand and mind: what gestures reveal about thought*. University of Chicago Press, Chicago.
- Morell, T. (2004). Interactive lecture discourse for university EFL students. *Engl. Specif. Purp.* 23: 325–338.
- Morell, T. (2015). International conference paper presentations: a multimodal analysis to determine effectiveness. *Engl. Specif. Purp.* 37: 137–150.
- Morell, T. (2018). Multimodal competence and effective interactive lecturing. *System* 77: 70–79.
- Morell, T., and Pastor-Cesteros, S. (2018). Multimodal communication in academic presentations by L2 Spanish students. *J. Span. Lang. Teach* 5: 125–138.

- Norris, S. (2004). *Analysing multimodal interaction: a methodological framework*. Continuum, New York, London.
- Norris, S. (2011). *Identity in (inter)action. introducing multimodal (inter)action analysis*. De Gruyter Mouton, Boston.
- Norris, S. (2019). *Systematically working with multimodal data: research methods in multimodal discourse analysis*. Wiley & Sons, Hoboken, NJ.
- Norris, S. (2020). *Multimodal theory and methodology: for the analysis of interaction and identity*. Routledge, NY.
- Norte Fernández-Pacheco, N. (2018). The impact of multimodal ensembles on audio-visual comprehension: implementing vodcasts in EFL contexts. *Multimodal Commun.* Available at: <https://doi.org/10.1515/mc-2018-0002> (Accessed 15 March 2019).
- O'Halloran, K. (2011). Multimodal discourse analysis. In: Hyland, K. and Paltridge, B. (Eds.). *The continuum companion to discourse analysis*. Bloomsbury, New York, pp. 120–137.
- Poyatos, F. (1983). Language and nonverbal sign systems in the structure of face to face interaction. *Lang. Commun.* 3: 129–40.
- Querol-Julián, M. (2011). *Evaluation in discussion sessions of conference paper presentations: a multimodal approach*. Lambert Academic Publishing, Saarbrücken, Germany.
- Querol-Julián, M., and Fortanet-Gómez, I. (2019). Collaborative teaching and learning of interactive multimodal spoken academic genres for doctoral students. *Int. J. Engl. Stud.* 19: 61–82, doi:<https://doi.org/10.6018/ijes.348911>.
- Ruiz-Madrid, N. and Fortanet-Gómez, I. (2015). A multimodal discourse analysis approach to humour in conference presentations: the case of autobiographic references. *Procedia – Soc. Behav. Sci.* 173: 246–251.
- Valeiras-Jurado, J., Ruiz-Madrid, N., and 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.