

Speech-Gesture-Slide (SGS) Interplay in Academic Oral Presentations

Tetyana Smotrova, National University of Singapore, Singapore
Anita Toh, National University of Singapore, Singapore

The Asian Conference on the Social Sciences 2024
Official Conference Proceedings

Abstract

The multimodal aspects of oral presentation (OP) skills have been an important focus in studies on public speech and academic communication (e.g., Hood & Forey, 2005; Morell, 2015). However, within this body of research, there has been a lack of pedagogical perspective on how to teach these multimodal aspects. Moreover, most existing studies consider either the use of gestures (e.g., Carney, 2014; Masi, 2019) or the use of slides in OPs (e.g., Dubois, 1980; Rowley-Jolivet, 2004) separately. Meanwhile, Harrison (2021) posits that it is the interplay between speech, gestures, and slides in an OP that helps to maintain the audience's attention and aid their comprehension. Building upon Harrison's work and other studies on multimodality in OPs, this pilot study analyses the speech-gesture-slide interplay in post-graduate students' OPs. In this talk, we present preliminary findings from a qualitative analysis of a small corpus of student OP videos. Our participants are post-graduate students from different disciplines taking an academic communication course at a public university in South-East Asia. Having identified the specific ways students use such functions of speech-gesture-slide interplay as “draw attention”, “depict”, “decompose”, “disclose”, and “animate”, we draw on Goodwin's (1995, 2000, 2014) framework to conclude that when presenting their research, students should strive to employ the different semiotic systems in a coherent way “so that they mutually elaborate each other in a way relevant to the accomplishment of the [communicative] actions” (Goodwin, 2014, p. 238). Other pedagogical recommendations for teaching multimodal aspects of OPs will also be discussed.

Keywords: Multimodal, Oral Presentation, English for Specific Purposes, Gesture, Teaching

iafor

The International Academic Forum
www.iafor.org

1. Introduction

Multimodal aspects of oral presentation (OP) skills have been an important focus in studies on public speech and academic communication (e.g., Hood & Forey, 2005; Morell, 2015). Within this body of research, some studies have examined conference-style presentations from multimodal discourse perspectives (e.g., Hood & Forey, 2005; Morell et al., 2008; Morell, 2015), others have specifically focused on presenters' use of hand gestures (e.g., Carney, 2014; Masi, 2019) and hand gestures in TED talks (e.g., García Pinar & Palleja Lopez, 2018; Masi, 2019, 2020; Wu & Qu, 2020). A number of studies have also analysed presenters' interaction with the visuals (Knoblauch, 2008, 2013; Morton, 2006; Rendle-Short, 2006; Dubois, 1980; Morell, 2015; Rowley-Jolivet, 2002). These studies point to the importance of the presenters' use of such multimodal aspects as hand gestures and visuals (especially the ones appearing on the slides) in conveying the content and engaging the audience. However, most of them consider the multimodal aspects of focus in separation from the other ones rather than in their interplay. Although some studies do examine a combination of modes, such as verbal and gesture/body language (Hood & Forey, 2005; Tsang, 2020), verbal and visuals/slides (Charles and Ventola, 2002; Fortanet Gómez & Edo Marzá, 2022; Rowley-Jolivet, 2012), and verbal and gaze (Ruiz-Madrid & Valeiras-Jurado, 2020), they only focus on two modes, rather than the interplay of multiple modes as a coherent whole.

Within the research on multimodality on OPs, there are only a few studies that take a pedagogical perspective on the way student presenters employ multimodal aspects in delivering their content to the audience. A relevant study by Morell, T., Garcia, M., & Sanchez, I. (2008) highlights the link between multimodality and language competence. It found that speakers with higher English language competence tend to use a greater variety and combination of modes (spoken, written, body language, image), whereas lower competence speakers seem more focused on the spoken mode, with less use of body movement. The study also discovered that many of its participants had not paid attention to the interactive features of their OPs before the study. The authors therefore conclude that courses designed to aid international English-speaking academics should emphasize the use of multimodality. Thus, more studies are needed to reveal how to teach the multimodal aspects of OPs in their interaction with each other.

One of the few studies that does take a pedagogical perspective in considering the use of spoken language, visuals on the slides, hand-gestures, eye-gaze shifts, and other body movement in academic presentations (a TED talk) has appeared recently - Harrison (2021). He emphasizes that it is the interplay of the different modes that facilitates engaging the audience and enhancing their understanding of the presentation. This "person-environment relation" is described in Goodwin's terms (2018) as an 'embodied participation framework', "a small ecology in which different signs in different media (talk, the gesturing body and objects in the world) dynamically interact with each other" (p. 199). This means that in OPs, according to Harrison (2021), students should be instructed to treat their presentation as 'showing' Streeck (2017) rather than just speaking, where showing is understood "as an environmentally-situated, embodied, enlanguaged, and socially implicative presentation skill" (Harrison, 2021, p. 17).

In sum, prior research on the use of multimodal aspects in OPs and on teaching such aspects to students points to the need of instructing them on how to effectively integrate different modalities as an ecology in conveying the meaning to the audience, and how to treat oral

presentations as a ‘single textual space’ (Rowley-Jolivet, 2002) where the different modalities interplay with each other, allowing the audience to perceive the OP as a coherent whole. To address that need, this study aims to identify how students use speech-gesture-slide (SGS)¹ interplay in their academic OPs. Based on this, the study intends to offer pedagogical recommendations on the ways to enhance students’ use of SGS interplay in academic OPs.

2. Methods

The data for the study were collected in a post-graduate communication course at a public university in Southeast Asia. The course is intended for Masters and first-year PhD students to help them improve their academic writing and oral presentation skills. In this course, the students are tasked to write a research proposal and based on this, prepare and deliver an oral presentation in the format of a conference talk. All student presentations are video recorded for evaluation and self-reflection purposes with the camera only capturing the presenter. Our participants were ten post-graduate students (two female and eight male students) from the following disciplines: nursing, physics, architecture, food science and technology, statistics and data sciences, microbiology and immunology, chemistry, business, and computing. Eight of them come from China, and two of them come from Singapore. To ensure confidentiality, all the participants’ real names have been replaced with assigned pseudonyms.

The video-recorded presentations were first viewed and annotated. Then, excerpts of interest were identified and transcribed in more detail using conventions adopted from McNeill’s (2005) gesture notation and Smotrova’s (2017) micro analysis of classroom interactions. For more precision in identifying SGS interplay, the video excerpts of interest were viewed in slow motion using QuickTime Player. The presentation videos were analysed qualitatively with the focus on the ways students use SGS interplay as described in detail below.

3. Analysis

In identifying the ways students employ SGS interplay in their presentations, we focused on the following most relevant multimodal aspects included in our annotations and transcripts: speech, visuals appearing on the slides, hand gestures, body orientation/movement (e. g., stepping forward or backward), and eye-gaze. To qualitatively analyse the interplay of these aspects, we classified the functions of SGS interplay, using the categories proposed by Harrison (2021), Streeck (2008, 2009, 2017), and Dubois (1980) as reflected in Table 1. According to this classification, in a presentation, the SGS interplay can help to “animate” a static object or process depicted on the slide. It can also help to “disclose” - make apparent the features not visible to the audience when they view the slide or enable to “decompose” – indicate the separate parts that make up an object. Finally, SGS interplay can help to “depict” – mime the action or object shown on the slide or “draw attention” by pointing to specific elements on the slide.

In our qualitative analysis of the ways the different modes were used by student-presenters interactively, we also considered whether the information was conveyed through these modes in a repetitive or rather, complementary way. This is where the verbal-visual configurations identified by Xia (2023), who had drawn on Unsworth (2008), became relevant (Table 2).

¹ The term was used in Harrison (2021) with the author’s note that it was initially suggested to him by Sotaro Kita

Purpose/Function	Description
Animate	Use gesture to show movement of inanimate, static, time-invariant objects on slides.
Disclose	Make apparent features not visible to audience
Decompose	Show separate parts that make up an object
Depict	Mime action or object; e.g. show width/length, miming turning of steering wheel
Draw attention	Use gesture or pointer (laser) to draw attention to specific elements/content on slide

Table 1: Functions of SGS interplay

Configuration	Description
Concurrence	The verbal part repeats the text reflected on the slide.
Complementarity Extension	The verbal part highlights and adds information to the visual on the slide.
Complementarity Enhancement	The visual mode provides more detailed information. The verbal mode reproduces this information in a simplified manner.

Table 2: Types of visual-verbal configurations (Xia, 2023)

4. Findings and Discussion

In this section, we will present our preliminary findings based on our qualitative analysis of the ten student presentations accomplished using the categories described above. We will begin with an excerpt that reflects the simplest SGS interplay, we will then present two excerpts that show a more intricate and complex SGS interplay, and we will conclude with two excerpts that reflect a breakdown or lack of coherence in using the SGS interplay. For the sake of space, we will present detailed transcripts only for the two excerpts that show a more intricate use of SGS interplay. The analysis below will further allow us to discuss implications for teaching effective SGS interplay and provide some pedagogical recommendations in the concluding section.

4.1 *Disclosing Meaning of the Visual in Speech*

In Excerpt 1: “Caregiving”, our participant Ryan, a student from nursing faculty, introduces his research project focused on caregiving in multigenerational families with low income. After asking the audience whether they have had some caregiving experiences in their families, Ryan next introduces a typical Singapore family and the common types of caregiving relationships existing in such a family. He does so by mainly engaging three modalities: images – photos of family members, verbal explanations, and pointing hand-gestures (Fig. 1).



Figure 1. Excerpt 1 “Caregiving”

Interestingly, the slide in Fig. 1 does not contain any text as the visuals are not labelled so that the audience has to rely exclusively on Ryan’s verbal explanations in understanding the relevance of the images to the presented topic. In this sense, Ryan’s speech fulfils the function of *disclosing* the meaning of what is depicted in the visual, and therefore, the relationship between the verbal and visual modes in this case is complementary. That is, the content of the speech, describing the types of caregiving relationships in a typical local family, complements the information conveyed by the images. According to Xia (2023), this is a “complementary extension”, where “...the speech highlights the visually presented information and provides additional information on the highlighted aspect.”

In revealing and explaining the images, Ryan also *draws the attention* of the audience to each of the images in order to make it clear which one he is talking about at the moment. He interacts with the visuals and the audience through his gaze direction as well by alternating gazing at the screen, at the audience, and then back at the screen. This is what Streek (2017) calls a “[s]plit orientation: directing attention, monitoring response” (p. 193) because explaining or “showing requires watching.”

To summarize, in this excerpt we observe that the SGS interplay employed by the student fulfils the functions of disclosing and drawing attention. In this interplay, the verbal and visual modalities complement each other, creating “a single textual space...which has to be processed as an integrated whole by the audience” (Rowley-Jolivet, 2002, p. 21). We will examine a more sophisticated way of using SGS interplay with more modalities involved in the section below.

4.2 Disclosing and Decomposing the Image

This section will present the analysis of student’s use of multiple modalities in a coherent and mutually complementary way so that several SGS interplay functions are fulfilled: *draw attention*, *depict*, *disclose*, and *decompose*. In Excerpt 2 “Urban Village”, Ziqi, a student from Business, introduces the central concept of her study: “urban village.” Throughout the excerpt, she exhibits an active verbal and embodied interaction with the image that she reveals on her slide.

1 And this is the downtown
2 *-takes a step towards screen*
3 *-stretches LH towards screen, palm open, holds*
4 *-gazes at screen*
5 area
6 *-turns face towards A*
7 *-gazes at A*
8 *-rotates RH, holding clicker, twice*
9 of Shenzhen
10 *-turns face to screen*
11 *-holds LH pointing at screen*
12 and this is the dark area
13 *-rotates LH palm, pointing at screen*
14 *-clicks to reveal highlighted area (with white outline) on image*
15 and those
16 *-reshapes LH to point with index finger at screen, moving it closer to image*
17 *-gazes at screen*
18 Red part
19 *-moves LH palm slightly upward & downward twice*
20 *-turns face to A*
21 is urban village
22 *-shifts gaze further to right*
23 *-RH holds pointing at screen*
24 it is located in a very central area
25 *-RH holds pointing at screen*
26 *-gazes at screen*
27 *-gazes at A*
28 in the city
29 *-steps back into initial position at right of screen*
30 *-gazes at screen*

Figure 2. Transcript of Excerpt 2 “Urban Village”

Preceding the excerpt, Ziqi starts to introduce the concept of urban village by providing its definition as she talks through the bullet points on her slide. Following that, she reveals the image that visually illustrates the urban village - a photo of the urban area in Shenzhen, a city in China.



Figure 3. Lines 1-4 “And this is the downtown”

Ziqi first draws the audience’s attention to the photo of the downtown area on her slide (lines 1-4; Fig. 3). She does that through body movement by taking a step towards the screen to engage the audience with the image through proxemics. Simultaneously, she uses a pointing gesture by stretching her left hand towards the screen, with palm open, facing the audience. Interestingly, Ziqi chooses to point with her hand rather than with the laser, which makes the pointing much more prominent in attracting the audience’s attention. She enhances this function by also directing her gaze towards the image on the slide.



Figure 4. Lines 5-8 “area”

At the word “area”, Ziqi turns her face and redirects her gaze towards the audience, (lines 5-8). This is again the “split orientation” observed in Excerpt 1, where the presenter simultaneously directs attention and monitors response. Having established the mutual gaze, Ziqi illustrates the meaning of the word “area” by using an iconic gesture (depiction) – rotating her right hand (holding the clicker) twice. This movement not only imitates the shape of the downtown area depicted in the image, but also highlights the boundaries of the space, which are not obvious from the photo. In line 12, Ziqi also specifies, “this is the dark area” and rotates her hand again to depict the area. Thus, the iconic gesture complements the image by disclosing the bounded nature of the downtown area.



Figure 5. Line 14: “area” -clicks to reveal highlighted area (with white outline) on image

Ziqi then approaches the crucial point in her explanation by revealing the animation added to the photo of the downtown area – a white outline with a red-coloured area within it (Fig. 5; line 14). This outlined area depicts the central concept - urban village and helps to *decompose* the image into two elements: downtown area and urban village. As she says “those”, Ziqi also makes her pointing gesture more specific by stretching forward her index finger (line 15). At the same time, she changes the proxemics by moving closer to the screen with her gaze directed at the image. In this way, Ziqi enhances the function of drawing the audience’s

attention, adding an upward and downward gesture with the pointing hand as she says “red part” (lines 18-19). She finally discloses in speech that this is “urban village” and gazes at the audience when introducing this central term (lines 21-22). At the end of her explanation, Ziqi changes the proxemics again by stepping back into her initial position, slightly away from the screen, as if signalling that the current explanation is complete (lines 28-29). She then moves on to the next slide.

We can see how in this excerpt, the student-presenter employed the different modalities in a coherent and mutually complementary way: the visual, proxemics, hand-gesture, and gaze-direction were all well-aligned to serve the communicative purpose of introducing the concept of urban village. The interplay of proxemics, gesture, and gaze served the function of *drawing attention*; the verbal helped to *disclose* the meaning of the image; the iconic gesture *depicted* the bounded nature of the downtown area, and the animation-the outline helped to *decompose* the image. In the excerpt below we will examine another function of SGS interplay – *animating* a diagram depicted on the slide through iconic gesture.

4.3 Animating the Diagram

This section will discuss the *animating* function of SGS interplay as performed in Excerpt 3 “Spin up – Spin down”. In this excerpt, Wei Jie, a student from Physics department with specialization in the cutting-edge area of quantum physics, takes considerable effort to explain his project to an inter-disciplinary audience. To fulfil this challenging task, Wei Jie extensively uses different types of multimodal resources, including visuals, abundant hand-gestures and other body movement, and even real objects to convey the meaning of his project to the audience from outside his discipline. We will see how the student *animates* the diagram and converts it into a dynamic depiction through iconic gesture.

Preceding the excerpt, Wei Jie poses the question central to his project: how to create the “cubit” - the quantum information unit, alternative to the familiar “bit” in traditional computers. He then indicates that his project proposes to use ground state molecules and shows the diagram that depicts two such molecules in a dipole state (Fig. 7). Their relative position is marked as “ r ”, and the angle is marked as “theta”. Wei Jie will then explain that under the influence of electric/magnetic field these two molecules can rotate or “spin.”

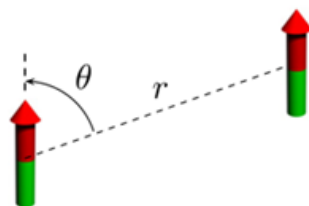


Figure 6. Two dipole molecules

1. The dipole moment could be controlled
2. *-faces screen, clicks to reveal image 2, makes beat with BH, RH index finger extended*
3. *-turns to face A, makes beat with BH at chest*
4. by uh extra electric field
5. *-turns and steps toward screen, extends BH upward toward diagram, palms open, facing each other,*
6. *-shifts body upward, "framing" part of image (p1 and angle)*
7. electric field
8. *-moves BH slightly away from screen surface, retaining the "frame"*
9. or magnetic field. So
10. *-holds BH over diagram, moves BH closer together towards angle*
11. you could see this theta here right,
12. *-moves RH holding clicker around angle twice*
13. *-swiftly points at theta symbol with RH index finger*
14. so we could use extra field
15. *-steps away from screen, faces A, rotates BH at chest*
16. to control theta
17. *-turns to screen, extends LH and points with index finger towards theta symbol*
18. to like maybe
19. *-moves BH back to chest, faces A*
20. spin down
21. *-shapes BH as fists, rotates BH: RH index finger points downward, LH index finger points upward*
22. spin up
23. *-rotates BH: now RH index finger points upward, LH index finger points downward*
24. (pause)
25. *-moves BH, index fingers extended, back to screen*
26. uh the molecule
27. *-brings Both pointing Hands back to chest*
28. up and down
29. *-holds BH, index fingers pointing upward*
30. *-flips RH, index finger pointing downward*
31. uh to like control
32. *-right index fingers imitates writing in the air*
33. the qubit, if you remember
34. *-points towards screen as if indicating position of previous diagram he's talking about*
35. the probability of the one state and zero state. So (pause)
36. *-points with RH index finger upward*
37. *-reshapes as feast (no pointing)*
38. that's the one way we could use the ground state molecule to make qubit
39. *-extends RH index finger, pointing at screen (in diagram's direction), makes beats*

Figure 7. Excerpt 3 "Spin up – Spin down"

The excerpt begins with Wei Jie’s crucial explanation that “the dipole moment could be controlled by extra electric field” (lines 1-9). During this explanation, he tightly interacts with the diagram to attract the audience’s attention. In lines 3-4, Wei Jie steps towards the screen and almost completely turns towards the screen, away from the audience. Importantly, he uses his both hands to “frame” the most relevant part of the diagram - the angle that is changed by the electric field (Fig. 8).

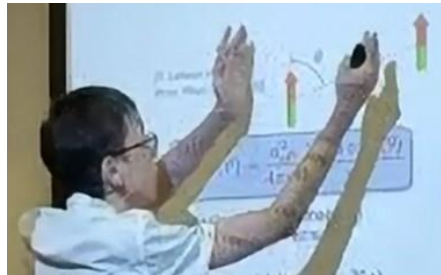


Figure 8. Line 4: “electric field”

This container gesture (palms facing each other), seems to fulfil three functions – iconically portraying the applied electric field and disclosing/drawing attention to the part that is most influenced by the field (the angle). Wei Jie next circles the angle itself as he says “so you could see this theta” (lines 11-12). After such a tight engagement with the screen, Wei Jie orients back to the audience as he steps away from the screen, faces the audience, and establish mutual eye-gaze with them as if checking whether they understood his complex embodied explanation.

Wei Jie then produces his crucial animation of the diagram through iconic gesture, disclosing its dynamic nature (lines 18-23). He transitions in speech “to like maybe” and positions himself facing the audience and prepares to produce the gesture by bringing both hands to chest (lines 18-19). Wei Jie then clenches both palms as fists with index fingers extended and rotates both wrists so that the index fingers alternately point upward and downward. This rotating movement is synchronised with the words “spin down; spin up” (lines 20-23).

In this iconic gesture, the student dynamically demonstrated the spin of the molecules under the impact of the electric field and *animated* the static diagram. With this gestural movement, he also converted a 2-dimensional image on screen into a 3-dimensional gestural movement, the movement which is impossible to be reflected dynamically in a visual. In this sense, he also *disclosed* the dynamic features hidden behind the 2-D image.



Figure 9. Line 20: “spin down”



Figure 10. Line 22: “spin up”

4.4 Showing Lack of Coherence in SGS Interplay

In this section we will briefly discuss the cases when, unlike the presenters above, students exhibit breakdowns or lack of coherence in using the multiple modes.

In Excerpt 4, “Norovirus”, a student from school of medicine, Winston, introduces the disease caused by norovirus. The textual part of the slide, which focuses on the ways of transmitting the disease, is accompanied by multiple, not labelled, images. The images are probably intended to illustrate the text by showing some symptoms and ways of the disease transmission. However, the other modes that Winston employs in presenting this slide do not seem to cohere with the visuals: the speaker does not refer to the images verbally and does not attempt to direct the audience’s attention to them through gesture or eye-gaze. In fact, his body and gaze are often directed away from the screen (Fig. 11), which signals lack of the “split orientation” observed in the excerpts above (drawing attention and monitoring response). For example, when Winston mentions such a symptom as diarrhoea, which is depicted on two images appearing on the slide, his body is turned away from the screen, while the gesture involves only container beat gestures, marking the rhythm of his speech. Thus, the images are not made part of “a single textual space” that should be created through coherent SGS interplay (Rowley-Jolivet, 2002).



Figure 11. Excerpt 4 “Norovirus”

A similar lack of coherence can be seen in Excerpt 5 “Time”, where a student from Physics, Kei Wen, presents the textual information reflected in multiple bullet points. These points are revealed all at the same time on the slide without the use of animation to help the audience focus. Even though Kei Wen uses multiple pointing gestures, he stands too far from the screen, and such proxemics seems to make it difficult for the audience to identify which specific line on the slide the presenter is pointing at. The challenge of understanding his content seems to increase once Kei Wen starts to explain the formula shown on the slide since his pointing is not close and specific enough for indicating the specific elements of the

formular. In this excerpt, we can see a lack of coherence in employing gestures and proxemics for drawing attention and decomposing the numerical visual.

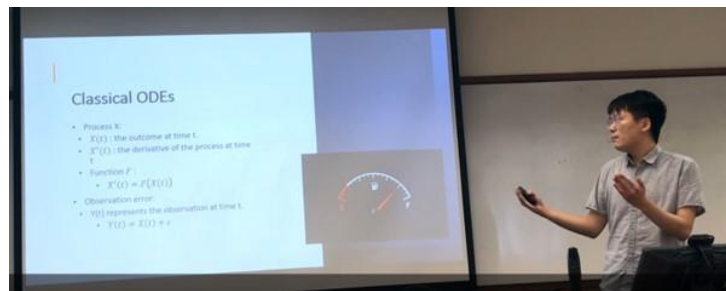


Figure 12. Excerpt 5 “Time”

5. Pedagogical Recommendations

Our findings indicate that in conveying the meaning of research projects, it is important for students to employ the different modalities as a coherent whole, a unified ecology in which these modalities complement and enhance each other. While some students are able to use the SGS interplay quite effectively in fulfilling its functions of *drawing attention, animating, disclosing, and decomposing*, other students exhibit lack of coherence in aligning the different modes to create a single textual space. This points to the necessity of making students aware of and instructing them on how the meanings of one element in the multimodal ensemble should emphasize, reinforce, or restate the meanings of another (Masi, 2020).

In the view of our findings, we suggest the following pedagogical recommendations for helping students integrate the speech, gestures, and slides coherently to convey the intended meaning effectively in their academic presentations.

1. Raise awareness about semiotic modes interaction.
Highlight the importance of the interaction between the different semiotic modes – speech, gestures, and slides – to help students understand how these elements work together to convey meaning effectively.
2. Provide specific instruction on SGS interplay functions.
Teach students the different functions of SGS interplay (animate, disclose, depict, decompose, and draw attention) with specific video excerpts to exemplify how effective presenters orchestrate different modalities to achieve these functions.
3. Analyse and discuss effective and ineffective uses of SGS interplay.
Show video excerpts of both effective and ineffective uses of SGS interplay and engage students in a discussion to identify best practices and common pitfalls; discuss how these impact audience engagement and comprehension.
4. Practice with peer feedback.
Conduct practice sessions where students present to their peers and receive constructive peer feedback. These sessions should focus on SGS interplay. Structured peer feedback forms could be used to ensure that the feedback is comprehensive and focused on the demonstration of coherent SGS interplay.

5. Video record presentations for self-evaluation.
Encourage students to video record their presentations and review these recordings with the help of structured reflection forms that focus on the purposeful and coherent use of speech, gesture and slide interplay.
6. Create a repository of videos for instructional purposes
It could be helpful for instructors to begin collecting and cataloguing video excerpts that exemplify effective and ineffective use of SGS interplay for future teaching purposes. This would ensure that there is a supply of relevant videos for students from different disciplines.

6. Conclusion

The literature has consistently highlighted the value of speakers' coherently employing different semiotic systems "so that they mutually elaborate each other in a way relevant to the accomplishment of the [communicative] actions" (Goodwin, 2014, p. 238). Our study has shown that while some students are more effective in achieving this goal, others need more guidance in this respect and therefore, it appears important to actively "sensitise and engage students more directly and implicitly with the ecology of oral presentations" (Harrison, 2021, p. 1) by raising awareness of the importance of SGS interplay in sense-making. We hope that our study and recommendations could provide helpful steps for instructors to help students demonstrate such SGS interplay more effectively in their academic oral presentations.

Acknowledgements

The authors are grateful to the Centre for English Language Communication (CELC), National University of Singapore, for its support of this study and for providing the funding necessary for our in-person participation at the conference.

References

- Carney, N. (2014). Is there a place for instructed gesture in EFL? In N. Sonda & A. Krause (Eds.), *JALT2013 Conference Proceedings* (pp. 413–421). https://jalt-publications.org/sites/default/files/pdf-article/jalt2013_041.pdf
- Charles, C., & Ventola, E. (2002). A multi-semiotic genre: The conference slide show. In E. Ventola, C. Shalom, & S. Thompson (Eds.), *The language of conferencing* (pp. 169–210). Peter Lang: Frankfurt am Main.
- Dubois, B. L. (1980). The use of slides in biomedical speeches. *The ESP Journal*, 1, 45–50. [https://doi.org/10.1016/0272-2380\(80\)90009-8](https://doi.org/10.1016/0272-2380(80)90009-8)
- Fortanet Gómez, I., & Edo Marzá, N. (2022). New genres and new approaches: Teaching and assessing product pitches from a multimodal perspective in the ESP classroom. *Porta Linguarum Revista Interuniversitaria de Didáctica de Las Lenguas Extranjeras*, 38, 65–81. <https://doi.org/10.30827/portalin.vi38.21621>
- García Pinar, A., & Palleja Lopez, C. (2018). TED talks: A multimodal tool for students of technological English. *Docencia e Investigación*, 29, 6–24. <https://repositorio.upct.es/server/api/core/bitstreams/e89962b8-aeec-4948-a4c7-2f04ca40b732/content>
- Goodwin, C. (1995). Co-Constructing meaning in conversations with an aphasic man. *Research on Language and Social Interaction*, 28(3), 233–260. <https://home.cs.colorado.edu/~martin/Csci6402/Papers/goodwin2.pdf>
- Goodwin, C. (2000). Pointing and the collaborative construction of meaning in aphasia. *Proceedings of the Seventh Annual Symposium About Language and Society*, 43, 67–76. <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=a95db04cc36d3855721c70daf6bc08106c60903f>
- Goodwin, C. (2014). The intelligibility of gesture within a framework of co-operative action. In M. Seyfeddinipur & M. Gullberg (Eds.), *From Gesture in Conversation to Visible Action as Utterance* (pp. 199–216). John Benjamins Publishing Company. <https://doi.org/10.1075/z.188.10goo>
- Goodwin, C. (2018). *Co-Operative Action*. Cambridge University Press.
- Harrison, S. (2021). Showing as sense-making in oral presentations: The speech-gesture-slide interplay in TED talks by Professor Brian Cox. *Journal of English for Academic Purposes*, 53, 101002. <https://doi.org/10.1016/j.jeap.2021.101002>
- Hood, S., & Forey, G. (2005). Introducing a conference paper: Getting interpersonal with your audience. *Journal of English for Academic Purposes*, 4(4), 291–306. <https://doi.org/10.1016/j.jeap.2005.07.003>
- Knoblauch, H. (2008). The performance of knowledge: Pointing and knowledge in powerpoint presentations. *Cultural Sociology*, 2(1), 75–97. <https://doi.org/10.1177/1749975507086275>

- Knoblauch, H. (2013). *PowerPoint, Communication, and the Knowledge Society*. Cambridge University Press.
- Masi, S. (2019). Complex mapping of words and gestures in TED Talks. In V. Bonsignori, G. Cappelli, & E. Mattiello (Eds.), *Worlds OF words: Complexity, creativity, and conventionality in English language, literature and culture* (Vol. 1, pp. 115–126). Pisa University Press. <https://arpi.unipi.it/bitstream/11568/1026717/9/Contributo%20Masi%20compresso.pdf>
- Masi, S. (2020). The multimodal representation of “Ideas worth spreading” through TED talks (1.0) [dataset]. University of Salento. <https://doi.org/10.1285/I22390359V36P155>
- McNeill, D. (2005). *Gesture and Thought*. University of Chicago Press. <https://doi.org/10.7208/chicago/9780226514642.001.0001>
- 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>
- Morell, T., García, M., & Sánchez, I. (2008). Multimodal strategies for effective academic presentations in English for non-native speakers. In R. Monroy & A. Sanchez (Eds.), *25 years of applied linguistics in Spain: Milestones and challenges* (pp. 557–568). Universidad de Murcia. <https://www.um.es/lacell/aesla/contenido/pdf/4/morell.pdf>
- Morton, J. (2006). The integration of images into architecture presentations: A semiotic analysis. *Art, Design & Communication in Higher Education*, 5(1), 21–37. https://doi.org/10.1386/adch.5.1.21_1
- Rendle-Short, J. (2006). *The Academic Presentation: Situated Talk in Action*. Routledge. <https://doi.org/10.4324/9781315241500>
- Rowley-Jolivet, E. (2002). Visual discourse in scientific conference papers: A genre-based study. *English for Specific Purposes*, 21(1), 19–40. [https://doi.org/10.1016/S0889-4906\(00\)00024-7](https://doi.org/10.1016/S0889-4906(00)00024-7)
- Rowley-Jolivet, E. (2004). Different visions, different visuals: A social semiotic analysis of field-specific visual composition in scientific conference presentations. *Visual Communication*, 3(2), 145–175. <https://doi.org/10.1177/147035704043038>
- Rowley-Jolivet, E. (2012). Oralising text slides in scientific conference presentations: A multimodal corpus analysis. In A. Boulton, S. Carter-Thomas, & E. Rowley-Jolivet (Eds.), *Studies in Corpus Linguistics* (Vol. 52, pp. 135–166). John Benjamins Publishing Company. <https://doi.org/10.1075/scl.52.06row>
- Ruiz-Madrid, N., & Valeiras-Jurado, J. (2020). Developing multimodal communicative competence in emerging academic and professional genres. *International Journal of English Studies*, 20(1), 27–50. <https://doi.org/10.6018/ijes.401481>

- Smotrova, T. (2017). Making pronunciation visible: Gesture in teaching pronunciation. *TESOL Quarterly*, 51(1), 59–89.
<https://www.jstor.org.libproxy1.nus.edu.sg/stable/44986977>
- Streeck, J. (2008). Depicting by gesture. *Gesture*, 8(3), 285–301.
<https://doi.org/10.1075/gest.8.3.02str>
- Streeck, J. (2009). *Gesturecraft: The manufacture of meaning* (Vol. 2). John Benjamins Publishing Company. <https://doi.org/10.1075/gs.2>
- Streeck, J. (2017). *Self-Making Man: A Day of Action, Life, and Language* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/9781139149341>
- Tsang, A. (2020). Enhancing learners' awareness of oral presentation (delivery) skills in the context of self-regulated learning. *Active Learning in Higher Education*, 21(1), 39–50.
<https://doi.org/10.1177/1469787417731214>
- Unsworth, L. (2008). Multiliteracies and metalanguage: Describing image/text relations as a resource for negotiating multimodal texts. In J. Coiro, M. Knobel, C. Lankshear, & D. J. Leu (Eds.), *Handbook of Research on New Literacies* (0 ed., pp. 395–424). Routledge. <https://doi.org/10.4324/9781410618894-24>
- Wu, A., & Qu, H. (2020). Multimodal analysis of video collections: Visual exploration of presentation techniques in TED Talks. *IEEE Transactions on Visualization and Computer Graphics*, 26(7), 2429–2442. <https://doi.org/10.1109/TVCG.2018.2889081>
- Xia, S. (2023). Explaining science to the non-specialist online audience: A multimodal genre analysis of TED talk videos. *English for Specific Purposes*, 70, 70–85.
<https://doi.org/10.1016/j.esp.2022.11.007>

Contact email: elcts@nus.edu.sg