

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

English for Specific Purposes

journal homepage: <http://ees.elsevier.com/esp/default.asp>

From the Editors

The genre of PechaKucha presentations: Analysis and implications for enhancing multimodal literacy at university

Vicent Beltrán-Palanques^{a,*}, Mercedes Querol-Julián^b^a Universitat Jaume I, Campus Riu Sec, 12071, Castellón, Spain^b Universidad Internacional de La Rioja (UNIR), Av. de la Paz, 137, 26006, Logroño, La Rioja, Spain

ARTICLE INFO

Article history:

Available online 27 May 2024

Keywords:

PechaKucha presentations
 Multimodal discourse analysis
 Multimodal literacy
 ESP pedagogy

ABSTRACT

ESP students need to develop their multimodal literacy to become literate in today's professional spaces. For this purpose, ESP teachers should revisit pedagogy practices to best engage students in the navigation and construction of multimodal genres. As a case in point, we explore PechaKucha (PK) presentations. This multimodal genre consists of 20 slides, which are automatically advanced every 20 s. PK presentations entail a complex format that requires speakers to choose how to convey content, design suitable visuals, and engage audiences. The dataset for the study consists of 7 PK presentations delivered during a social event at an architecture conference. Adopting a multimodal discourse analysis lens, we analyse this set of PK presentations in terms of rhetorical structure and the way in which intersemiotic relations unfolded (synchronisation between speech and visuals and the modal density of slides). The analysis demonstrates that PK presentations entail an intricate multimodal composition consisting of three moves in which professional and personal narratives intertwine. The examination of intersemiotic relations reveals how speech and visuals interplay effectively to transmit meaning and engage the audience. The results of this study provide critical information to design a research-informed pedagogy to enhance ESP students' multimodal literacy.

© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Professional and academic communication is constantly changing, likely largely due to advances in technology and audiovisual resources. The way of conceiving communication has also evolved, becoming increasingly digital, immediate and, in some cases, even brief. This situation has reshaped day-to-day professional and academic discursive practices, which in turn gives rise to the remediation and emergence of genres (Luzón and Llantada, 2019). Such changes in professional and academic genres are not trivial and have a major impact on communication. This is particularly observable in audiovisual and digital genres, in which multiple semiotic resources are combined to create a single communicative event (Luzón and Pérez-Llantada, 2022). In this study, we focus on the genre of PechaKucha (PK) presentations to explore their rhetorical organisation and the way varied modes of communication unfold to construct intersemiotic relations to promote audience' comprehension and engagement.

* Corresponding author.

E-mail addresses: vbeltran@uji.es (V. Beltrán-Palanques), mercedes.querol@unir.net (M. Querol-Julián).

While genres are shown to be made up of diverse communicative modes (e.g., [Feng, 2021](#); [Luzón, 2023](#)), early approaches to genre analysis ([Halliday, 1978](#); [Swales, 1990](#)) were usually oriented towards the exploration of a single communicative mode (written/spoken language), which reduces the scope of analysis. More recent approaches to genre analysis tend to also consider the multimodal representation of meaning, which allows a much broader and more holistic exploration. The analysis of genres from the perspective of multimodality helps identify the variety of semiotic modes, along with their affordances or potentialities ([Kress, 2010](#)), that contribute to the meaning-making process. The multimodal approach to genre analysis has been endorsed by a number of research studies, resulting in a growing body of literature that follows this view. For example, [Hiippala \(2014\)](#) argues for the use of methodological approaches that expand the analysis of a single mode of communication (i.e., written or spoken language) to the analysis of varying semiotic resources. Along the same line, [Tardy and Swales \(2014\)](#) claim the idea of adopting a multimodal/visual genre analysis approach that permits the identification of non-textual elements such as images. In line with this, [Bateman \(2008\)](#) proposes the Genre and Multimodality (GeM) model that allows the exploration of visual style (e.g., layout) in terms of an expanded notion of multimodal genre ([Jewitt et al., 2016](#)). Drawing on systemic functional linguistics ([Halliday, 1978](#)), [van Leeuwen \(2005\)](#) explores how genre analysis should be modified to also incorporate multimodal perspectives and discusses the following three main assumptions: 1) a stage may be performed by several modes, 2) a stage can be multimodal itself, and 3) the relation between modes within a stage can be elaborative (content performed in one mode is restated through another mode, e.g., adding an explanation) or extensive (i.e., “one mode adds new, related content to the content expressed in another mode” (p. 77). [O’Halloran \(2021\)](#) describes the paradigm of systemic functional multimodal discourse analysis (SF-MDA) in which the study of discourse encompasses language in combination with other semiotic modes. Central in SF-MDA are intersemiosis and resemiotisation. Intersemiosis refers to the process by which semiotic choices interact and combine to make meaning. Resemiotisation involves the reconstruction of semiotic choices within and through multimodal phenomena ([Iedeman, 2001](#)). Particularly relevant for the present study is the notion of intersemiotic relations, especially in terms of visual and speech (e.g., [Bateman, 2014](#); [Lim & O’Halloran, 2012](#); [O’Halloran and Lim, 2009](#)). In this regard, [Royce \(1998\)](#) contends that visuals and verbal modes complement each other to create a single unit. For [O’Halloran \(2005\)](#), intersemiotic relations describe the meaning that emerges from semiotic choices. [Lim \(2019\)](#) adds that these relations refer to the ways in which modes of communication are combined to construct meaning.

Due to the multimodal nature of genres, speakers/writers seem to face new communicative challenges that require the knowledge and mastery of new forms of literacy, different from traditional ones. In the context of education, the conventional notion of literacy is understood primarily in terms of reading and writing. However, this vision of literacy may no longer be sufficient to meet the communicative demands of the 21st century. Concerning this, the [New London Group \(1996\)](#) questioned the conservative view of literacy and promoted multiliteracies. In general terms, this movement was intended to serve as a driving force to transform and reframe pedagogy, as well as to respond to the remarkable increase of technology. In so doing, this group expanded the traditional concept of literacy by emphasising the multimodal nature of communication and calling for a pedagogy based on multiliteracies. This viewpoint has implications for language teaching and learning (e.g., [Hafner and Ho, 2020](#); [Lim, 2018](#)). In line with this, [Lim and Tan-Chia \(2022\)](#) also argue for the need to expand on the well-known four language skills to incorporate “viewing” to reading and listening and “representing” to writing and speaking. Viewing involves the interpretation and understanding of multimodal texts and representing entails the representation of ideas and points of view through multiple texts ([Lim et al., 2020](#)). These research developments seem to point to a paradigm shift in which the approach to teaching language is only challenged to encompass other modes of communication ([Jewitt et al., 2016](#)). This paradigm shift involves a much broader sense of understanding literacy that is becoming increasingly relevant in the field of language teaching and learning. In essence, while traditional literacy continues to be addressed, multimodal text production is on the rise ([Maagerø and Tønnessen, 2022](#)).

The term multimodal literacy came to the fore to explicate the way meaning is made through diverse communicative modes ([Jewitt and Kress, 2003](#)). The term refers to the ability to use and combine semiotic modes (resources used to understand, navigate and construct discourse) modes in a particular way to construct meaning ([van Leeuwen, 2017](#)) and encompasses the knowledge and the skills students need to navigate and produce multimodal texts ([Eisenlauer and Karatza, 2020](#)). In addition to this, a cornerstone aspect of multimodal literacy is the building of students’ semiotic awareness ([Towndrow et al., 2013](#)), understood as “critical attention to relational, multimodal aspects of meaning design, without which meaningful assessment schemes can neither be conceived nor implemented” (p. 328). The development of such features will enable students to be multimodal literate, which in turn will help them successfully engage in the understanding, construction and representation of multimodal texts ([O’Halloran and Lim, 2011](#)). In response to the call for a change in language teaching, several researchers have started to incorporate multimodality in the language classroom (e.g., [Diamantopoulou and Ørevik, 2022](#); [Jiang, 2017](#); [Jiang et al., 2019](#); [Kim and Belcher, 2020](#)). Research has shown the need to explore new literacy practices such as the involvement of students in the navigation and production of multimodal texts. Enhancing students’ multimodal literacy is essential at all educational levels; at each level, however, there will be specific requirements and learning objectives. So far, researchers have focused their attention mainly on primary and secondary education ([Lim et al., 2020](#)), and on tertiary education to a lesser extent ([Querol-Julián and Fortanet-Gómez, 2025](#)). Although multimodal literacy in language teaching is still at a nascent stage at university, growing research interest in this field is emerging, particularly in English for Specific Purposes (ESP) (e.g., [Beltrán-Palanques, 2024](#); [Crawford Camiciottoli, 2019](#); [Hafner and Miller, 2018](#); [Incelli, 2021](#)).

Integrating new forms of literacy into the curriculum involves revising ESP pedagogy with the aim of designing learning experiences aligned with the professional communicative demands of students and with the goal of supporting them in

becoming multimodal literate. For this purpose, ESP teachers should create rich learning situations centred on the development of students' multimodal literacy (e.g., engage students in the construction and deconstruction of professional genres). Through the development of students' multimodal literacy, teachers can go beyond the building of students' linguistic repertoire to also promote their awareness of the role diverse semiotic resources play in the meaning-making process. In addition, through the development of multimodal literacy, ESP teachers can also foster critical thinking skills (Querol-Julián and Fortanet-Gómez, 2019), visual skills, or analytical skills (Feng, 2021). In order for ESP teachers to integrate multimodal literacy in ESP, they have to first fully understand genres. In this sense, we maintain that undertaking a multimodal analysis of genres is critical to best inform the design of learning experiences aimed to develop multimodal literacy.

As a case in point, in this study, we address the genre of PK presentations. This genre falls under the category of blitz presentations, also known as rapid-fire presentations (Harinck and van Leeuwen, 2020). These presentations are characterised by their brevity. PK is a presentation concept that appeared in 2003 with the work of Astrid Klein and Mark Dytham, founders of Klein Dytham Architecture in Japan. The original idea for this presentation format was born within the framework of an event in which designers and architects engagingly presented their work. The purpose was to reduce the time each presenter devoted to each slide. Therefore, it was decided that each slide would be 20 s long. A PK presentation consists of 20 slides, each one lasting 20 s (20 × 20) that are advanced automatically, with a total duration of 6 min and 40 s. This type of presentation is becoming increasingly popular to disseminate knowledge to a broad mix of audiences, including experts and non-experts. They differ from other types of presentations such as conference presentations in that they tend to be fast-paced and less formal. The design of PK presentations is distinguished by its simplicity as they only contain static slides with no moving images and additional sound (Frost, 2020). Moreover, the slides rely mainly on visual support (e.g., images and pictures) while written information is reduced to the minimum. The format of PK presentations forces presenters to convey meaning without digression (Courtney-Klentzin et al., 2010) within a limited amount of time. This is quite challenging since each slide is pre-programmed, and hence synchronising speech and slides can be critical. In addition, presenters should make decisions regarding the visual representation of the topic and reflect on how to use diverse semiotic resources to effectively disseminate knowledge. Due to the diversity of audiences, presenters should also adapt their speech in terms of content complexity. This implies, for example, considerations regarding lexical choices. By the same token, it is key for presenters to create interpersonal meaning to engage and establish rapport with the audience throughout the presentation. In this sense, research has shown the relevance of establishing interpersonal meaning across different kinds of presentations such as conference presentations (Morell, 2015) and Three-Minute Thesis presentations (Hyland and Zou, 2022). Engagement can be achieved through the expression of, for example, specific linguistic strategies such as hearer mentions, appeals to shared knowledge, questions (Hyland, 2005), which may lead to establish rapport between the presenters and the audience. Thus, creating a PK presentation can be difficult as it requires presenters to consider several factors such as slide design, language choices, intersemiotic relations, or the enactment of interpersonal meaning.

As in other dissemination genres such as Three-Minute Thesis (3 MT) presentations (e.g., Carter-Thomas and Rowley-Jolivet, 2020), performing a PK presentation involves not only the expression of content but also arousing the interest and attention of the audience through multiple semiotic resources. Against this backdrop, the present study attempts to carry out a multimodal genre analysis of PK presentations to determine the rhetorical structure and the way intersemiotic relations unfold to promote audience comprehension and establish interpersonal meaning. The research questions guiding the present study are as follows:

RQ1 What rhetorical moves are used to construct PK presentations?

RQ2 How is the intersemiotic relation unfolded in PK presentations to facilitate comprehension and promote engagement?

2. Methodology

The dataset for this study comprises 7 PK presentations delivered at a night event, ACA 18 TOKYO Design Talk, that was part of the social programme of a professional conference in architecture (ACA18 TOKYO - 18th Asian Congress of Architects) held in Tokyo in 2018. The conference is organised annually by *Arcasia* (the Architects Regional Council Asia) and the social event, which is known as a “PechaKucha night”, by PechaKucha. PK, apart from a presentation format, is also “a free online and offline global storytelling platform that celebrates people, passion, and creative thought” (PechaKucha, 2023). The recordings of the 7 PK presentations were uploaded onto the PechaKucha website.¹ These were chosen for the study as they were part of the same event in which each participant talked about their favourite city from a professional viewpoint. As shown in Table 1, the multimodal data entailed a 48-min video recording of the presentations, which include the presenters' voices and the visuals. That is a total of 140 slides and around 8,000 words.

The dataset was closely examined by adopting a corpus-driven approach. That is, the dataset served as an empirical basis from which we extracted information without prior assumptions or expectations (Tognini-Bonelli, 2001). To answer RQ1, we identified the organisation of rhetorical moves (Swales, 1990) within each PK presentation. The identification of rhetorical

¹ <https://www.pechakucha.com/events/aca-18-tokyo-design-talk-powered-by-pechakucha>.

Table 1
Dataset of PechaKucha presentations.

Presenter	Pecha Kucha's title	Duration (min. sec.)	# words
1	My Favorite City: Archaeology of the Future City	6.49	1,229
2	My Favorite City: Singapore through the lens of Architecture	6.46	1,295
3	My Favorite City: Vertical city - Hong Kong	6.50	1,307
4	My Favorite City: UNStudio's Amsterdam	7.08	892
5	My Favorite City: Islamabad the beautiful	6.44	1,340
6	My Favorite City: Undulations of the Scared Cities	6.49	634
7	My Favorite City: 20 × 20 of Indonesian young architects	6.54	916
Total		48.00	7,613
Mean		6.57	915

moves served the purpose of establishing the macrostructure of the genre. This approach has extensively been used in the exploration of other dissemination talks as we illustrate in the presentation of the results.

To respond to RQ2, we explored how the intersemiotic relation between speech and visuals unfolded and how this relation might be intended to facilitate audience's comprehension and engagement. First, we focused on the synchronisation of speech and visuals. For this purpose, we identified the topic units of the speech, their duration and if they were temporally and conceptually aligned with the content presented on the slides. Topic units were identified based on cues such as discourse markers, silence, and mainly changes in the topic of the presentation. Then, we analysed the modal density of the slides and the intersemiotic relations developed between the speech and the different degrees of modal density of the slides. To determine the modal density of the slides, we considered the number of communicative modes that composed each slide and quantified their modal intensity. Modal intensity refers to the "weight or importance of specific modes" (Norris, 2004, p. 80). Thus, in this study, we define the modal intensity of each communicative mode as the area it occupies on the slide. GeoGebra² software assisted in the quantitative analysis of modal density (see Beltrán-Palanques and Querol-Julián, 2018; Querol-Julián, 2023), as it automatically calculates areas defined by the user. Thus, we determined on the slide the area of each communicative move including written content, images, pictures, etc. According to the findings we classified the slides as having low, medium and high modal density. We used ELAN³ to examine the rhetorical moves and topic units and conduct the multimodal analysis. This software allows multilayer annotation and facilitates the systematic analysis of video and audio data. We coded the rhetorical moves and identified the topic units, when each slide was on the screen, the content of the slides, and the (de)synchronisation of speech and visuals. ELAN helps with the qualitative and quantitative examination of the annotations since we can retrieve the frequency and duration, as well as the beginning time and the end time. We annotated the dataset separately (percentage of inter-rater agreement of 90%) and discrepancies were discussed.

3. Results

The PK presentations analysed showed some structural features similar to other dissemination talks, academic (e.g., conference paper (CP) presentations, research pitches (RP) or 3 MT presentations) and non-academic (e.g., TED talks). Nevertheless, PK presentations differ notably from them in length and focus. First, PK presentations are not as long as CP presentations and TED talks, which commonly last 15–20 min, and as short as RPs and 3 MT presentations, which last no more than 3 min. Second, as mentioned, these PK presentations were part of the social programme of a professional conference, and all 7 shared the same topic "My favourite city". Thus, speakers referred to professional information, but personal and professional spheres intertwined throughout the presentations. Hence, the presenters talked about their professional projects but also shared personal information and experiences. In the case of CP and 3 MT presentations, which focus on academic issues, personal narratives are occasionally used to open CP presentations (Viera and Williams, 2020) and are rarely employed as an engagement strategy in 3 MT presentations (Hyland and Zou, 2022). TED talks, on the other hand, are described as personal storytelling characterised by emotional and authentic personal language (MacKrill et al., 2021). These two features, length and focus, seem to have influenced the macrostructure and intersemiosis of the genre, as we present in the next sections.

3.1. Rhetorical moves

The PK presentations were structured in three moves: set-up, topic development and closing. Topic development was the longest and most elaborated move, while closing was the shortest (Table 2).

The set-up move consisted of 3 other optional moves: listener orientation (greetings/thanking the audience), speaker presentation/self-introduction, and content orientation/topic introduction. These opening moves are also found in CP

² <https://www.geogebra.org/>.

³ EUDICO Linguistic Annotator (<http://www.lat-mpi.eu/tools/elan/>), developed at the Max Planck Institute for Psycholinguistics (MIP) (Nijmegen, The Netherlands).

Table 2

Pecha Kucha presentation's moves.

	Set-up		Topic development		Closing	
	Duration (min. sec.)	# words	Duration (min. sec.)	# words	Duration (min. sec.)	# words
Mean	0.25	72.7	6.22	1,008.4	0.02	6.4
Maximum	0.43	142	6.40	1,263	0.06	19
Minimum	0.02	7	6.01	597	0.01	4

presentations (Hood and Forey, 2005), 3 MT presentations (Hu and Liu, 2018), RPs (Ruiz-Madrid, 2021), and TED talks (Chang and Huang, 2015). As shown in Figure 1,⁴ results revealed that some presenters scarcely devoted time to this stage, such as Presenter 1 who performed the shortest set-up move of the dataset (see transcription conventions in Appendix).

He just introduced himself by saying his name while a picture of New York was on the screen. In this case, there was no intersemiotic relation between speech and visuals as the speech was completely unrelated to the content shown on the slide. Furthermore, the slide could have created false expectations of a topic introduction move as New York was not this presenter's favourite city. However, with this picture, the presenter started the topic development move. Indeed, only in 3 PK presentations the first slide showed information directly related to the set-up move. For example, Presenter 5's speech and the content on the slide were partially related (Figure 2). Unlike Presenter 1, he performed the three opening moves: the listening orientation, through the speech; the speaker presentation, through the visuals; and the topic introduction, through the speech and visuals.

First, the speaker made the listener orientation move by saying, "konnichiwa and as-salaam-alaikum from Pakistan to all of you it is lovely to be here". He mentioned the audience and evaluated positively taking part in the event. What we found interesting is the two cultural references. He greeted in Japanese and Arab, possibly to create a good rapport with the audience from the outset when referring to the host culture, and to project his own personal identity when alluding to his origins. As part of this first move of the set-up, the slide also included the name of the conference and of the social event ("ACA-18 TOKYO DESING TALK" – ACA stands for Asian Congress of Architects). Then, to make the speaker's presentation move he included the following information on the slide: his name preceded by the title "AR" and his membership ("IAP, Pakistan" – IAP stands for Institute of Architects of Pakistan). He did not self-present just as an architect but as part of a professional association of architects in his country, which may have helped the audience to build up his professional identity. Finally, the content orientation move was made first through the slide, which also contained the name of a city as part of a kind of motto ("Islamabad is beautiful") and a picture of the Faisal Mosque, the largest mosque in Pakistan and an iconic symbol of Islamabad throughout the world. The positive evaluation of the city, "beautiful", made clear that Islamabad was the speaker's favourite city. Indeed, he did not even mention its name when introducing the topic of the talk, but he presented it through a touching narrative of his personal and professional relationship with the city, "the city that I grew up (...) that I designed (...) places where I've always been part of growing up walking jogging and cycling around". Sharing these memories of the city as part of his childhood and life may be also an engagement strategy to evoke positive feelings in the audience.

**Figure 1.** The shortest set-up move.

⁴ The number of the slide and the text outside the picture in the figures have been added by the authors of the paper.

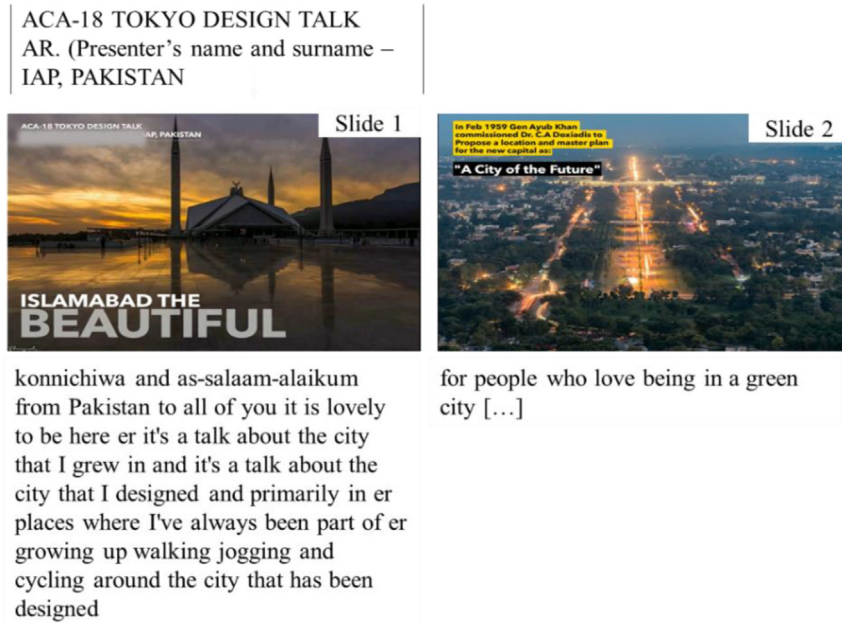


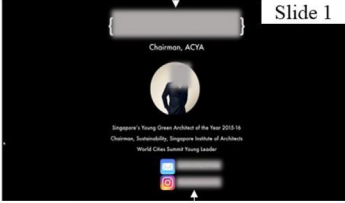
Figure 2. Set-up move with average length.

Besides, the presenter closed the move by stating the green preferences of its inhabitants, “the city that has been designed for people who love being in a green city”. He projected the image of a city not only aesthetically beautiful, as one could see in the picture, but his words were deeper inside conveying the message that the city is what it is because of its people. In line with current global environmental concerns, the idea of environmental awareness and sustainability will be a recurrent topic later in this presentation and for all the speakers. In this example, self-introduction and content orientation moves merged visually on one slide.

Presenter 2 also made the three moves to perform the longest set-up stage found in the dataset (Figure 3); however, he chose to design two slides to present himself and to introduce the topic of the presentation.


First, he thanked the audience, PechaKucha, and the organisation of the conference, “what a moment thank you guys it’s a big honour to be here thank you to (GIA) thank you PechaKucha and thank you Arcasia because that’s where all began in some ways for me”. The positive evaluation of the experience, the informal way to address the audience (“guys”) and the special mention of Arcasia, looking back on his early career as an architect, make us think that a good rapport with the audience and engagement were developed. Then, he made an aside about time, “I’ve realised this slide also runs for 20 s”. The speaker revealed how conscious he was of the format of the PK presentation and specifically of time constraints. He mentioned the structure of the self-introduction move, “I’m gonna talk about er er who I am what I do for the next 10 s”. His name and a picture, as well as professional and contact information, were displayed on slide 1 in white letters on a black background. That was the way the speaker projected his professional image visually. Nonetheless, despite his level of awareness of time, he had just started self-introduction when slide 2 showed up. The speech did not repeat visual information but highlighted the speaker’s professional profile. First, he commented on his past connection with Arcasia, “I happened to chair a committee fantastic committee in Arcasia called the Arcasia committee on young architects we (shall er too guys okay)”. The interpersonal meaning of this part of the speech was prominent. The presenter showed a positive attitude towards the committee (“fantastic committee”). Besides, he addressed those who were also part of it, uttering, “we shall er too guys okay”; afterward, he waited for 2 s for their response, an ovation, and welcomed it laughing. Then, he mentioned his activity as an architect and sustainability work, “I also do architecture and sustain sustainability work”. Finally, a discourse marker introduced the topic of the presentation, “back in my favourite city”. However, the audience already knew what the city was since slide 2 was on the screen: a drawing of a city surrounded by green areas and the statement “Singapore. Probably the closest to Utopia you’ll ever get”. In this statement, once more, the presenter showed his stance, first mitigating its intensity with a downtoner (“probably”) and then expressing his positive attitude. An engagement strategy of mentioning the audience was employed. Despite showing this information, the name of the city was also uttered, “that’s er Singapore”. A connection could be established between the mention of his sustainable work and the green image of Singapore. Finally, he interacted again with the audience through the question, “anyone has been there?”, waited for an answer and made a joke about the enthusiastic positive response of some of them who he guessed were from Singapore, “pretty much alright and the big shout out to anyone from Singapore I see as well come on”.

Presenter's name and surname
Chairman, ACYA




Slide 1

Singapore's Young Green Architect of the Year 2015-16
Chairman, Sustainability, Singapore Institute of Architects
World Cities Summit Young Leader
Gmail address
Instagram address



Slide 2

{ Singapore }
Probably the closest to Utopia you'll ever get.



Slide 3

Singapore
Probably the closest to Utopia you'll ever get.

what a moment thank you guys it's a big honour to be here thank you to (GIA) thank you PechaKucha and thank you Arcasia because that's where all began in some ways for me I've realised this slide also runs for twenty seconds so I'm gonna talk about er er who I am what I do for the next ten seconds er I happened to chair a committee fantastic committee

in Arcasia called the Arcasia committee on young architects we (shall er too guys okay) <pause: 2 sec> [Audience: ovation] <laughter> and I also do architecture and sustain sustainability work back in my favourite city and that's er Singapore anyone has been there? <pause: 3 sec.> [Audience: yes] pretty much alright and the big shout out to anyone from Singapore I see as well come on

five of you come on <laughter> [...]

Figure 3. The longest set-up move.

The topic development move was found to be a required move in the PK presentations analysed in the study as the presenters elaborated on the main theme of the talk. The structure of this move seems to differ from the structure of short academic presentations such as 3 MT presentations in which presenters used various moves to present the main topic of the talk (Hu and Liu, 2018). This is probably because the PK presentations examined were professionally oriented. Thus, the structure of this stage was more similar to that found in TED talks (Chang and Huang, 2015): present an argument, offer an explanation, and describe a process/series of events. Throughout the topic development move, the PK presenters generally addressed their vision of future cities and described projects. To accomplish this move, the presenters tended to combine both personal and professional information. This permitted them to show their professional position by referring to themselves as architects and their studies, to include some personal asides to comment on what has been uttered (Hyland, 2005), or to share experiences that belonged to the personal and professional spheres. In our dataset, the topic development move resembles a narration in which the presenters dealt with the main theme of the talk. Most topics addressed within this stage referred to the notion of the future city. As we can see in the following examples, the speakers approach the notion of the future city by relating it, for example, to sustainability (1), the idea of maintaining a healthy city (2) and tradition (3). The speakers also showcased their position as architects by expressing their vision of what a city would look like (1 and 3) or mentioning their architecture projects (1 and 4). As shown in the examples, this move is featured by the expression of interpersonal meaning too.

- (1) it was designed primarily around what was really there and made sure that was about green planning it has been designed on the basis of the ideal city of the future (Presenter 5)
- (2) we all of us here today have part of the future of our favourite cities in our hands and the question is what what can we do as design professionals to safeguard er a healthy future for all our cities (Presenter 4)
- (3) my idea of the ideal is the er archaeology of the future er city so how can we think not only new idea new design new things but actually more learning from the deeper longer er past to be able to come up or our let's say future idea for the city (Presenter 1)
- (4) near Jakarta I have a project er I can play some geometric umm er geometric er shape and I played this kind of interior and this is er this is er a typical type of Indonesian new Jakarta architecture project (Presenter 7)

In the third stage, the presenters signalled the closing of their talk by showing gratitude to the audience for having attended the presentation. The presenters tended to just thank the audience with "thank you very much", as in other blitz presentations like 3 MT presentations (Hu and Liu, 2018).

- (5) er that's it [laughter] okay thank you [laughter] very much (Presenter 1)
 (6) thank you very much (Presenter 2, 4, 5, 6)
 (7) thank you very much arigato gozaimasu (Presenter 7)

It is worth noticing Example 7, where Presenter 7, who was not Japanese, used both English and Japanese. This was possibly an attempt to connect with and engage the audience by alluding to their culture and showing deference to the country in which the event was hosted. This strategy was also used by Presenter 5 in the set-up move (Figure 2).

3.2. Intersemiotic relations

The analysis of the intersemiotic relations between speech and the content of the slides aimed to answer RQ2. Disclosing how they related helped us comprehend how meaning was conveyed and might be intended to facilitate audience's comprehension and engagement. First, we examined how speech and visuals were temporally and conceptually synchronised, i.e., the degree of alignment. Then, the focus was on the modal density of the slides and the intersemiotic relation of their presentation.

3.2.1. Synchronisation of speech and visuals

The content of the PK presentations was conveyed through 140 slides and 231 oral topic units. The intersemiotic analysis showed that 127 of the topic units were synchronised with the content of the slides. The total duration of the dataset was 48 min, and speech and visuals were synchronised for 36.30 min. Table 3 presents the percentage of (de)synchronisation of speech and visuals.

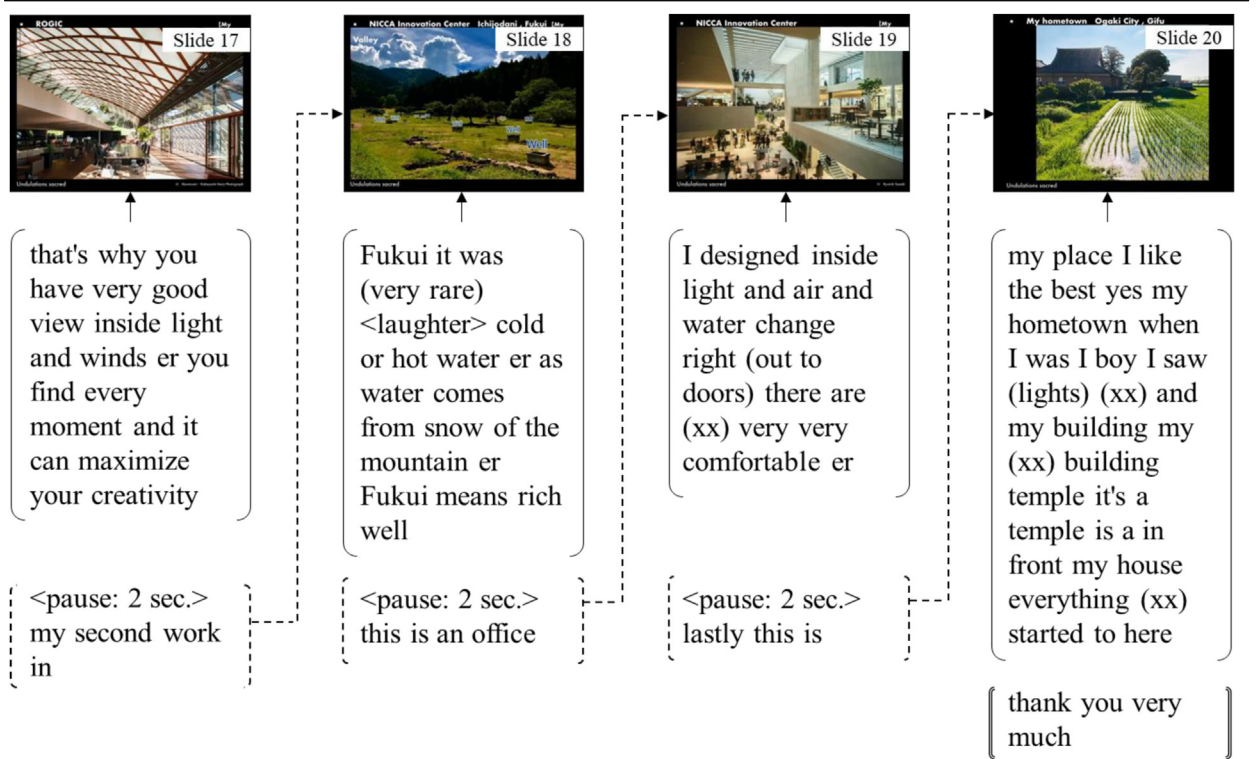
Findings revealed that around half of the oral topic units and visuals were desynchronised due to three different situations: the speech was behind, the visuals were behind, or the speech and the visuals were completely unrelated. When the speech was behind, the presenter was talking about a topic related to the content of a slide that was no longer on the screen. In contrast, when the visuals were behind, the topic of the speech was related to the content of a slide that had not yet appeared. The third desynchronised situation, unrelated speech and visuals, means that the speech had no visual support since the topic of the speech was unrelated to the content of the slides. The two first desynchronised situations occurred during the topic development move. However, the third type appeared during the set-up move in 4 presentations, during the topic development move in 1, and during the very brief closing move in all of them. We expected this lack of intersemiotic relations in such a blitz presentation where the slides appear automatically on the screen. Therefore, being time a crucial factor in PK presentations, its analysis could be significant to infer the influence that a lack (or not) of these intersemiotic relations may have on the audience.

Interestingly, the analysis of time distribution informed that PK presentations were mostly synchronised (75% of the total time). The results were encouraging for audience comprehension and engagement since a clear alignment of the topic of the speech and the content of the slides possibly favoured them. This degree of synchronisation could be particularly important for the audiences of these rapid-fire presentations where each slide is only 20 s in view and the time for processing information is short. Nonetheless, we foresee that the lack of intersemiotic relations could cause more difficulties in decoding the message when the speech was behind than when the presenter anticipated the content of the next slide. Figure 4 illustrates the different (de)synchronised situations.

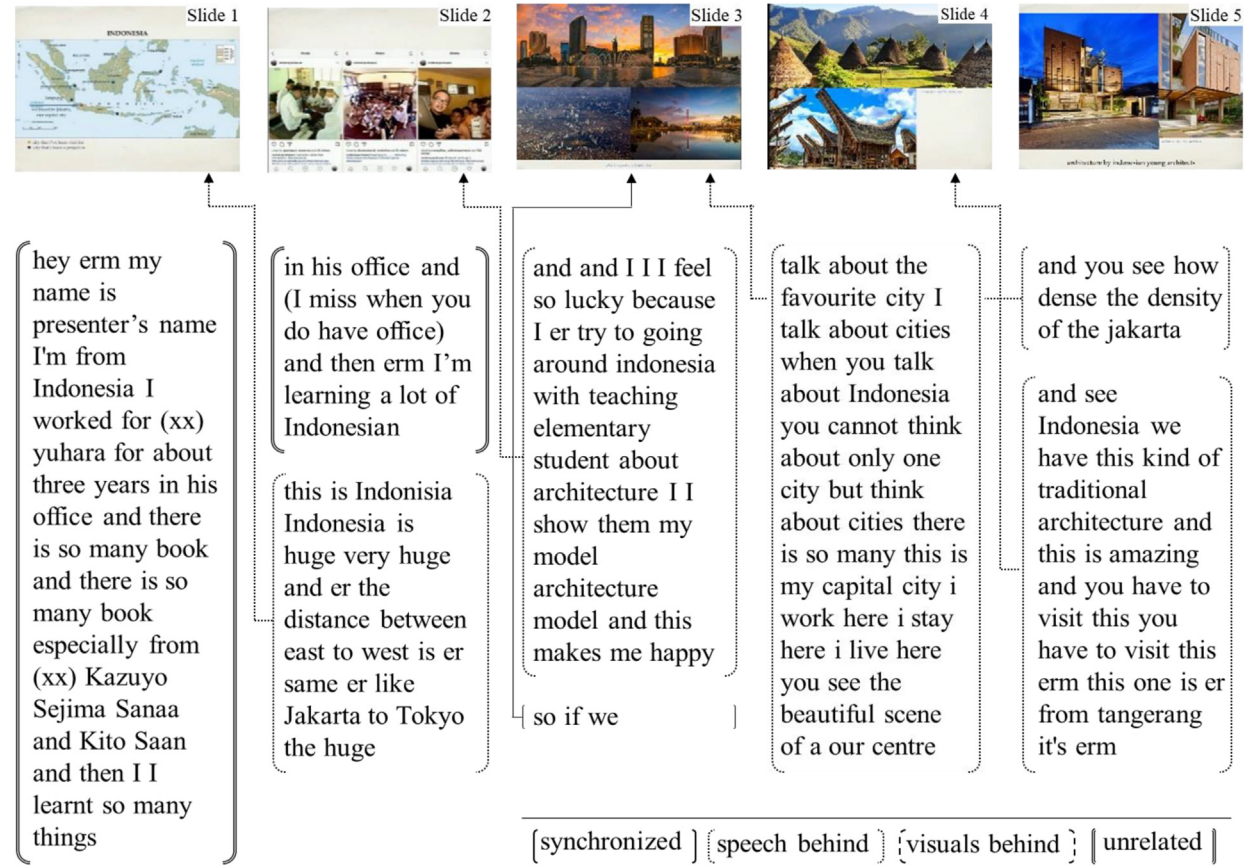
Speech and visuals were unrelated during the set-up (Presenter 7) and the closing (Presenter 6) moves. As mentioned, this lack of conceptual synchronisation was general in the dataset during the short closing stage as the presenters just thanked the audience for their attention ("thank you very much" mentioned while showing slide 20 – Presenter 6). The low information load of the last move seems to justify the presenters' decision of the presenters not to invest 20 s and one slide to perform this stage. The set-up move was a matter of a different order. When this move lasts around 20 s, we claim that a slide is needed to support the speech. Otherwise, the PK presentation may have a "bad" start with undesirable consequences for the rest of the presentation; i.e., the lack of intersemiotic relations could be accumulated as time and slides pass by. That is what happened to Presenter 7: the set-up move lasted almost 1.5 slides and, consequently, the speech was behind for half of the presentation (see slides 1–5). In particular, this lack of intersemiotic relations was even more evident when the presenter used deictics to refer to the content of the slide that was no longer on the screen. For example, he said "this is Indonesia Indonesia is huge very huge (...)" when slide 2 was on the screen; however, he was referring to the content of slide 1, where a map of Indonesia was shown. The misuse of "this" in this situation may be confusing for the audience, and it could be denoting that the presenter

Table 3
(De)synchronisation between speech and visuals.

	% frequency	% time
Synchronised	55.0	75.0
Desynchronised	45.0	25.0
Speech behind	13.4	15.1
Visuals behind	26.8	5.1
Unrelated speech and visuals	4.8	4.3



(Presenter 6)



(Presenter 7)

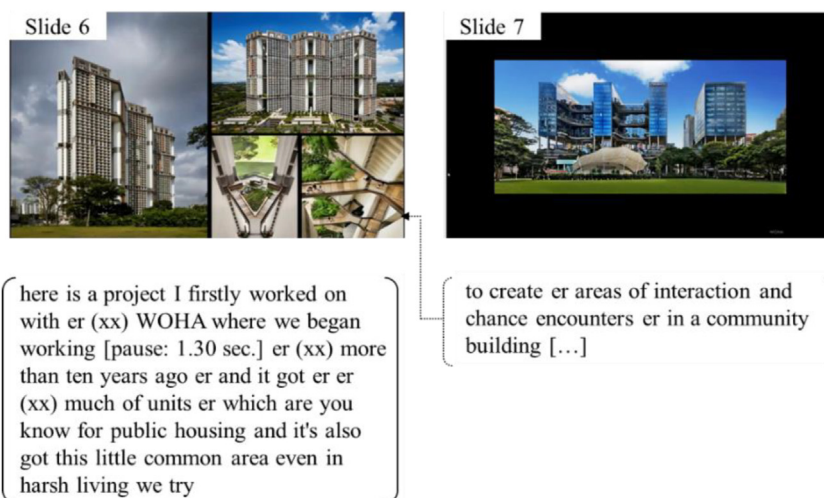
Figure 4. Synchronisation of speech and visuals.

possibly learned the speech by heart and did not adapt it to the new situation. Instead of trying to repair it by saying, e.g., “on the previous slide (...)”, he kept on talking as planned and overlooked what was shown on the screen. This may put the presenter and the audience in an uncomfortable communicative situation. On the one hand, although the presenter continued talking as if nothing had happened, the situation could have affected his emotional state and therefore his communicative skills. This is an issue that could become particularly worrisome for presenters when English is not their L1 but a lingua franca. On the other hand, possibly, a higher level of attention was required to follow the presentation than when speech and visuals were synchronised (see an example of synchronisation in Presenter 6, slides 17–20). Finally, when the visuals were behind, as the presenters were talking about something that was going to be visually supported straightaway, might not be problematic for comprehension and engagement. In this desynchronised situation, which in the dataset lasted between 0.30 and 5 s, presenters either kept on talking as if the next slide were in view; or waited in silence for the next slide for about 2 s, then, if it did not appear, they introduced it with expressions such as the ones used by Presenter 6: “my second work in” (at the end of slide 17), “this is an office” (slide 18), and “lastly this is” (slide 19). Here, the employment of “this” was probably less problematic than when the pronoun was used in a speech-behind situation since what it was referring to appeared on the screen immediately.

3.2.2. Modal density of the slides

We also approached the analysis of the intersemiotic relations through the study of the modal density of the slides and their presentation. A close examination of the content of the slides revealed frequent use of written content and picture(s) (36.6%) and only picture(s) (30.7%). The presenters also employed a combination of written and other visual content (22.6%) such as images, drawings, maps, diorama maps, blueprints, screenshots from Instagram/Twitter/Google, charts, graphs, and shapes (i.e., arrows and lines). In the remaining slides (12.8%), they used a map, prototype(s), a blueprint, only written content, and a picture and an image. We have classified the slides as having low, medium or high modal density, according to the number of communicative modes and their modal intensity. Slides with low modal density were designed using a maximum of two communicative modes having one of them low modal intensity (e.g., one slide with only a picture or one slide with a picture and few words). Slides with medium modal density involved two communicative modes with high modal intensity (e.g., a slide with one picture and different areas with written content). Finally, slides with high modal density were composed of more than two communicative modes with high modal intensity (e.g., a slide with drawings, a map, and different areas with written content). Findings showed that presenters mostly designed slides with low modal density (86.5%), if compared with medium (7.1%) and high (6.4%) modal density. Figure 5 illustrates the presentation of a slide with low modal density which involved the use of only pictures and speech.

As we can see, speaker 2 was talking about a project he worked on with WOHA, a prestigious Singaporean architecture studio, apparently well-known by the audience, which focuses on “conceiving integrated architectural and urban solutions to tackle the problems of the 21st century such as climate change, population growth and rapidly increasing urbanisation” (WOHA, 2023). The presenter conveyed the philosophy of the company through speech and visuals. He talked about public housing and areas of interaction in community buildings and illustrated these ideas visually with details that could not be expressed through words. Additionally, he also tried to engage the audience by uttering “you know”. The use of “you” in this example played the role of creating elaborate positive politeness, as it was identified in CP presentations (Polo Fernández, 2018), by emphasising in-group membership (Brown and Levinson, 1987). That is, the presenter’s response to the



(Presenter 2)

Figure 5. Intersemiotic relation between speech and slide with low modal density.

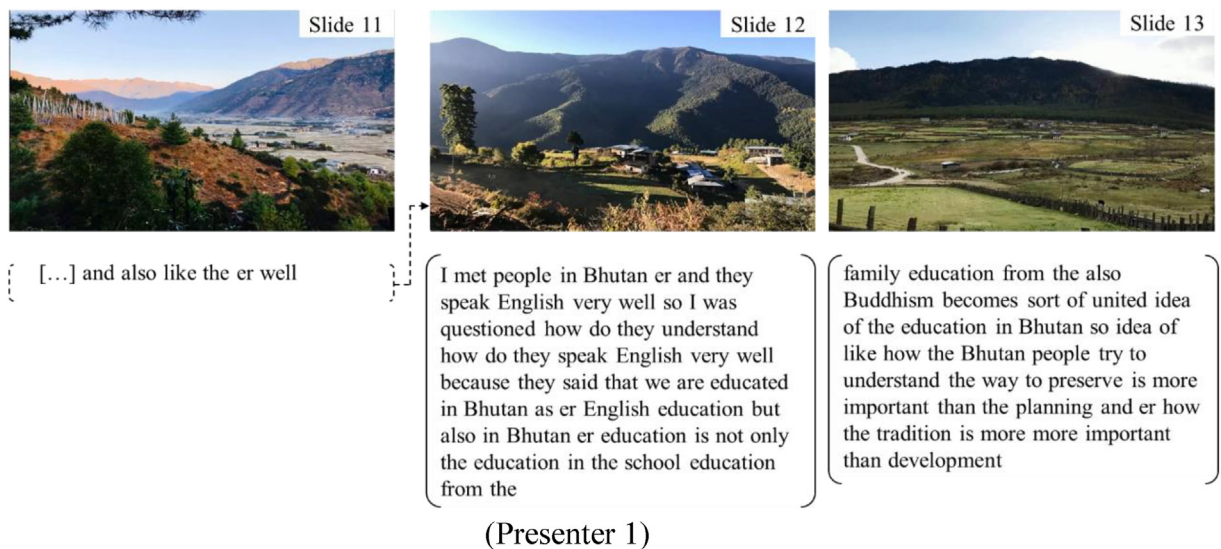


Figure 6. Lack of intersemiotic relation between speech and slide with low modal density.

audience's wish to be recognized as part of a group sharing the same wants. In this example, the presenter also provided information about his long relationship with WOHA ("more than ten years") showing also his professional identity.

Low modal density of the slides did not necessarily imply a clear intersemiotic relation between visuals and speech as in this example, which could facilitate comprehension and engagement. Thus, as illustrated in Figure 6, Presenter 1's speech was also conveyed through the visual support of only pictures (slides 12 and 13); nonetheless, there was a lack of intersemiotic relation between the two communicative modes.

The conceptual desynchronisation in this example could be confusing for the audience since there was not a direct connection between what he was saying and the pictures. This is an example in which the speakers' personal and professional experiences were interwoven. It illustrated how the two spheres were inseparable and fed off each other to build up the philosophy of this presenter's work. He explained that the people he met in Bhutan spoke good English and that education in this country comes from school, families and religion; apparently two unrelated topics to architecture. Then, he connected education in Bhutan with how they understand preservation and that "tradition is more important than development". This is indeed the philosophy of this presenter, as it was explained and illustrated with his projects later during the presentation: the idea of renewing traditional forms of architecture. "Tradition" was the only aspect that could be related to the visuals, but it was mentioned at the end of slide 13. Thus, the audience could be puzzled trying to find out the connection between speech and visuals. It seems the pictures in this example had the function of just contextualising speech, visually recalling to the audience how rural and traditional this country is.


Low modal density of the slides was also considered when a few written words and a picture were used (Figure 7).⁵

On slide 3, the presenter showed a picture he took at mount Yarigatake in Japan and explained how the experience of climbing it changed his life. This is another example of the presenter narrating a personal experience that later connected with the essence of his architecture designs, in this case, his love for nature and mountains. Thus, during this PK presentation, he showed how he integrated and tried to reproduce nature and undulations in his projects. The low density of the presentation of the slide and the intersemiotic relations, as well as the personal narrative, could have served to foster engagement with the audience.

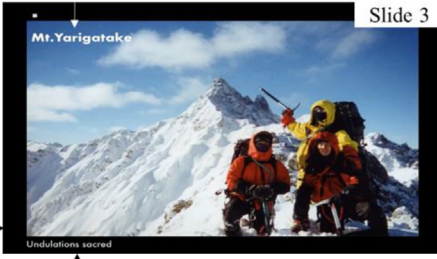
A medium modal density of the slides involved, for example, the use of more written content and pictures, as in Figure 8, slide 2.

In this example, Presenter 4 introduced her favourite city, Amsterdam, through personal and professional information, "my own city Amsterdam which I know very well I've lived and worked there for thirty years". Then, she referred to the visuals, "this is UNstudio in Amsterdam" ("UNstudio Amsterdam" was the only piece of information repeated on the slide), and described the study as being "very Dutch (...) very international (...) as Amsterdam is". Some of these features were somehow represented in the picture taken from outside of the studio in Amsterdam, as well as in the written content. For example, the international feature is visually conveyed with the data given about the number of nationalities and the reference to UNStudio Hong Kong. More detailed information was provided on the slide, such as that the presenter cofounded it, the number of projects, staff and nationalities, and that there are other two UNstudio in the world. That is, the slide provided more information than the speech. This was done through the use of a picture, a strategy also employed in longer presentations. It was striking, however, that the written content on the slide introduced information not mentioned orally. In

⁵ Hereafter, boldface letters indicate that the content is repeated on the slide and orally.



Slide 2



Slide 3


Undulations sacred

for people who love [...] this is **mount Yarigatake** [Pause: 2 sec.]


it changed my life (xx) I (xx) today
I climbed this mountain and discovered it's a beautiful (xx) one in nature and **undulation**
I took many pictures

(Presenter 6)

Figure 7. Intersemiotic relation between speech and slide with low modal density.



Slide 1



Slide 2

[...] but as my

favourite city I've chosen my own city Amsterdam which I know very well I've lived and worked there for thirty years this is **UNstudio in Amsterdam** and it's very Dutch practice in many ways but of course also very international practice which is also what Amsterdam is

(Presenter 4)

Figure 8. Intersemiotic relation between speech and slide with medium modal density.

longer presentations, the slides are used to support the speech, not the other way around. Presenters used keywords/phrases or short sentences that were explained orally. It is likely that, due to time constraints, some speakers used the slides to expand the content of their speech.

The slides with high modal density were not common in our dataset. Unlike low and medium modal density, high modal density of the slides may be challenging for the audience. Figure 9 illustrates the presentation of slide 4 by speaker 5.

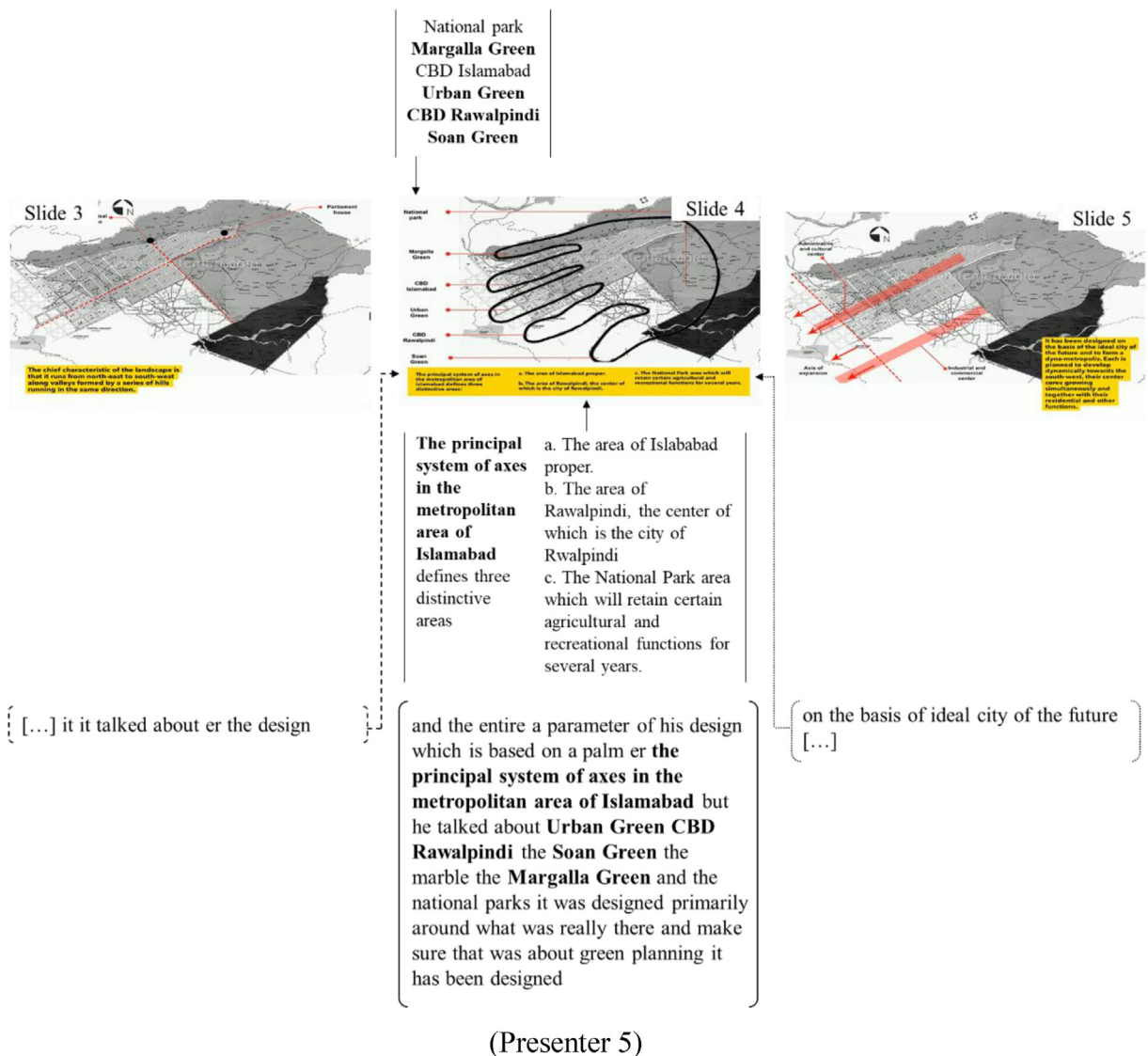


Figure 9. Lack of intersemiotic relation between speech and slide with high modal density.

The modal density of the slide was very high with 5 different communicative modes covering it: a map, which occupied 59.4% of the space; a drawing of a palm representing the city, 22.6%; written content in a yellow box, 11.1%; written content in labels for each finger of the palm, 1.2%; and the rest of the slide, blank space, which was only 5.7%. The design of this slide may be challenging for the audience's comprehension, as they are expected to process and relate multimodal information (speech and visual support) within a very short time frame. Besides, it could be confusing, not only because he just mentioned some of the written information included in the yellow box and the labels (in bold letters), but also because he referred to the labels in a different order.

4. Discussion and conclusions

In this study, we focused on the genre of PK presentations, characterised by specific time constraints and the number of slides (20 × 20). Specifically, we first identified the variety of rhetorical moves presenters used to construct PK presentations. Then, we explored how intersemiotic relations were deployed throughout these presentations to facilitate comprehension and foster engagement with the audience.

The first analysis showed that PK presentations consisted of three main rhetorical moves (set-up, topic development and closing) through which the presenters elaborated on their talk, shared their ideas and engaged the audience. Overall, the presenters took a more professional view, although they also expressed their personal positions. That is, the presenters built and shared with the audience a collective professional identity, as well as personal views and ideologies on the topic addressed. As for the set-up move, findings showed that it consisted of a series of optional steps that permitted the presenters to introduce and orient themselves towards the audience and the content. Results also revealed that the way the presenters constructed this rhetorical move could vary in terms of content, visual representation and length. More than half of the presenters performed the set-up move without visual support. Some moves within the setting up were presented just visually, others only orally, and others both visually and orally. In addition, while some introductions appeared to be rather brief, others were much longer. Despite the brevity of the set-up move, it may be problematic for the audience to understand the multimodal input and engage themselves in the presentation if visual support and speech are desynchronised. A different situation was observed in the closing move, which was found to be the shortest one. Due to its brevity, it was realised without visual support. Unlike the set-up move, adding visual support to end the presentation, even at the cost of missing intersemiotic relation, would not be necessary as it did not cause any comprehension problems for the audience. The longest rhetorical move was the topic development, which was central for the presenters to cover the main topic. This move can be conceived as a narrative in which the presenters shared their professional and personal experiences on architecture and their vision of future cities. Forming both a professional and personal identity is also a distinctive feature of the PK presentations analysed, which can be of relevance in professionally oriented PK presentations. Decisions regarding how to project presenters' identity may vary according to the type of PK presentations chosen (i.e., PK night, academic or professional, and for educational purposes).

The second analysis focused on intersemiotic relations, which was pivotal in identifying the complexity of synchronising speech and visuals, and the decision-making about the visual representation of content. For this purpose, we looked at both the degree of alignment (synchronisation) and modal density of slides. On the one hand, synchronisation was found to be critical to effectively transmit content and engage the audience. Nevertheless, desynchronisation was identified in three different situations: speech behind (the presenter talks about a topic no longer shown on slide), visual behind (the speech was related to a slide still not shown) and unrelated speech and visuals. The lack of synchronisation could negatively impact on the development of intersemiotic relations and, therefore, on audience understanding and engagement. As for the design of visuals, results showed that they were mainly based on pictures and brief written information, even though other types of visual input (e.g., images, maps, graphs) were included. Intersemiotic relations were also explored from the perspective of modal density, determined in terms of the number of communicative modes and modal intensity. We categorised modal density into three levels: low, medium or high. Each modal density level helped understanding and describing the way the presenters made use of communicative modes to convey meaning and engage their audiences. On a scale, these three parameters are gradually increasing according to the complexity and intensity of the communicative modes involved. For example, situations classified as low modal density consisted of, for example, one picture or a picture and few words, yet not necessarily with clearly connected speech (e.g., conceptual desynchronisation). Medium density would refer to situations with greater intensity of modes that contribute to the configuration of the overall communicative event. A construction classified as high modal density can be understood as particularly complicated for the audience. This is due to the intensity of the communicative modes intertwined on a single slide with a fixed duration of 20 s. In turn, this situation can lead to misunderstandings and a decreased level of engagement throughout the presentation.

This study served to unveil some of the major characteristics of PK presentations in terms of rhetorical structure and intersemiotic relations, which lends support to discuss pedagogical implications to enhance multimodal literacy in the ESP context. The development of students' multimodal literacy is found to be central for them to learn how to view, navigate and construct texts making use of different resources in a cohesive and coherent manner (Lim, 2018). Various approaches can be adopted to deal with genres from a multimodal perspective. To name a few, the multimodal genre-based approach (see Querol-Julián and Beltrán-Palanques, 2021), for an example based on PK presentations, the Learning by design framework (Cope and Kalantzis, 2015) extensively used to develop multiliteracies, or the Multimodality–Entextualisation Cycle (MEC) (Lin, 2016). Each of these approaches, with their unique specificities, has pedagogical potential for addressing genres in ESP and preparing students for effective communication. In the following lines, drawing on the findings of the study, we propose the design of a research-informed pedagogy to enhance ESP students' multimodal literacy through PK presentations.

Despite the brevity and apparent simplicity of PK presentations, they are much more intricate than they may appear and, therefore, it may be challenging for ESP students to construct a sample effectively. Regarding the rhetorical structure, as reported here, PK presentations involve three main different moves, with varying communicative purposes and length. Students should understand the communicative purposes of each of the moves and learn how to synchronise speech and visuals to transmit content and engage their audiences appropriately. It is advisable that students decide the type of content they intend to share and how the synchronisation between visuals and speech will be carried out effectively. This is applied not only to the topic development, in which most of the speech occurs but also to the set-up move where support can be necessary if it lasts about 20 s. Failing to synchronise both speech and visuals may impede the audience's comprehension and engagement, which is vital in this type of presentation.

Another key aspect to consider is that ESP students use English as a lingua franca, making it arguably more demanding for them to both transmit content and engage their audiences. Using a different language can, in fact, lead to increased stress and anxiety for students (Khan, 2015), which in turn can result not only in desynchronisation but also in the poor expression of

interpersonal meaning. For example, in our study, we found that some presenters not only failed to synchronise speech and visuals but also encountered language problems that could negatively affect the audience's overall comprehension and engagement. In addition to this, during the learning process, teachers should present students with both engagement strategies and repair strategies that can help them refocus discourse, if necessary. For instance, as shown in this study, the presenters shared both professional and personal information on the topic development move. While this may be used as a strategy to engage the audience, students in ESP may perhaps find it difficult to share at that point their professional identity as they are still developing it. Nevertheless, it would be relevant for them to be aware of the possibility of mixing professional and personal spaces throughout a PK presentation. On the other hand, making students aware of the relevant use of repair strategies (e.g., “as we saw on the previous slide”) may be also beneficial for them to refocus their discourse. As we found in this study, some presenters had problems connecting aural information with visual information as the slide shown corresponded to a different theme. This situation reflected desynchronisation, which should be avoided as much as possible so that the audience does not lose the thread of the speech.

Furthermore, the results also provided information on how intersemiotic relationships unfolded in terms of the modal density of the visuals. As findings suggest, modal density can be critical in the effective delivery and promotion of engagement through PK presentations. In this sense, ESP students should carefully think about how to represent visual content to make it comprehensible and engaging for the audience. The aim is to make students reflect on the modal density of the slides, which is associated with the complexity and intensity of the modes. Modal density can have a direct impact on the audience's comprehension and engagement (Querol-Julián and Fortanet-Gómez, 2025) as they are expected to decode multimodal input (speech and visuals) in only 20 s. As prospective presenters, students should try to design slides with a low or medium modal density that can facilitate the audience's decoding, comprehension and engagement.

Bearing in mind the complexity of designing and delivering PKs, we consider it especially important for students to devote some time to rehearsing. This can be of great help not only in familiarising themselves with time constraints but also with the complexities of adequately relating visual and auditory content, as well as determining whether the visual support selected may be suitable.

This paper invites reflection on both the analytical and pedagogical treatment of PK presentations. As reported, their analysis from a rhetorical and intersemiotic perspective has allowed us to understand and unveil some of their major characteristics. This study has some limitations. The number of PK presentations could be expanded to more effectively determine how presenters organize their discourse and construct intersemiotic relationships. Arguably, however, such a detailed multimodal analysis can only be conducted with small representative samples such as those presented in this study. Further research would be needed to explore other aspects of these presentations, such as the way presenters take stance and promote engagement with their audiences from a multimodal lens. Pedagogically, it would be necessary to explore both the implementation and assessment of multimodal literacy using PK presentations as a case in point. Therefore, this study is to be understood as a first approach to explore the genre of PK presentations from multimodal lenses and further research and pedagogical reflection and practice are needed to determine “good practices” in PK presentations.

Declarations of interest

None.

CRedit authorship contribution statement

Vicent Beltrán-Palanques: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Mercedes Querol-Julián:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Acknowledgements

The authors would like to thank PechaKucha Inc. for granting permission to use the images and transcriptions of the presentations included in this paper. Open access of this paper has been funded by Universitat Jaume I (Spain).

Appendix. Transcription conventions

(xx)	indicates that	one or more words that are completely unintelligible.
(words surrounded by parenthesis)		the transcription is uncertain.
<laughter>		the speaker is laughing.
[...]		some portion of the speech has been omitted.
[Audience: yes]		audience's response.
<pause: 2 s>		speaker's pauses

References

- Bateman, J.A., 2008. *Multimodality and Genre: A Foundation for the Systematic Analysis of Multimodal Documents*. Springer.
- Bateman, J.A., 2014. *Text and Image: A Critical Introduction to the Visual/verbal Divide*. Routledge.
- Beltrán-Palanques, V., Querol-Julián, M., 2018. English language learners' spoken interaction: what a multimodal perspective reveals about pragmatic competence. *System* 77, 70–80. <https://doi.org/10.1016/j.system.2018.01.008>.
- Beltrán-Palanques, V., 2024. Assessing video game narratives: implications for the assessment of multimodal literacy in ESP. *Assessing Writing*, 60, 100809. <https://doi.org/10.1016/j.asw.2024.100809>.
- Brown, P., Levinson, S.C., 1987. *Politeness: Some Universals in Language Usage*. Cambridge.
- Carter-Thomas, S., Rowley-Jolivet, E., 2020. Three Minute Thesis presentations: recontextualisation strategies in doctoral research. *J English Acad. Purposes*, 48. <https://doi.org/10.1016/j.jeap.2020.100897>.
- Chang, Y.J., Huang, H.T., 2015. Exploring TED talks as a pedagogical resource for oral presentations: a corpus-based move analysis. *English Teaching Learning*, 39 (4), 29–62. <https://doi.org/10.6330/ETL.2015.39.4.02>.
- Cope, B., Kalantzis, M., 2015. *A Pedagogy of Multiliteracies: Learning by Design*. Springer.
- Courtney-Klentzin, J., Bounds Paladino, E., Johnston, B., Devine, C., 2010. Pecha Kucha: using “lightning talk” in university instruction. *Reference Services Rev.* 38 (1), 158–167.
- Crawford Camiciottoli, B., 2019. Harnessing multimodal literacy for knowledge dissemination in ESP teaching. In: Bonsignori, V., Cappelli, G., Mattiello, E. (Eds.), *World of Words: Complexity, Creativity, and Conventionality in English Language, Literature and Culture*. Vol 1. Language. Pisa University Press, pp. 47–52.
- Diamantopoulou, S., Ørevik, S. (Eds.), 2022. *Multimodality in English Language Learning*. Routledge.
- Eisenlauer, V., Karatza, S., 2020. Multimodal literacies: media affordances, semiotic resources and discourse communities. *J. Visual Literacy*, 39 (3–4), 125–131. <https://doi.org/10.1080/1051144X.2020.1826224>.
- Feng, D., 2021. Genre, pedagogy, and PowerPoint design. A multimodal move analysis of linguistics lecture slides. In: LIN, K.L., Mwinlaaru, I.N., Tay, D. (Eds.), *Approaches to Specialized Genres*. Routledge, pp. 177–197.
- Frost, P., 2020. *Jazz on the Line. Improvisation in Practice*. Routledge.
- Hafner, C.A., Ho, W.Y.J., 2020. Assessing digital multimodal composing in second language writing: towards a process-based model. *J. Second Language Writing*, 47, 100710. <https://doi.org/10.1016/j.jslw.2020.100710>.
- Hafner, C.A., Miller, L., 2018. *English in the Disciplines: A Multidimensional Model for ESP Course Design*. Routledge.
- Halliday, M.A.K., 1978. *Language as Social Semiotic: the Social Interpretation of Language and Meaning*. Edward Arnold.
- Harinck, F., van Leeuwen, E., 2020. *The Art of Presenting. Delivering Successful Presentations in the Social Science and Humanities*. Cambridge University Press.
- Hiippala, T., 2014. Multimodal genre analysis. In: Norris, S., Maier, C.D. (Eds.), *Interactions, Images and Texts: A Reader in Multimodality*. De Gruyter Mouton, pp. 111–124. <https://doi.org/10.1515/9781614511175.111>.
- Hood, S., Forey, G., 2005. Introducing a conference paper: getting interpersonal with your audience. *J. English Academic Purposes*, 4 (4), 291–306. <https://doi.org/10.1016/j.jeap.2005.07.003>.
- Hu, G., Liu, Y., 2018. Three minute thesis presentations as an academic genre: a cross-disciplinary study of genre moves. *J. English Academic Purposes*, 35, 16–30. <https://doi.org/10.1016/j.jeap.2018.06.004>.
- Hyland, K., Zou, H., 2022. Pithy persuasion: engagement in 3 minute thesis presentations. *Applied Linguist.* 43 (1), 21–44. <https://doi.org/10.1093/applin/amab017>.
- Hyland, K., 2005. Stance and engagement: a model of interaction in academic discourse. *Discourse Studies*, 7, 173–192. <https://doi.org/10.1177/1461445605050365>.
- Iedema, R., 2001. Analysing film and television: a social semiotic account of hospital: an unhealthy business. In: van Leeuwen, T., Jewitt, C. (Eds.), *Handbook of Visual Analysis*. Sage, pp. 183–204.
- Incelli, E.A., 2021. Engaging students in multimodal literacy practices in a university ESP context: towards understanding identity and ideology in government debates. *Multimodal Communication*, 11 (1), 49–61. <https://doi.org/10.1515/mc-2021-0004>.
- Jewitt, C., Kress, G. (Eds.), 2003. *Multimodal Literacy*. Peter Lang.
- Jewitt, C., Bezemer, J., O'Halloran, K., 2016. *Introducing Multimodality*. Routledge.
- Jiang, L., 2017. The affordances of digital multimodal composing for EFL learning. *ELT J.* 71 (4), 413–422. <https://doi.org/10.1093/elt/ccw098>.
- Jiang, L., Yu, S., Zhao, Y., 2019. Teacher engagement with digital multimodal composing in a Chinese tertiary EFL curriculum. *Language Teaching Res.* 25 (4), 613–632. <https://doi.org/10.1177/1362168819864975>.
- Khan, S.M., 2015. Influence of speech anxiety on oral communication skills among ESL/EFL learners. *Adv. Language Literary Studies*, 6 (6), 49–53.
- Kim, Y., Belcher, D., 2020. Multimodal composing and traditional essays: linguistic performance and learner perceptions. *REL C J.* 51 (1), 86–100. <https://doi.org/10.1177/0033688220906943>.
- Kress, G., 2010. *Multimodality. A Social Semiotic Approach to Contemporary Communication*. Routledge.
- Lim, F.V., 2018. Developing a systemic functional approach to teach multimodal literacy. *Functional Linguist.* 5 (13), 1–17. <https://doi.org/10.1186/s40554-018-0066-8>.
- Lim, F.V., 2019. Investigating intersemiosis: a systemic functional multimodal discourse analysis of the relationship between language and gesture in classroom discourse. *Visual Communication*, 20 (1), 34–58. <https://doi.org/10.1177/147035721882069>.
- Lim, F.V., O'Halloran, K.L., 2012. The ideal teacher: analysis of a teacher recruitment advertisement. *Semiotica* 189, 229–253. <https://doi.org/10.1515/sem-2012-0032>.
- Lim, F.V., Tan-Chia, L., 2022. *Designing Learning for Multimodal Literacy Teaching Viewing and Representing*. Routledge.
- Lim, F.V., Weninger, C., Chia, A., Nguyen, T.T.H., Tan, J.M., Adams, J.L., Tan-Chia, L., Peters, C.M., Towndrow, P.A., Unsworth, L., 2020. *Multiliteracies in the Singapore English Language Classroom: Perceptions and Practices*. National Institute of Education, Nanyang Technological University.
- Lin, A.M.Y., 2016. *Language Across the Curriculum & CLL in English as an Additional Language (EAL) Contexts: Theory and Practice*. Springer.
- Luzón, M.J., Pérez-Llantada, C., 2022. *Digital Genres in Academic Knowledge Production and Communication: Perspectives and Practices*. Multilingual Matters.
- Luzón, M.J., Pérez-Llantada, C. (Eds.), 2019. *Science Communication on the Internet. Old Genres Meet New Genres*. John Benjamins.
- Luzón, M.J., 2023. Multimodal practices of research groups in Twitter: an analysis of stance and engagement. *English for Specific Purposes*, 70, 17–32. <https://doi.org/10.1016/j.esp.2022.10.006>.
- Maagerø, E., Tønnessen, S., 2022. Multimodal literacy in English as an additional language. In: Diamantopoulou, S., Ørevik, S. (Eds.), *Multimodality in English Language Learning*. Routledge, pp. 27–38.
- MacKrell, K., Silvester, C., Pennebaker, J.W., Petrie, K.J., 2021. What makes an idea worth spreading? Language markers of popularity in TED talks by academics and other speakers. *J. Assoc. Inform. Sci. Technol.* 72 (8), 1028–1038. <https://doi.org/10.1002/asi.24471>.
- Morell, T., 2015. International conference paper presentations: a multimodal analysis to determine effectiveness. *English for Specific Purposes*, 37, 137–150. <https://doi.org/10.1016/j.esp.2014.10.002>.
- New London Group, 1996. *A pedagogy of multiliteracies: designing social futures*. *Harvard Educational Rev.* 66 (1), 60–92.
- Norris, S., 2004. *Analyzing Multimodal Interaction: A Methodological Framework*. Routledge.
- O'Halloran, K.L., 2005. *Mathematical Discourse: Language, Symbolism and Visual Images*. Continuum.

- O'Halloran, K.L., 2021. Multimodal discourse analysis. In: Hyland, K., Paltridge, B., Wong, L. (Eds.), *The Bloomsbury Handbook of Discourse Analysis*, 2a ed. Continuum, pp. 249–266.
- O'Halloran, K.L., Lim, F.V., 2011. Dimensions of multimodal literacy. *Viden om Læsning, (Knowledge about Reading)*. 10, 14–21.
- O'Halloran, K.L., Lim, F.V., 2009. Sequential visual discourse frames. In: Ventola, E., Guijarro, A.J.M. (Eds.), *The World Told and the World Shown: Multi-semiotic Issues*. Palgrave Macmillan, pp. 139–156.
- PechaKucha, 2023, February 19. About PechaKucha. <https://www.pechakucha.com/about>.
- Polo Fernández, J.F., 2018. Functions of “you” in conference presentations. *English for Specific Purposes*. 49, 14–25. <https://doi.org/10.1016/j.esp.2017.10.001>.
- Querol-Julián, M., Beltrán-Palanques, V., 2021. PechaKucha presentation to develop multimodal communicative competence in ESP and EMI live online lectures: a team-teaching proposal. *CALL-EJ* 22 (2), 73–90. <http://callej.org/journal/22-2/Querol-Julian-Beltran-Palanques2021.pdf>.
- Querol-Julián, M., Fortanet-Gómez, I., 2019. Collaborative teaching and learning of interactive multimodal spoken academic genres for doctoral students. *Int. J. English Stud.* 19 (2), 61–82. <https://doi.org/10.6018/ijes.348911>.
- Querol-Julián, M., 2023. Multimodal interaction in English-medium instruction: how does a lecturer promote and enhance students' participation in a live online lecture? *J. English Academic Purposes*. 61, 101207. <https://doi.org/10.1016/j.jeap.2022.101207>.
- Querol-Julián, M., Fortanet-Gómez, I. (Eds.), 2025. *Multimodal Literacy in English as an Additional Language in Higher Education*. Routledge.
- Royce, T., 1998. Synergy on the page: exploring intersemiotic complementarity in page-based multimodal text. *JASFL Occasional Articles*. 1, 25–49.
- Ruiz-Madrid, M.N., 2021. A multimodal discourse approach to research pitches. *J. English Academic Purposes*. 52. <https://doi.org/10.1016/j.jeap.2021.101003>.
- Swales, J., 1990. *Genre Analysis: English in Academic and Research Settings*. Cambridge University Press.
- Tardy, C.M., Swales, J., 2014. Genre analysis. In: Schneider, K.P., Barron, A. (Eds.), *Pragmatics of Discourse*. De Gruyter Mouton, pp. 165–188. <https://doi.org/10.1515/9783110214406-007>.
- Tognini-Bonelli, E., 2001. *Corpus Linguistics at Work*. John Benjamins.
- Towndrow, P.A., Nelson, M.E., Yusuf, W.F.B.M., 2013. Squaring literacy assessment with multimodal design: an analytic case for semiotic awareness. *J. Literacy Res.* 45 (4), 327–355. <https://doi.org/10.1177/1086296X13504155>.
- van Leeuwen, T., 2005. Multimodality, genre and design. In: Jones, R.H., Norris, S. (Eds.), *Discourse in Action. Introducing Mediated Discourse Analysis*. Routledge, pp. 73–94.
- van Leeuwen, T., 2017. Multimodal literacy. *Viden om læsning (Knowledge About Reading)*. 21, 4–11.
- Viera, C., Williams, S.A., 2020. Corpus analysis of engagement discourse strategies in academic presentations. *Research in Corpus Linguist.* 8 (1), 105–130. <https://doi.org/10.32714/ricl.08.01.07>.
- WOHA, 2023, February 19. WOHA about. <https://woha.net/woha/>.

Vicent Beltrán-Palanques is a Lecturer in the Department of English Studies, Universitat Jaume I (Castellón, Spain). He is a member of the GRAPE (Group for Research on Academic and Professional English) group. His research interests include discourse analysis, multimodal literacy, ESP/EAP pedagogy, English-medium instruction and teacher professional development.

Mercedes Querol-Julián is a Senior Lecturer in Applied Linguistics at the Department of English Teaching, Universidad Internacional de La Rioja (Spain). Her main research interests include multimodal literacy, multimodal analysis, digital interaction, digital teaching, English-medium instruction, and teacher professional development. She leads the research group Personal and Professional Development through Digital Genres (PRODIGI) at Universidad Internacional de La Rioja.