
**Designing Knowledge Dissemination in a Digital Era – Analyzing TED Talks’
Multimodal Orchestration**

**Concevoir la diffusion des connaissances à l'ère numérique- Analyse de
l'orchestration multimodale de TED Talks**

*Jingxin Jiang, Chongqing Open University, China; National Institute of Education, Nanyang
Technological University, Singapore*

Fei Victor Lim, National Institute of Education, Nanyang Technological University, Singapore

Abstract

Online learning has gained increasing attention during the COVID-19 pandemic. Teachers face social exigencies to design ways of knowledge dissemination in online instruction. We posit that understanding how knowledge can be represented in successful online academic genres can inform teachers on how they can design students’ online learning experiences. This study examined how scientific knowledge is disseminated in one of the most widespread academic genres, TED Talks, which share discursal similarities with other academic genres such as online lectures. This study adopted a systemic functional multimodal discourse analysis approach to explore how a presenter used speech, images, and gestures to disseminate knowledge. The analysis shows that a presenter orchestrates speech, images, and gestures strategically to clarify the scientific ideas and engage the audience. Based on understanding how the three semiotic modes are used to disseminate scientific knowledge in accessible and engaging ways, this paper discusses how insights on multimodal orchestration can function as a heuristic tool to inform design in online learning.

Keywords: Online learning; TED Talks; Multimodal discourse analysis; Speech; Image; Gesture

Résumé

L'apprentissage en ligne a fait l'objet d'une attention croissante pendant la pandémie de COVID-19. Les enseignants sont confrontés à des exigences sociales pour concevoir des moyens de diffusion des connaissances dans l'enseignement en ligne. Nous postulons que la compréhension de la façon dont les connaissances peuvent être représentées dans des genres académiques en ligne réussis peut informer les enseignants sur la façon dont ils peuvent concevoir les expériences d'apprentissage en ligne des étudiants. Notre étude a examiné la manière dont les connaissances scientifiques sont diffusées dans l'un des genres académiques les plus répandus, les TED Talks, qui partagent des similitudes discursives avec d'autres genres académiques tels que les cours magistraux en ligne. Cette étude a adopté une approche systémique d'analyse fonctionnelle du discours multimodal pour explorer la manière dont un présentateur utilise la parole, les images et les gestes pour diffuser des connaissances. L'analyse montre qu'un présentateur orchestre la parole, les images et les gestes de manière stratégique afin de clarifier les idées scientifiques et d'impliquer le groupe. Sur la base de la compréhension de la façon dont les trois modes sémiotiques sont utilisés pour diffuser des connaissances scientifiques de manière accessible et engageante, cet article discute de la façon dont les idées sur l'orchestration multimodale peuvent fonctionner comme un outil heuristique pour informer la conception de l'apprentissage en ligne.

Mots-clés : apprentissage en ligne ; TED Talks ; analyse multimodale du discours ; discours ; image ; geste

Introduction

Advancement in digital technology and the prevalence of new media have led to a fundamental change in the education sector, with online learning and teaching now considered at the forefront of 21st century pedagogy (Ali, 2020; Miller et al., 2020). The outbreak of COVID-19 has propelled a massive transition from face-to-face classrooms to online learning systems around the world and has made online learning a new normal (Martin, 2021; Teo et al., 2021). However, the effectiveness of online learning in the pandemic is debatable (Adedoyin & Soykan, 2020; Hamid et al., 2020). For example, many students are dissatisfied with their online learning experiences (Maqableh & Alia, 2021) and report difficulties understanding the lecturers' instruction and materials (Allo, 2020). Meanwhile, teachers and instructors are also faced with challenges such as how to communicate their content in online instruction, how to keep students engaged in online learning, and how to use multimedia resources effectively (Oyedotun, 2020; Pei & Wu, 2019).

It has been widely accepted that students' learning experiences are more influenced by the design of instruction and materials than by the medium and technology deployed (Ally, 2008; Mayer, 2019). For online learning especially, one important factor to consider is what resources are available and how these resources can be used to make instruction more accessible and engaging (Niess, 2012; Shalev-Shwartz, 2011). Investigating how knowledge is represented in successful online academic genres can help teachers and instructors use the various semiotic modes afforded by online platforms to disseminate knowledge well. To this end, this study examined how scientific knowledge is disseminated multimodally in one of the most widespread academic genres, the TED Talks. We argue that given the discursial similarities between TED Talks and online lectures, understanding the ways of multimodal orchestration can function as a heuristic tool to inform the design of knowledge dissemination in online learning.

TED is the acronym for technology, entertainment, and design. It is a non-profit organization that aims to spread inspired ideas and has gained centrality in disseminating scientific knowledge. It provides a platform for scientists and researchers from different disciplines to popularize their latest research, inventions, and creative ideas¹. With their easy accessibility, TED Talks have long been used in teaching and learning and recognized as an academic genre (D'Avanzo, 2015; Harrison, 2021). As a prestigious genre, TED Talks share discursial similarities in knowledge representation with other academic genres such as academic oral presentations and lectures (Chang & Huang, 2015; Harrison, 2021). Thus, the delivery style and techniques can shed light on the designing of knowledge dissemination in online lectures and inform teachers and instructors how semiotic modes ("a socially organized set of semiotic resources for making meaning" such as image, speech, and gesture [Jewitt et al., 2016, p.157]) can be orchestrated in accessible and engaging ways in online learning. Taking one of the most popular TED Talks (based on number of views) as an illustrative example, this study adopted a systemic functional multimodal discourse analysis approach to explore how the presenter orchestrated different semiotic modes—speech, images on slides, and hand gestures—to achieve effective knowledge dissemination, and discuss how this can inform teachers' and instructors' design in online learning. In this paper, we use the terms *presenter* to indicate the speaker who delivers the TED Talk and *audience* to indicate the people who view the talks. While TED Talks are open to the public, most audience members are non-expert viewers, so they are often referred to as the lay audience or lay public. The term *participant* indicates the persons (or entities) who perform an action in the analysis of the meaning representation by the three modes.

¹ www.ted.com

Research Background

Online Learning

Online learning, also referred to as e-learning, digital learning, and computer-based learning, has been broadly defined as “instruction delivered on a digital device that is intended to support learning” (Mayer, 2019, p. 152). It involves using the Internet and digital technologies to access the learning materials and the instructional delivery to achieve effective learning (Adedoyin & Soykan, 2020; Ally, 2008). Online learning can be either synchronous or asynchronous. Synchronous online learning involves real-time engagement between the teacher and students in a virtual classroom. Asynchronous online learning is self-paced and students can access the prepared digital instructional materials and perform learning activities at their preferred time (Ogbonna et al., 2019). Both synchronous and asynchronous online learning involve teachers’ use of different semiotic resources, like slides and gestures to accompany speech in an online lecture or videos in asynchronous learning.

Mayer (2002) and Mayer and Moreno (2002) developed the theory of multimedia learning which suggests that people’s processing capacity of working memory is limited and, hence, it is important to reduce students’ cognitive load when designing instructions (Bannert, 2002; Sweller, 2005). The cognitive theory of multimedia learning assumes that people have different apparatus in processing different semiotic modes (speech or images). Different semiotic modes can be used synchronically to disseminate knowledge. As processing capacity is limited in working memory, effective learning occurs when people are involved with appropriate cognitive processing, which necessitates the strategic orchestration of different semiotic modes when designing online instruction (Mayer, 2019).

TED Talks as an Academic Genre

A *genre* refers to a series of communicative events with common purposes agreed upon by a particular group of people (Swales, 1990). The aim of a genre is to realize its communicative intention with recognizable generic features within a community. Many genres are used for academic purposes, such as lectures, online courses, academic presentations, and science popularization. As an academic genre, TED Talks share rhetorical and discursal features with other academic genres. For instance, the vocabulary representation, lexical density, and speech rate of TED Talks are similar to those in lectures (Wingrove, 2017). They also resemble university lectures in the use of words and multimedia resources to express a speaker’s stance and avoid misunderstanding (D’Avanzo, 2015; Mattiello, 2019). With the common communicative purpose of knowledge dissemination in an online medium, we argue that a discourse analysis of TED Talks can inform the ways online learning can be designed.

Presentations are also a common way to disseminate knowledge in academic genres, such as seminars, lectures, and classroom teaching (Tardy, 2005). Because of their easy accessibility, TED Talks have been used as models of effective communication. For instance, in Li et al. (2016), TED Talks were used to teach the organization of ideas in oral presentations and were reported to be effective in enhancing students' thinking ability and presentation skills. In Salem (2019), TED Talks were used as model presentations in a business English presentation course, and students' presentation skills improved. TED Talks are thus a useful resource for teaching the design of oral presentations in physical classrooms. In the same way, TED Talks can be useful in online instruction, which has similar purposes and audience. We posit that the findings from the multimodal analysis of TED Talks can also illuminate the ways different semiotic modes can be used in online instruction, such as the representation of ideas in a presentation during synchronous online learning and a recorded video presentation as material for asynchronous online learning.

Use of Multimodality in Knowledge Dissemination

One commonality shared between TED Talks and other academic genres is the crucial role of multimodality to disseminate knowledge (Jewitt et al., 2016). Previous studies have found that images (including photos, tables, and scripts) play an essential role in representing scientific knowledge (Tardy, 2005) and can help an audience get a better understanding of texts (Bucchi & Saracino, 2016; Rowley-Jolivet, 2012). There are also many studies investigating the use of gestures in face-to-face classroom teaching. For example, Alibali et al. (2014) examined teacher's gestures in middle school mathematics classes and claimed that "gestures are an integral part of teachers' communication during mathematics instruction" (p. 65). Chue et al. (2015) investigated the functions of gestures in teaching chemistry. They found that iconic gestures can help illustrate abstract ideas and present a more complete version of meaning than when represented by speech alone.

Previous Studies on the Multimodal Nature of TED Talks

The multimodal nature of TED Talks has been described in previous studies (Harrison, 2021; Masi, 2016, 2020). By analyzing the synchronization of gestures and speech, Masi (2016, 2020) observed that hand gestures are often used to reinforce meaning in abstract concepts and promote audience involvement. Harrison (2021) investigated the speech-gesture-slide interplay in two sample talks through the lens of showing as sense-making and found that the orchestration of the three modes helped the audience make sense of images on slides, understand, and remain engaged in the talk. However, the multimodal orchestration involved the interaction and interplay across semiotic modes and warrants further examination. In this study, we analyzed the orchestration of speech, images, and

hand gestures to describe the interplay and discuss how the orchestration can inform design in online learning. We addressed these research questions:

1. How does the presenter use speech, images, and gestures to disseminate scientific knowledge in the TED Talk?
2. How can the analysis of TED Talks serve as a heuristic to inform the design of knowledge dissemination in online learning?

Methods

A Systemic Functional Multimodal Discourse Analysis Approach

From the multimodal perspective, different semiotic modes have different sets of rhetorical devices and could be understood in their design and functions (Harrison, 2003). In systemic functional theory, language, in speech or writing, is regarded as a semiotic mode and a resource for meaning-making (Halliday, 1994). Language makes meanings in combination with other semiotic modes to form an integrated whole. This understanding is aligned with the interconnected cognitive processing in different channels in the cognitive theory of multimedia learning (Mayer, 2002). The systemic functional multimodal discourse analysis approach (SF-MDA) (O'Halloran, 2008) was adopted in the present study. SF-MDA aims to understand and describe the functions of different semiotic modes as systems of meaning and the emergent meanings arising in the orchestration of these modes (Jewitt et al., 2016).

Framework to Annotate Speech

Semiotic modes serve the ideational metafunction to represent ideas and experiences, the interpersonal metafunction to enact social relation, and the textual metafunction to organize the discourse (Halliday, 1994). Systemic functional grammar (SFG) was adopted to annotate speech. In SFG, the ideational meanings are realized by the system of transitivity, which describes the type of action process that the participant conducts in a clause. Interpersonal meanings are realized by a system of mood block, mood types, and modality. Mood block includes two parts: the subject and the finite. The persons or entities who carry out the social practices (exchange of information or goods/service) are known as the subject, and the transaction's validity is described by the term finite. The verb used in exchanging information or goods/services is described by mood types, comprising declarative, interrogative, indicative, and imperative moods. Another indicator, modality, describes the extent of probability, usuality, obligation, and inclination. Textual meanings are realized by the system of theme

(the first constituent of a clause) and rheme (the remaining part of the clause), which indicates the first element of a clause, serving to examine the organization of a text.

Framework to Annotate Images

In images, ideational meanings are realized by representational structures, i.e., the relationship between participants in an image and the nature of the actions. Interpersonal meanings, in terms of “symbolic contact, social distance, power relations, and involvement between viewers and visual participants” (Feng & O’Halloran, 2013), are realized by how participants gaze at viewers, the shot distance (long, medium, or close), vertical camera angles (high, eye-level, or low), and horizontal camera angles (frontal or oblique), respectively. Textual meaning is expressed in the composition and layout of the image through three systems: information value, salience, and framing. Information value is viewed through the layout and placement (i.e., top and bottom, left and right, centre and margin), salience through the prominence and importance of visual elements, and framing through the connection and disconnection between visual components.

Framework to Annotate Gestures

Gestures were annotated according to Martinec’s (2000) scheme for classifying actions, including presenting action, representing action, and indexical action. Presenting actions indicate gestures that do not represent meaning, such as holding the hands in front of the body. Representing action has a conventional signifying function, which can express meaning. Based on the relationship between gesture and language, Lim (2019) classified representing action into two categories: language correspondent gestures and language independent gestures. Indexical action often occurs with speech and should be interpreted in synchronization with the utterances (Martinec, 2000). It is also described as language correspondent gesture (Lim, 2019). The ideational meanings of presenting actions, such as those in SFG, are realized through transitivity processes such as the material, behavioural, state, verbal, and mental processes. To describe the ideational meaning of representing actions, congruent entities (i.e., participants, processes, or circumstances) and metaphorical concepts are examined. Indexical actions are considered through importance, receptivity, and relation. The interpersonal meanings of gestures are annotated through dimensions of attitude (positive, negative, and neutral), engagement (expansion, contraction, neutral, and possibility), and graduation (fast, medium, and slow). Textual meanings in gestures are examined through wavelength (size and rhythm) and pointing (directionality and specificity).

How Orchestration Across Speech, Images, and Gestures Was Annotated

The orchestration of speech, images, and gestures was examined through the ways they interplayed and functioned. Informed by O'Halloran's (2008) framework in the study of text-image interplay and Lim's (2021) description of verbal-gestural interrelations, the presenter's orchestration of the three modes was annotated according to their meaning convergence and divergence. The interplay across semiotic modes can work together to express congruent meanings through convergence and different or even contradictory meanings through divergence (O'Halloran, 2008, p. 90).

The Data

This study is part of a larger research project which focuses on multimodal analysis of TED Talks. The video analyzed was downloaded from the TED website (<https://www.ted.com>). The talks on this website are classified using categories such as “topic,” “the most viewed,” and “the latest.” The talk under analysis is social psychologist Amy Cuddy's (2012) talk: *Your Body Language May Shape Who You Are* (https://www.ted.com/talks/amy_cuddy_your_body_language_may_shape_who_you_are). It explains how body language changes people's behaviour and consequently changes their life. It appears on the home page in the playlist “The most popular talks of all time,” with 63,730,150 views (7 January 2022) and can be considered representative of the TED Talks genre. It is also popular among viewers, suggesting interest and indicating that the talk is successful. Transcripts were annotated clause by clause, images slide by slide, and gestures action by action.

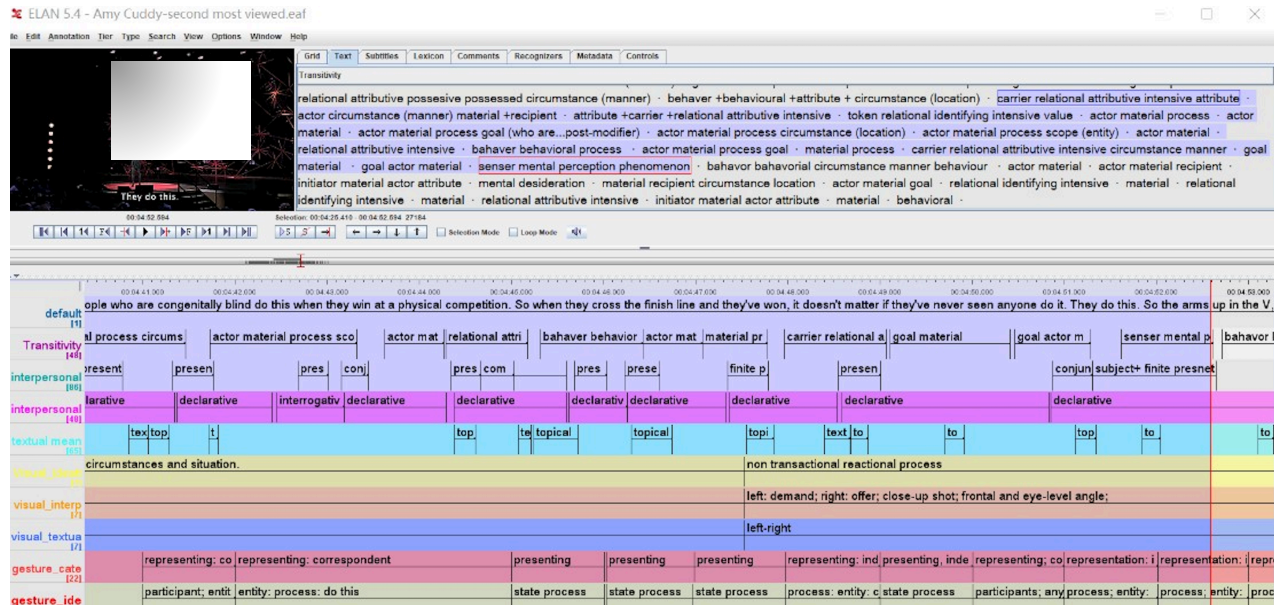
Data Analysis

In the study, we performed a detailed annotation of the data, with iterative viewing and zooming into instances of multimodal orchestration. Clauses of the transcripts and screenshots of special moments were extracted for fine-grained analysis and detailed description. Reproduction of screenshots in this paper is supported by the TED Talks Usage Policy and licensed under a Creative Commons license (CC BY–NC–ND 4.0 International [<https://www.ted.com/about/our-organization/our-policies-terms/ted-talks-usage-policy>]).

ELAN (Version 5.4) was used for the analysis (Figure 1). It provides functions, such as time-aligned annotation, fragmentation, and labelling. It is organized in tiers and is hierarchically interconnected, allowing exploration of the annotation's interplay on each tier (<https://tla.mpi.nl/tools/tla-tools/elan/elan-description/>). Documentation and search functions also aid in the iterative reading of data and identifying characteristics, as well as locating and capturing a screengrab at an interesting moment.

Figure 1

Screenshot of ELAN



Note. The image is a screenshot of ELAN (V.5.4)[Computer software]. (2018). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. <https://archive.mpi.nl/tla/elan/>

Findings

Our first research question focused on how the presenter used speech, images, and gestures to disseminate scientific knowledge in this TED Talk. TED Talks aim to spread science to a lay audience, making it necessary to represent sophisticated scientific ideas in an accessible way. Meanwhile, like other online academic genres, such as science news and online courses, TED Talks have long competed for an audience’s attention (Lorés, 2020). Most of the online audience bears a “scan and go” attitude (Zhang et al., 2015, p. 239), though it is vital to keep the audience involved. Thus, the presenter must orchestrate speech, images, and gestures effectively. In what follows, we describe the meanings made in the three modes, analyze their interplay, and interpret how multimodal orchestration contributes to the effective dissemination of scientific knowledge in the focal TED talk.

Use of Speech, Images, and Gestures to Clarify Meaning

In terms of speech, it can be seen from the annotation that the presenter used a large amount of material process in transitivity, that is, the process involving physical actions to explain concepts,

describe phenomena or observations, and report results and findings. For example, between 5:07 and 6:03 of the video clip, the presenter claimed that when we are in class and encounter a person using high-power gestures, we tend to make low-power gestures. She first described a series of actions that students of high power perform in the material process: “they get right into the middle of the room,” and “they raise their hands like this.” Then she discussed students’ low-power gestures, such as “they sit in their chairs, and they make themselves tiny.” The use of material process to describe specific actions can instantiate abstract concepts and vivify the description of phenomena or statements, and therefore help clarify ideas and facilitate audience understanding.

Correspondingly, images were displayed to support the meanings made in the speech. For example, to explain the concept of nonverbal expressions of power and dominance (3:48–5:09), the presenter deployed a series of images with participants performing a specific expanding or shrinking action. She first put forward the question of “What are nonverbal expressions of power and dominance?” to raise curiosity and appeal to the audience’s interest. In what followed, she answered this question using a guiding statement, “Well, this is what they are,” with a series of images (Figures 2 and 3). In most of these images, the participants’ gaze is directed away from the audience, forming a non-transactional reactional process, which allows the audience to focus on the actions displayed instead of participating in the event. While describing the images and elaborating the concept, the presenter used the two modes to form a meaning convergence. For instance, the presenter displayed an image (Figure 4) of a participant (the winner) in a running race. The winner is crossing the finish line while expanding his arms high in the air, forming the letter V. Combined with a detailed description of “the arms up in the V, the chin is slightly lifted,” the meaning of “non-verbal expression of pride” was vividly conveyed to the audience. The orchestration of the two modes clarified the concept, helped ease the burden in cognitive processing, and enhanced the intake of information and understanding of the content.

Figure 2

Access the Image of Participants with Expanding Gestures through:

https://www.ted.com/talks/amy_cuddy_your_body_language_may_shape_who_you_are at 4:01-4:09

Note. This image presents two gorillas. One gorilla is holding its two arms high up in the air, the other is standing with arms akimbo. The transcript reads, “So in the animal kingdom, they are about expanding. So you make yourself big, you stretch out, you take up space.” In the public domain. (Cuddy, 2012).

Figure 3

Access the Image of Participants with Contrasting Gestures through:

https://www.ted.com/talks/amy_cuddy_your_body_language_may_shape_who_you_are at 5:10-5:19

Note. This image presents two people. The lady is standing with arms akimbo, and the man is standing with two hands together in front of the body. The transcript reads, “And this is what happens when you put together high and low power. So, what we tend to do when it comes to power is that we complement the other’s nonverbals.” In the public domain. (Cuddy, 2012).

Meanwhile, when speech and gestures were exhibited in the video, hand actions were also used to disambiguate meaning. The language independent actions were mostly used to exemplify abstract concepts and make statements. When explaining the concepts, the presenter performed corresponding gestures, such as expanding (Figure 5) to demonstrate high power status and wrapping two arms in front of the chest (Figure 6) to show a lower power status. These gestures, in themselves, could express meanings of power and timidity. The meanings represented in the gestures concurred with those made in the speech: “really power” and “make ourselves small.” Thus, speech and gestures converged to form a whole and reinforce the meaning reciprocally in the two modes. In this way, the use of gestures allowed the audience to visualize and experience the specific and vivid “non-verbal expressions of power and dominance” expressed in the speech and enhanced the audience’s understanding of these concepts.

Figure 4

Access the Image of a Winner in a Running Race through:

https://www.ted.com/talks/amy_cuddy_your_body_language_may_shape_who_you_are at 4:53- 4:57

Note. This image presents a running race. The athlete is rushing through the finish line with two arms high up in the air. The transcript reads, “So the arms up in the V, the chin is slightly lifted.” In the public domain. (Cuddy, 2012).

Figure 5

Access the Screenshot When the Speaker Performed an Expanding Action through:

https://www.ted.com/talks/amy_cuddy_your_body_language_may_shape_who_you_are at 5:12-5:13

Note. The screenshot presents the presenter and the screen in the background. The presenter performs a gesture of high power: expanding the two arms in a horizontal position with her shoulder. The transcript reads, “So if someone is being really powerful with us.” In the public domain. (Cuddy, 2012).

At times, the three modes converged to clarify the message in the video. For example, in Figure 7, the image displays gestures of powerless status where participants are seated with lowered heads and shrinking bodies. Meanwhile, the presenter was acting out a similar gesture, hunching up, with the verbal description, “You’re hunching up, making yourself small.” At this moment, the meanings represented in the three modes were in convergence, and the audience could process the information through different channels synchronically. The meanings were parallel and worked together as a whole, reinforcing the ideas. In this way, the audience could hear, visualize, and experience the statement through the repeated representation in the three modes, thus facilitating the lay audience to understand the ideas.

Figure 6

Access the Screenshot When the Speaker Hunched Shoulders and Crossed Arms through:

https://www.ted.com/talks/amy_cuddy_your_body_language_may_shape_who_you_are at 5:14-5:15

Note. This screenshot shows the presenter and the screen at the background. She performs a shrinking gesture with two arms wrapping herself. The transcript reads, “We tend to make ourselves smaller.” In the public domain. (Cuddy, 2012).

The three modes can also supplement one another in meaning. Figure 8 shows such an interplay. At this moment, the speaker was describing a job interview and stating that “for numerous reasons, no, no, no, don’t do that,” with two hands beating downward to emphasize the “do that.” While the speech and gestures were semantically parallel in their expression of the meaning in the action process, it was still unclear what “that” indicated. However, the image on the screen in the background supplemented the other modes and illustrated the meaning of “that” (hands high up in a V shape and feet on the desk). Thus, in this kind of interplay, the image has supplemented the meaning constructed by spoken and gestural instances and, consequently, the presenter achieved clarity and disambiguation of the meaning. Converging the three modes in the way of supplementation clarifies the meaning represented, eases the burden of comprehension, and improves the accessibility of this talk.

Figure 7

Access the Screenshot of an Orchestration of the Three Modes in Convergence (Concurrence) through:

https://www.ted.com/talks/amy_cuddy_your_body_language_may_shape_who_you_are at 13:46

Note. The screenshot shows the whole platform with a slide in the background. The presenter is standing with her two arms wrapping herself. The image in the background presents three people who

are sitting with their heads down. The orchestration happens at 13:46. The transcript reads, “You’re hunching up, making yourself small.” In the public domain. (Cuddy, 2012).

Figure 8

Access the Screenshot of Orchestration of the Three Modes in Convergence (Supplementation) through: https://www.ted.com/talks/amy_cuddy_your_body_language_may_shape_who_you_are at 13:28

Note. The screenshot presents the speaker and the image in the background. The speaker is moving her hands in a upward and downward in a beating down motion. The image in the background shows an interviewee with her two arms high and legs on the desk. The orchestration happens at 13:28. The transcript reads, “For numerous reasons, no, no, no, don’t do that.” In the public domain. (Cuddy, 2012).

According to Valeiras-Jurado et al. (2018), the more easily a message is understood, the more accessible and persuasive it comes to be. Only when an idea is clearly conveyed and easily understood can it be accessed, accepted, and disseminated. Hence, it is of utmost importance to make ideas clear to persuade the audience. Moreover, it may be difficult for lay audiences to comprehend sophisticated scientific concepts, terminologies, rationales, and statements in a specific field. Thus, making reference to specific objects, actions, or concrete examples, represented by the three modes coherently, could help an audience visualize abstract ideas and enhance their comprehension.

Use of Speech, Images, and Gestures to Engage an Audience

The presenter appeared to make certain choices of image and gesture in order to appeal to the audience. The interpersonal meanings made in the images selected for this talk helped engage the audience. According to Kress and Van Leeuwen (2006), a medium close-up shot and eye-level frontal angle can enact a sense of intimacy with an audience (Figures 2, 3, and 4). The participants’ actions of spreading out or shrinking represented in these images are relatable to the daily experiences of the audience and thus created familiarity and enhanced engagement.

Gestures also enhance audience involvement. Indexical actions that occur with speech would necessarily require speech for interpretation (Lim, 2019). Although less frequently used, the indexical action of pointing is observed in this talk (Figures 9 and 10). In the process of asking the question, “So what is your body language communicating to me?”, the presenter pointed to the audience when she said “you” and to herself when she said “me.” The action of pointing with the personal pronouns heightened the sense of invitation and participation, thus building a sense of involvement. This helps bring about stronger audience engagement.

It was also found that the three modes were used in deviation or contradiction to one another in order to involve the audience. Figure 11 illustrates this orchestration. Between 13:17 and 14:40, the presenter described an experiment of an interview, aiming to test whether the practice of high-power gestures could change people’s lives in the real world. The screenshot is from 13:29, when the presenter described an imagined scenario that contradicts people’s conventional expectations and showed the image of an interviewee expanding arms and legs. Meanwhile, she joked, “Okay, so this is what you do when you go in for the job interview, right?” In this orchestration, the participant’s gesture displayed in the image complemented what “this” indicates.

Figure 9

Access the Screenshot of Indexical Action Converged with Speech #1 through:

https://www.ted.com/talks/amy_cuddy_your_body_language_may_shape_who_you_are at 2:00

Note. The screenshot shows the presenter pointing downward with her index finger. The orchestration happens at 2:00. The transcript reads, “So what is your body language...” In the public domain. (Cuddy, 2012).

Figure 10

Access the Screenshot of Indexical Action Converged with Speech #2 through:

https://www.ted.com/talks/amy_cuddy_your_body_language_may_shape_who_you_are at 2:02

Note. The screenshot shows the presenter pointing herself with two hands. The orchestration happens at 2:02. The transcript reads, “...communicating to me?” In the public domain. (Cuddy, 2012).

However, the gesture was of a state process and did not carry ideational meaning, and thus was disparate from the meanings represented in the images and speech. The three modes did not represent meanings congruently. In other words, they were in meaning divergence. Although the presenter’s gesture could not add ideational meaning to what the speech and the images represented, the relaxed state of hand gesture, combined with the ideational meaning that contradicted people’s expectation created a relaxed atmosphere by using humour to keep the audience involved.

Figure 11

Access the Screenshot of Orchestration of the Three Modes in Divergence through:

https://www.ted.com/talks/amy_cuddy_your_body_language_may_shape_who_you_are at 13:29

Note. In this screenshot, the presenter is standing with two hands in front of her body. The image shown in the background is an interviewee with hands and legs fully expanding. The orchestration happens at

13:29. The transcript reads, “Okay, so this is what you do when you go in for the job interview, right?” In the public domain. (Cuddy, 2012).

In an online academic genre, it is critical to maintain engagement between presenter and audience. The interpersonal meanings expressed by the medium close-up shot of the stage with a frontal eye-level angle helped to enact a sense of familiarity and involvement. The indexical actions with corresponding spoken utterances also created a feeling of dialogue and participation. In addition, the instances where the three modes represented incongruent meanings in divergence created a relaxed atmosphere and humour. Thus, the image and gesture instances as well as the artful orchestration of meaning divergence contributed to enacting a sense of participation and enhancing audience engagement.

Discussion

Analysis of the TED Talk reveals the ways in which the speaker used speech, images, and gestures strategically to disseminate scientific knowledge in accessible and engaging ways by clarifying ideas and engaging the audience. Many linguistic choices, for example, material process and declarative, were used to represent scientific knowledge in a direct manner. Converging with speech, images, and gestures worked in combination to explain abstract concepts, reinforce key statements, and form a meaning whole. This could have eased the burden of cognitive processing and enhanced the audience’s understanding. Meanwhile, a medium close-up shot from frontal and eye-level angles was deployed to enact a sense of involvement. Indexical actions combined with pronouns in speech created a sense of belonging and facilitated the presenter’s engagement with the audience. The three modes also worked in divergence to express humour which kept the audience involved.

We posit that the findings from this analysis can inform teachers’ design of online learning to address the currently reported challenges, such as difficulties in students’ understanding of online instruction and materials (Allo, 2020). According to Mayer (2002), people’s processing capacity is limited in the working memory, and only appropriate cognitive processing can bring about learning. As such, a teacher’s intentional use of semiotic modes can clarify meaning in online teaching and reduce unnecessary cognitive load on students and facilitate learning. For example, a teacher can use gestures and images to vivify the description made in speech. Gestures and images can also be used to illustrate abstract concepts expressed in teachers’ speech through meaning convergence. As observed in the analysis of the TED Talk, orchestration across semiotic modes can ease the burden on cognitive processing and enhance the intake of information and the understanding of content. When a teacher uses the three semiotic modes to supplement one another, the interplay can clarify meaning, ease the

burden of comprehension, and improve the accessibility of ideas being disseminated (Valeiras-Jurado et al., 2018). This, we argue, is superior to the sole use of language, whether a teacher's speech or words on a slideshow presentation, to disseminate knowledge in online teaching.

Another challenge related to online learning is the need to keep the students engaged (Oyedotun, 2020). We argue that the artful orchestration of the semiotic modes of speech, images, and gestures can serve to better engage students during online instruction. When the modes are used to express meaning divergence, where each mode seems to express contradictory meaning, students' interest could be piqued and attention drawn. For example, a teacher could use speech and gestures to express seemingly oppositional meanings to create humour. Such playful orchestration of semiotic modes to express meaning divergence contributes to a structured informality which builds solidarity and rapport between the teacher and students, creates a convivial learning environment, and encourages student engagement (Lim, 2021).

With greater semiotic awareness (Towndrow et al., 2013), teachers can design online learning experiences that disseminate knowledge in an accessible and engaging manner. Teachers, as designers of learning experiences, can apply semiotic awareness for orchestration of their speech, images, and gestures in embodied teaching (Lim, 2020). A more intentional use of semiotic modes in online instruction can mitigate the challenges of not being able to meet and interact with students in the physical classroom (Tanis, 2020). Online learning cannot be a substitute for face-to-face. Online learning offers new ways of learning and when designed well, it can be superior (Kalantzis & Cope, 2020). In this paper, we argue that the design of online learning begins with a more intentional orchestration of the teacher's speech, gestures, and images to disseminate knowledge in accessible and engaging ways.

Conclusion

This study has discussed how an understanding of the use of speech, images, and gestures to disseminate scientific ideas in TED Talks can inform the design of online instruction. Using one of the most viewed videos as an illustrative example, the fine-grained analysis has described the interplay across the three modes, revealed the ways of strategic orchestration, and demonstrated how they functioned to clarify the ideas and achieve audience engagement. TED Talks and online instruction have in common a communicative purpose of knowledge dissemination. They also share the online medium and seek to engage non-expert audiences. As such, an understanding of how the presenter orchestrates speech, images, and gestures in TED Talks can also illuminate the ways knowledge can be delivered in online instruction. Based on understanding how the three semiotic modes are used to

disseminate scientific knowledge in accessible and engaging ways, our paper argues that insights into multimodal orchestration can function as a heuristic tool to inform the design of knowledge dissemination in online learning.

While the present analysis is based on one popular TED Talk, we seek to demonstrate how any similar fine-grained analysis can reveal the strategic use of multimodal orchestration to fulfill the communicative purposes of TED Talks and to draw illustrative applications for online lectures. Further study on a larger database of TED Talks may be pursued. Future research could also be conducted on TED Talk video production, considering choices such as camera angles and post-production editing. As our paper focuses on online learning, we have not discussed how the ways of multimodal orchestration in TED talks could also inform the design of face-to-face learning, which could also be a subject for future research.

Acknowledgement

This study was supported by Nanyang Technological University Research Scholarship. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the university.

References

- Adedoyin, O. B., & Soykan, E. (2020). Covid-19 pandemic and online learning: The challenges and opportunities. *Interactive Learning Environments*. Advance online publication. 1-13.
<https://doi.org/10.1080/10494820.2020.1813180>
- Ali, W. (2020). Online and remote learning in higher education institutes: A necessity in light of COVID-19 pandemic. *Higher Education Studies*, 10(3), 16-25.
<https://doi.org/10.5539/hes.v10n3p16>
- Alibali, M. W., Nathan, M. J., Wolfgram, M. S., Church, R. B., Jacobs, S. A., Johnson Martinez, C., & Knuth, E. J. (2014). How teachers link ideas in mathematics instruction using speech and gesture: A corpus analysis. *Cognition and Instruction*, 32(1), 65-100.
<http://dx.doi.org/10.1080/07370008.2013.858161>
- Allo, M. D. G. (2020). Is the online learning good in the midst of Covid-19 pandemic? The case of EFL learners. *Jurnal Sinestesia*, 10(1), 1-10. <https://sinestesia.pustaka.my.id/journal/article/view/24>
- Ally, M. (2008). Foundations of educational theory for online learning. In T. Anderson & F. Elloumi (Eds.), *Theory and practice of online learning* (pp. 3-31). Athabasca University Press.
- Bannert, M. (2002). Managing cognitive load—Recent trends in cognitive load theory. *Learning and Instruction*, 12(1), 139-146. [https://psycnet.apa.org/doi/10.1016/S0959-4752\(01\)00021-4](https://psycnet.apa.org/doi/10.1016/S0959-4752(01)00021-4)
- Bucchi, M., & Saracino, B. (2016). “Visual Science Literacy” images and public understanding of science in the digital age. *Science Communication*, 38(6), 812-819.
<https://doi.org/10.1177/1075547016677833>
- 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.
<http://dx.doi.org/10.6330/ETL.2015.39.4.02>
- Chue, S., Lee, Y.-J., & Tan, K. C. D. (2015). Iconic gestures as undervalued representations during science teaching. *Cogent Education*, 2(1), Article 1021554.
<https://doi.org/10.1080/2331186X.2015.1021554>
- Cuddy, A. (2012, June). Your body language may shape who you are [Video]. TED Conferences.
https://www.ted.com/talks/amy_cuddy_your_body_language_may_shape_who_you_are
- D’avanzo, S. (2015). Speaker identity vs. speaker diversity: The case of TED Talks corpus. In G. Balirano & M. C. Nisco (Eds.), *Linguaging diversity: Identities, genres, discourses* (pp. 279-296). Cambridge Scholars Publishing.
- ELAN (Version 5.4)[Computer software]. (2018). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. <https://archive.mpi.nl/tla/elan/>

- Feng, D., & O'Halloran, K. L. (2013). The visual representation of metaphor: A social semiotic approach. *Review of Cognitive Linguistics. Published under the auspices of the Spanish Cognitive Linguistics Association, 11(2)*, 320-335. <https://doi.org/10.1075/rcl.11.2.07fen>
- Halliday, M. A. K. (1994). *An introduction to functional grammar* (2nd ed.). Edward Arnold.
- Hamid, R., SENTRYO, I., & HASAN, S. (2020). Online learning and its problems in the Covid-19 emergency period. *Jurnal Prima Edukasia, 8(1)*, 86-95. <https://doi.org/10.21831/jpe.v8i1.32165>
- Harrison, C. (2003). Visual social semiotics: Understanding how still images make meaning. *Technical Communication, 50(1)*, 46-60.
- 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, Article 101002*. <https://doi.org/10.1016/j.jeap.2021.101002>
- Jewitt, C., Bezemer, J., & O'Halloran, K. (2016). *Introducing multimodality*. Routledge.
- Kalantzis, M., & Cope, B. (2020). After the COVID-19 crisis: Why higher education may (and perhaps should) never be the same. *Access: Contemporary Issues in Education, 40(1)*, 51-55. <https://doi.org/10.46786/ac20.9496>
- Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). Routledge.
- Li, Y., Gao, Y., & Zhang, D. (2016). To speak like a TED speaker—A case study of TED motivated English public speaking study in EFL teaching. *Higher Education Studies, 6(1)*, 53-59. <https://doi.org/10.5539/HES.V6N1P53>
- Lim, F. V. (2019). Analysing the teachers' use of gestures in the classroom: A systemic functional multimodal discourse analysis approach. *Social Semiotics, 29(1)*, 83-111. <https://doi.org/10.1080%2F10350330.2017.1412168>
- Lim, F. V. (2020). *Designing learning with embodied teaching: Perspectives from multimodality*. Routledge.
- Lim, F. V. (2021). 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/10.1177/1470357218820695>
- Lorés, R. (2020). Science on the web: The exploration of European research websites of energy-related projects as digital genres for the promotion of values. *Discourse, Context & Media, 35*, Article 100389. <https://doi.org/10.1016/j.dcm.2020.100389>
- Maqableh, M., & Alia, M. (2021). Evaluation online learning of undergraduate students under lockdown amidst COVID-19 Pandemic: The online learning experience and students' satisfaction. *Children and Youth Services Review, 128*, Article 106160. <https://doi.org/10.1016/j.childyouth.2021.106160>

- Martin, B. A. (2021). Teachers perceptions of Google Classroom: Revealing urgency for teacher professional learning. *Canadian Journal of Learning and Technology*, 47(1).
<https://doi.org/10.21432/cjlt27873>
- Martinec, R. (2000). Construction of identity in Michael Jackson's Jam. *Social Semiotics*, 10(3), 313-329. <https://doi.org/10.1080/10350330050136370>
- Masi, S. (2016). Gestures in motion in TED Talks: Towards multimodal literacy. In V. Bonsignori & B. C. Camiciottoli (Eds.), *Multimodality Across Communicative Settings, Discourse Domains and Genres* (pp. 146-165). Cambridge Scholars Publishing.
- Masi, S. (2020). Exploring meaning-making practices via co-speech gestures in TED Talks. *Journal of Visual Literacy*, 39(3-4), 201-219. <https://doi.org/10.1080/1051144X.2020.1826223>
- Mattiello, E. (2019). A corpus-based analysis of scientific TED Talks: Explaining cancer-related topics to non-experts. *Discourse, Context & Media*, 28, 60-68.
<http://dx.doi.org/10.1016/j.dcm.2018.09.004>
- Mayer, R. E. (2002). Multimedia learning. *Psychology of Learning and Motivation*, 41, 85-139.
[https://doi.org/10.1016/S0079-7421\(02\)80005-6](https://doi.org/10.1016/S0079-7421(02)80005-6)
- Mayer, R. E. (2019). Thirty years of research on online learning. *Applied Cognitive Psychology*, 33(2), 152-159. <https://doi.org/10.1002/acp.3482>
- Mayer, R. E., & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction*, 12(1), 107-119. [http://dx.doi.org/10.1016/S0959-4752\(01\)00018-4](http://dx.doi.org/10.1016/S0959-4752(01)00018-4)
- Miller, T., MacLaren, K., & Xu, H. (2020). Online learning: Practices, perceptions, and technology. *Canadian Journal of Learning and Technology/La revue canadienne de l'apprentissage et de la technologie*, 46(1). <https://doi.org/10.21432/cjlt27894>
- Niess, M. L. (2012). Teacher knowledge for teaching with technology: A TPACK lens. In R. N. Ronau, C. R. Rakes, & L. N. Margaret (Eds.), *Educational technology, teacher knowledge, and classroom impact: A research handbook on frameworks and approaches* (pp. 1-15). IGI Global.
<https://doi.org/10.4018/978-1-60960-750-0.ch001>
- Ogbonna, C. G., Ibezim, N. E., & Obi, C. A. (2019). Synchronous versus asynchronous e-learning in teaching word processing: An experimental approach. *South African Journal of Education*, 39(2), 1-15. <https://doi.org/10.15700/saje.v39n2a1383>
- O'Halloran, K. L. (2008). Systemic functional-multimodal discourse analysis (SF-MDA): Constructing ideational meaning using language and visual imagery. *Visual Communication*, 7(4), 443-475.
<https://doi.org/10.1177/1470357208096210>
- Oyedotun, T. D. (2020). Sudden change of pedagogy in education driven by COVID-19: Perspectives and evaluation from a developing country. *Research in Globalization*, 2, Article 100029.
<https://doi.org/10.1016/j.resglo.2020.100029>

- Pei, L., & Wu, H. (2019). Does online learning work better than offline learning in undergraduate medical education? A systematic review and meta-analysis. *Medical Education Online*, 24(1), Article 1666538. <https://doi.org/10.1080/10872981.2019.1666538>
- 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.), *Corpus-informed research and learning in ESP* (pp. 135-166). John Benjamins.
- Salem, A. A. (2019). A sage on a stage, to express and impress: TED Talks for improving oral presentation skills, vocabulary retention and its impact on reducing speaking anxiety in ESP settings. *English Language Teaching*, 12(6), 146-160. <https://doi.org/10.5539/ELT.V12N6P146>
- Shalev-Shwartz, S. (2011). Online learning and online convex optimization. *Foundations and Trends in Machine Learning*, 4(2), 107-194. <https://doi.org/10.1561/22000000018>
- Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.
- Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19-30). Cambridge University Press. <https://doi.org/10.1017/CBO9780511816819.003>
- Tanis, C. J. (2020). The seven principles of online learning: Feedback from faculty and alumni on its importance for teaching and learning. *Research in Learning Technology*, 28, Article 2319. <https://doi.org/10.25304/rlt.v28.2319>
- Tardy, C. M. (2005). Expressions of disciplinarity and individuality in a multimodal genre. *Computers and Composition*, 22(3), 319-336. <https://doi.org/10.1016/j.compcom.2005.05.004>
- Teo, C. L., Tan, S. C., & Chan, C. K. (2021). Pedagogical transformation and teacher learning for knowledge building: Turning COVID-19 challenges into opportunities. *Canadian Journal of Learning and Technology*, 47(4). <https://doi.org/10.21432/cjlt28057>
- Towndrow, P. A., Nelson, M. E., & Yusuf, W. F. B. M. (2013). Squaring literacy assessment with multimodal design: An analytic case for semiotic awareness. *Journal of Literacy Research*, 45(4), 327-355. <https://doi.org/10.1177/1086296X13504155>
- Valeiras-Jurado, J., Ruiz-Madrid, M. N., & Jacobs, G. (2018). Revisiting persuasion in oral academic and professional genres: Towards a methodological framework for multimodal discourse analysis of research dissemination talks. *Ibérica: Revista de la Asociación Europea de Lenguas para Fines Específicos (AELFE)*, 35, 93-118.
- Wingrove, P. (2017). How suitable are TED Talks for academic listening? *Journal of English for Academic Purposes*, 30, 79-95. <https://doi.org/10.1016/j.jeap.2017.10.010>

Zhang, Y., Machin, D., & Song, T. (2015). Visual forms of address in social media discourse: The case of a science communication website. *Journal of Multicultural Discourses*, 10(2), 236-252.
<http://dx.doi.org/10.1080/17447143.2015.1042384>

Authors

Jingxin Jiang is an Associate Professor at Chongqing Open University, China, and Ph.D. candidate at the National Institute of Education, Nanyang Technological University, Singapore. Her research interest is multimodal discourse analysis, multiliteracy, and online learning. *Email:* jingxinjiang@hotmail.com

Fei Victor Lim is an Assistant Professor at the National Institute of Education, Nanyang Technological University, Singapore. He researches and teaches on multiliteracies, multimodal discourse analyses, and digital learning. *Email:* victor.lim@nie.edu.sg



© 2022 Jingxin Jiang, Fei Victor Lim

This work is licensed under a Creative Commons Attribution-NonCommercial CC-BY-NC 4.0 International license.