32
INTERPERSONAL MEANING AND AUDIENCE ENGAGEMENT IN ACADEMIC PRESENTATIONS

A multimodal discourse analysis perspective

Gail Forey and Dezheng Feng

Introduction

‘Start with a quote’, ‘start with a hook’, ‘capture the attention of your audience in the first few seconds’ and many other ‘tips’ can be found about how to give a good presentation, and there are a plethora of websites, YouTube videos and textbooks offering tips for better presentations (www.TED.com; Duarte, 2010; Gallo, 2014). Academic presentations (APs) are a key mode of sharing information both in the workplace (Evans, 2013; Kline, 2003) and academia (Ventola & Charles, 2002; Hood & Forey, 2005; Thompson, 2003). Students’ presentations are often a form of assessment and studies related to student presentations have discussed group and individual performance (Chou, 2011), the projection of identity by the speaker (Zareva, 2013) and L1 versus L2 perceptions of APs (Zareva, 2009). At a postgraduate level, studies provide insights into such features as dissertation defenses (Mežek and Swales, this volume). For example, both Recski (2005) and Wulff et al. (2009) examine the discussion section (DS) of the dissertation defense and demonstrate the high stakes nature of APs. APs are high stakes for both students and faculty; that is, students’ presentations are linked to assessment and grades, and faculty presentations are informally assessed for their contribution to the field by their peers. For faculty members, APs are an integral feature of their annual workload, and these presentations may range from an invited plenary, informal seminar, parallel paper presentation to a large or a small audience. Due to space, we focus on APs given by academics rather than students, and suggest that the features identified in faculty APs will have currency and value for students and other public speakers.

Within the field of EAP an overwhelming number of studies favor written rather than spoken EAP (Hood & Forey, 2005). It could be argued that some of the written modes are ‘spoken-like’, that perhaps the conference AP is often ‘written-to-be-spoken’, and that
some of the less stimulating APs sound extremely written-like (Halliday, 1985). As pointed out by Thompson (1997, 2003), one of the significant differences between a research article (written mode) and an AP (spoken mode) is the foregrounding of interpersonal and interactive meanings in the latter. It is the interpersonal element of APs that is one of the most striking features, and is the focus of the present chapter. Studies on APs include a wide range of features and, more recently, have tended to focus on multimodal interpersonal meaning. We present an overview of these current studies, and then move on to examine the multi-semiotic resources chosen by the speaker to develop the tenor (i.e. the relationship the speaker has with the audience) in an extract from an AP. Finally, we draw together the discussion of interpersonal meaning and suggest possible future directions for research related to APs.

**Studies of academic presentations: linguistic and multimodal analyses**

The dialogic aspects of APs, that is of having an audience to speak to, influence the choices the speaker makes from the initial decision to attend the conference and the drafting of the abstract proposal to the last word spoken. The academic, when presenting, aims to highlight chosen features, either current research, frameworks or other elements that they believe will be of interest to the audience. The speaker mediates the intended meaning through the ‘semiotic spanning’ between the linguistic, visual and audio choices (Ventola and Charles, 2002). Semiotic spanning ‘involves looking at how we link with other sources, with other experiences and their expressions, and still keep what we say as relevant and focused in terms of the presented paper and the point of discussion’ (Ventola 1999, p. 118). Semiotic spanning helps us understand the link between multiple semiotic resources found at a conference and beyond the conference; for example, how one speaker relates to the previous speaker, and perhaps how a presentation is referred to and discussed in the coffee break and at a later point. Ventola (1999) and Ventola and Charles (2002) point out that the discourse at a conference expands beyond what the speaker says to the complex world of discourses. The presentation lives on through formal and informal texts, quotations in other people’s texts, social exchanges such as around the dinner table, newspaper reports or written papers, etc.

The term ‘linguistic features’ refers to what is said during the presentation, and this is often the starting point for research focusing on APs. Some studies focus on general linguistic features such as Bhattacharyya (2014) who discusses communicative competence in engineering presentations. Other studies focus on specific language features such as Fernández-Polo (2014) who outlines the role of I mean by English as a lingua franca (ELF) speakers in their presentations. Using a corpus analysis approach, Fernández-Polo (2014) demonstrates the various use of I mean, such as self-repair, corrective or preemptive purposes, introducing background, foregrounding interpersonal solidarity with the audience, along with reinforcing and highlighting the ideas presented. Another approach is to focus on sections of the talk as in Hood and Forey (2005) who outline interpersonal meaning in introductions, and Querol-Julían and Fortanet-Gómez (2012), who analyse the discussion section of presentations. Querol-Julían and Fortanet-Gómez (2012) point out that there is a complex shift from the monologic elements of the main part of the presentation to the discussion (Q&A) section of a presentation. Both Hood and Forey (2005) and Querol-Julían and Fortanet-Gómez (2012) draw on a systemic functional linguistic (SFL) framework (Hood, this volume) and adopt Martin and White’s (2005) appraisal system to investigate the complementarity of the linguistic and visual choices that convey the relationship between the speaker and the audience with respect to how ‘solidarity’ is or is not achieved. Both
highlight the value of studying interpersonal meaning in the introductions or the discussion sections of APs. Both focus on language and highlight the importance of the relationship between the linguistic interpersonal meaning and gesture (Austin & Weller, 2014). However, a more comprehensive examination of the relationship between the linguistic and gesture is needed.

In order to understand the multi-semiotic resources we need to capture the multitude of resources available to a speaker when presenting. A number of studies have adopted the multimodal discourse analysis (MDA) approach (see Kress & van Leeuwen, 2006) which was developed from systemic functional linguistics (Halliday & Matthiessen, 2014). MDA scholars have explored an increasing range of domains: for example, visual image (Bateman, 2008; Kress & van Leeuwen, 2006; O’Toole, 1994), three-dimensional objects (Martin & Stenglin, 2007; O’Toole, 1994), websites (Djono, 2005) and film (Bateman & Schmidt, 2012; Feng & O’Halloran, 2013). In addition, classroom teaching has captured increasing attention from multimodal theorists, which is studied within the emerging field of ‘multiliteracies’ (New London Group, 1996; Cope & Kalantzis, 2000; Kress et al., 2001; Lim et al., 2012; O’Halloran et al., this volume). APs have also started to attract systematic multimodal analysis. The relationship between the visual (the slides) used in an AP and the linguistic (what the speaker says) has been discussed in a number of studies (Cassily & Ventola, 2002; Rowley-Jolivet, 2002, 2004, Rowley-Jolivet & Carter-Thomas, 2005). Bernardis and Gentilucci (2006) focus on the symbiotic nature of spoken word and gesture. Rowley-Jolivet (2002, 2004) examines the visual resources found in science and in particular geology, medicine and physics presentations. In an earlier paper, Rowley-Jolivet (2002) highlights the role of image as a key resource for making meaning in science, and the disciplinary differences between these three scientific subject areas (Rowley-Jolivet, 2004). She highlights the benefits of shared visual knowledge within a discipline, especially at international conferences where English is not the first language for many (ibid. 2002). Using the multimodal annotation tool ELAN, a free software developed by the Max Plank Institute (https://tla.mpi.nl/tools/tla-tools/elan), Rowley-Jolivet and Carter-Thomas (2005) demonstrate the importance of the relationship between the visual, paralinguistic and kinetic resources that are used to co-construct evaluative meaning in the discussion section of science presentations. Rowley-Jolivet (2004) also demonstrates the value of applying software to support the analysis of multimodal texts. Rowley-Jolivet (2004) uses ELAN to enable a detailed analysis of multiple complex choices. In this study, we adopt an alternative software tool, the Multimodal Lab Software (Smith et al., 2014), to analyze the data. The value of utilizing software ensures archival resources and ultimately can lead to a comparable corpus of a complex meaning-making in APs.

The final semiotic resource afforded in conference presentation is the audio; for example, the acoustics, the voice quality and other resources related to sound. However, studies related to the audio resources in APs are limited. These three elements – linguistic, visual and audio – co-construct the meanings made, and are only isolated for discussion purposes to deconstruct the text and better understand the choices made by the speaker. Although these are all resources found in APs, a discussion of the full range of linguistic, visual and audio semiotic resources will not be possible.

In this chapter we focus on one of the most salient features that distinguishes the conference presentation from other EAP modes: the dialogic nature of conference presentations and the issue of tenor. We illustrate, with reference to data, the semiotic spanning of the linguistic, visual and audio which together co-construct to engage the audience. Adopting this semiotic discursive perspective allows us to use the social semiotic
tools that are so well developed for systematically analyzing meaning construction in linguistic and multimodal discourse. From the perspective of multimodal studies, for empirical research of APs, we need a framework that is explicit enough to support reliable description and fine-grained enough for the description to be informative. In what follows, we outline a semiotic framework to elucidate the array of multimodal choices that are available for a speaker to engage with his/her audience.

**Engagement in academic presentations: an integrated framework**

The combination of discourse analysis, in particular MDA, and appraisal analysis is useful to deconstruct the complex reality of APs. Interpersonal meaning is a key resource in the introduction (Hood & Forey, 2005) and discussion section (Quel-Julian & Fortanet-Gomez, 2012) of a presentation, and interpersonal meaning is particularly important when the speaker is convincing the audience of the position in the body of the AP. Engagement is a key resource the speaker adopts to develop a positive relationship with the audience throughout the talk in order to strengthen his/her argument. Martin and White’s (2005) appraisal analysis, and particularly the engagement system, provides a framework to understand the space that is opened up or closed down within the text by the speaker in an AP. For example, in what has perhaps been incorrectly referred to as the ‘monologic’ section of the talk – the body – in this section, along with all other stages of the AP, the speaker continues to make choices that involve and persuade the audience. Thus, even though there is only one speaker when choosing what to say, the speaker engages with the audience to create solidarity through speech functions chosen (e.g. using interrogatives or imperatives), involvement through gaze, body direction, gesture or movement, and the proximic social distance that is selected: is the speaker at a close, median or distant position in relation to the audience.

For the systematic description of the discourse of an AP and to unpack the dynamic relationship between speaker and audience, it is useful to adopt the social semiotic notions of stratum and system, which have been developed for language in SFL (Halliday, 1978; Halliday & Matthiessen, 2014; see Hood, this volume, for an overview), and are extended to nonlinguistic semiosis (e.g. Kress & van Leeuwen, 2006; Martinec, 2001; Bateman & Schmidt, 2012). The APs which we are referring to are closely related to the research article and provide an opportunity for the speaker to showcase their current research and argue the feasibility of their work (Thompson, 1997, 2003). Therefore, an AP can be seen as a social activity with a particular purpose and structure (Martin, 1992; Hood & Forey, 2005). This genre is realized through the configuration of field, tenor and mode, which incorporate what the speaker presents, how he/she engages audiences’ attention and interest, and what is the mode of communication respectively. Viewer engagement is in turn realized through various parameters of interpersonal semantics; for example, speech functions (Halliday & Matthiessen, 2014) and appraisal resources in language (Martin & White, 2005), and interactive meaning in visual images (Kress & van Leeuwen, 2006). These aspects of meaning are then realized by multi-semiotic resources of language, gesture, gaze, movement, and so on. The notion of strata, which is illustrated in Table 32.1, allows us to elucidate how AP discourse is semiotically constructed and how the social purpose of AP discourse is realized. To describe exactly how interpersonal meaning is constructed, an integrated framework of the dimensions of interpersonal semantics that contribute to the construction of viewer engagement is provided in Figure 32.1.

The ‘systemic’ principle regards grammar as systems of paradigmatic choices, and meaning is created through making and combining choices from the systems, which are
Table 32.1 Strata in the analysis of viewer engagement in AP discourse

<table>
<thead>
<tr>
<th>Strata</th>
<th>Realization</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Genre</strong></td>
<td>Argumentation</td>
</tr>
<tr>
<td><strong>Register</strong></td>
<td>Tenor – viewer engagement</td>
</tr>
<tr>
<td></td>
<td>Mode – multimodal</td>
</tr>
<tr>
<td></td>
<td>Field – object of study</td>
</tr>
<tr>
<td>Discourse semantics</td>
<td>Prosodic features of interpersonal meaning</td>
</tr>
<tr>
<td>Lexicogrammar</td>
<td>Multi-semiotic</td>
</tr>
<tr>
<td></td>
<td>• linguistic resource</td>
</tr>
<tr>
<td></td>
<td>• visual resources</td>
</tr>
<tr>
<td></td>
<td>• acoustic resources</td>
</tr>
</tbody>
</table>

Figure 32.1 System network: choices of engagement identified in academic presentations
Interpersonal meaning and audience engagement in academic presentations

represented as system networks. Extending the notion of choice to the multimodal discourse of APs, we are able to classify the exact choices made by a speaker. With these two semiotic tools, we propose the system network as shown in Figure 32.1, which models the choice of multi-semiotic resources for the construction and negotiation of meaning in AP discourse. Synthesizing the work of Halliday and Matthiessen (2014), Kress and van Leeuwen (2006) and Martin and White (2005), our framework includes four parameters of interpersonal semantics, namely, speech function, involvement, social distance and affect, which are explained in detail in the following.

The speech function framework is based on Halliday and Matthiessen (2013) for language and Martinec (2001) for gesture. Halliday and Matthiessen (2013) categorize speech functions into four types: demanding information, demanding goods/services, offering information and offering goods/services. In language, these four speech functions are generally realized respectively by the mood types of interrogatives, imperatives, indicatives and any of these for the last category. As language is a fundamental resource for establishing speaker–audience relation, the analysis of mood types can offer insights into how a speaker negotiates meaning for optimal engagement. Another important resource for exchanging meaning is gesture, as shown in Figure 32.1. Martinec (2001) argues that gestures can perform these four functions as well and terms it ‘sign function’. The questioning function is mainly realized by an upward movement of the head and/or eyebrows; the commanding function can be realized through emblematic gestures (e.g. gestures of stopping, summoning, etc.) and pointing gestures; the offering function is typically realized by stretching an arm and a hand (offering an actual object or virtual one); the statement function is realized through the representation of things or concepts (e.g. the iconic gestures in McNeill (1992), for example, hands moving up to describe other upward movements). We will also term these functions ‘sign functions’ to refer to both the speech and the gestures in exchanging information. This framework affords the description of how a speaker uses language and gesture to negotiate meaning with the audience.

The involvement function is mainly realized by nonverbal resources, including gaze, body direction, posture, the use of a pointer (as an extension of body gesture) and movement (see Figure 32.1). In terms of gaze, we distinguish three targets in APs: the audience, the screen and other objects. When there is eye contact between the speaker and the audience, it is possible for the speaker to communicate affective or attitudinal information to engage the audience, and at the same time, the speaker can notice affective or attitudinal reactions from the audience. When the speaker looks at the screen, he/she directs the audience’s attention to it and can maintain a satisfactory degree of interaction; when the speaker looks at his/her own computer or notes, speaker–audience interaction is minimized. By calculating the time distribution of different speakers’ gaze, we can tell the degree of his/her interaction with the audience. Body direction, as one of the variables of involvement, can be categorized into different positions/angles. If we take the screen as the point between 0° and 180°, then a frontal angle is facing the audience at 90°; back is facing the screen at 90°; side is facing to the audience at 0° or 180°; frontal oblique, which is halfway between frontal and side at 45° with two categories of frontal oblique left and frontal oblique right; and back oblique left and back oblique right, which is what is in between the back and side. The frontal angle of the presenter’s body direction embodies an elicitation for the presenter, inviting the audience’s attention. The third resource is the pointer, which can be used to point at a single word or a group of words (up to a whole paragraph) through moving the pointer. The last category of involvement resource is movement, which includes the directions of left/right and front/back from the audience’s perspective. As human beings tend to focus more attention on moving objects, the speaker’s movement can attract
the audience’s attention. Movement also allows the speaker to adjust his/her distance with the audience at different locations of the room.

The parameter of social distance is mainly realized through the physical distance between the speaker and the audience. Studies in proxemics suggest that closer physical distance impacts the closeness of the social or affective distance (e.g. Hall, 1966). Based on these findings, Lim (2010) proposes a framework of space in his analysis of classrooms, and the distinction between personal space, authoritative space and interactional space is adopted here. The speaker stays within her personal space if he/she stands/sits behind the computer (normally in the corner) with half of his/her body hidden. In this case, the social distance between the speaker and the audience is distant, and if the speaker stands right in front of the audience without any object blocking his/her body, he/she is using the authoritative space, which constructs the speaker–audience relation as an institutional one of medium distance. If the speaker stands among the audience (e.g. checking how tasks are done, talking to individual audience members, etc.), he/she establishes an interactional space in which the speaker–audience relation is close. In public lectures, the speaker normally just stands on the platform or sometimes behind a lectern, but in regular classroom teaching, or in our case an informal AP, the strategic balance of authoritative space and consultative space is crucial for the construction of ideal speaker–audience relation for effective engagement.

Affect refers to the speaker’s engagement of audiences’ emotions. We distinguish between the emotional engagement achieved by provoking audience’s emotion and by inviting audience to empathize with the speaker, following Martin and White’s (2005) framework of appraisal. In APs, the range of affect is quite limited and we will only briefly talk about the use of humor (i.e. telling jokes) for the category of provoking emotion in the following section (identified through audience laughter) (e.g. Knight, 2010; Reershemius, 2012). The speaker may also just express his/her own emotions to invite the audience’s empathetic emotions. The primary and most effective resource for emotional engagement is facial expression (Ekman & Friesen, 1975). In addition, this is also achieved through the linguistic recounting of an event (e.g. sharing a happy event, an anecdote), and the vocal features along with it (e.g. pitch level and loudness) (see Feng and O’Halloran, 2013 for a comprehensive discussion of emotion expression).

To summarize this section, we have outlined a framework which captures the verbal and nonverbal resources that speakers use to engage the audience. With the framework, we can describe accurately what choices experienced speakers make. Echoing Bateman’s (2014) argument for empirical research in multimodal studies, the framework is explicit enough to support reliable annotation of a large quantity of data, as exemplified below.

Data and analysis

The AP examined in this chapter, titled Literature Meets Grammar, was delivered by an expert speaker. The talk focuses on the exploration of how an SFL approach to grammar can contribute towards an informed appreciation of literary works. The AP was audio and videotaped. However, the angle of the camera was not ideal as it was shot from an oblique angle, rather than a preferred frontal angle. Informed consent to use the data from the speaker was confirmed. The presentation was viewed multiple times and extracts that were seen to be highly interpersonal were identified, and the linguistic, visual and acoustic resources used in these extracts were transcribed and annotated. Due to space, we use one short extract from the talk to illustrate the range of engagement resources found in the AP, based on the framework set out in the previous section.
In order to be able to explore the co-deployment of multiple semiotic resources, we need a multimodal annotation tool that can time-synchronize the transcription and annotation with linguistic, visual and acoustic data. Among the available multimodal annotation tools (e.g. ANVIL, ELAN and Multimodal Analysis Video (MMAV)), MMAV, developed by the Multimodal Analysis Company in 2013 (http://multimodal-analysis.com/), was selected. MMAV provides greater opportunity to annotate and generate descriptive results across a range of data (see Smith et al., 2014 for a detailed discussion of the available functions). The software provides a range of graphical user interfaces (GUIs) so the analyst can import and view different media, enter systems networks, create time-stamped tier-based annotations and overlays and other useful analytical approaches for multimodal analysis (Smith et al., 2014, p. 273).

The interface is illustrated in Figure 32.2, where area (A) is the film strip and sound strip visualization area; (B) is transcription of verbal text; (C) is the system choices that we input in the software library (that is, the framework in Figure 32.1); and (D) is the annotated area where the systems are rendered as different tiers, and time-stamped annotation of the unfolding video can be conducted by choosing from the systems in (C) area. Using MMAV as an annotation tool and the explicit framework in Figure 32.1 as the annotation scheme, an analysis of extracts from the speaker was undertaken. The annotation can be exported to an Excel spreadsheet for statistical description, so that a synoptic view of the patterns of choices made can be generated (e.g. the percentage of a certain gesture and a certain speech function). Moreover, the time-stamp multi-tier annotation allows us to examine the multi-semiotic choices made at each point and the inter-semiotic relation based on empirical data, and to analyze dynamic change of the choices in the logogenetic development of the

Figure 32.2 Sample view of multimodal annotation in Multimodal Analysis Video
video (that is, the semiotic choice according to the contextual demand of each stage of the discourse). In the following section, we will briefly report some findings from our analysis.

As shown in Figure 32.2, the annotation includes the dimensions from the system network in Figure 32.1. In this section, we will briefly discuss our findings from the annotation of a three-minute clip about the role of grammar in teaching poetry, and specifically the development of the nominal group from the talk. The transcript is provided here for easy reference. We will first discuss the speech functions of the speaker’s utterances, focusing on how she unpacks and communicates abstract linguistic knowledge. Then we elucidate how different semiotic resources are co-deployed to maximize audience engagement.

Text 1

00:01: so that’s one way. that this poet. is playing with language
00:08: I call them collapsed clauses. its premodifiers
00:14: is that. are they any use to people?
00:18: so we take a whole clause and collapse it down into classifiers
00:24: and we did this with some kids, we demonstrated this, we modeled it and then asked them to talk about dogs in this case. and then we came up with these premodifiers
00:36: ‘an elegant, spotted, fluffy-bottomed, tootsie-feet, shiny-eyed, smelly-ankled, sley-pulling police Husky who was highly toilet trained’
00:51: ‘a jaw-shattering, drool-making, mouth-waterying bone’
00:57: ‘a scrunchy-faced dog with vampire-like teeth’
01:06: and then we could look at this, looking out at both premodification and postmodification in a nominal group so building up the description
01:16: so we could have just said. ‘a mouse’, but that wouldn’t have really built up much the picture about this mouse
01:25: but if it were ‘a timid mouse’, now we know a bit more about the mouse, don’t we?
01:30: (whispering) a timid mouse
01:32: ‘a rather timid mouse’
01:36: so we start to know this mouse a little bit better, but we could say more about it
01:41: ‘a rather timid mouse who’d made her home within the house’
01:47: so you can see how it’s building up the meanings around the nominal group
01:56: here’s another one, ‘we had birds, gold-fish, a fine dog, rabbits, a small monkey, and a cat. This latter was a remarkably large and beautiful animal that was entirely black, and sagacious to an astonishing degree.’
02:14: so again, here we got animal, at the heart of the nominal group
02:19: what can we see about this animal?
02:20: well it was beautiful
02:23: it was large and beautiful
02:27: it was a remarkably large and beautiful animal
02:32: and what more can we say about it?
02:35: well we could keep adding information… afterwards … as a ‘qualifier’
02:42: ‘a remarkably large and beautiful animal that was entirely black, and sagacious to an astonishing degree’
Interpersonal meaning and audience engagement in academic presentations

02:50: so again it’s that…the availability of…resources…ways of adding and enriching the description through the nominal group and that’s what it’s good for

03:04: it’s good for developing erm the description

Initially if we simply look at the speech functions of the text, we can see that there are a number of linguistic features that realize engagement; for example, the use of interrogatives, whether they are yes/no or tag or more open questions (e.g. are they any use to people? don’t we? what more can we say about it?). The speaker does not expect a response, and interrogatives are posed more as a rhetorical device for emphasis and to engage the audience. As the primary purpose of APs is explaining abstract concepts and findings (i.e. offering information), we will examine in detail how a skilled presenter unpacks linguistic knowledge and communicates to the audience. Two notable features include the use of pronouns, and the development of nominal groups. In terms of pronoun, the speaker often uses the personal pronoun we to involve the audience. It is possible that the speaker explicitly chooses we in order to be inclusive and embrace multiple potential listeners – e.g. we the research team, teachers and students and the audience of applied linguists – perhaps?

As the speaker highlights how learners developed control of the nominal group and when referring to examples from the learner’s texts, the speaker gestures frequently to elucidate and clarify information. When the body language is co-expressed with spoken language, the body can play an important role in ‘invoking attitude through the grading of the meanings’ (Hood, 2011, p. 44). The findings demonstrate that the presenter used her body language accompanying verbiage to grade the meaning of intensity. As illustrated in Figure 32.3, the tension in the speaker’s hand gesture and posture co-expresses the intensification of meaning

![Figure 32.3 Hand gesture and posture expressing intensification](image-url)

*Figure 32.3 Hand gesture and posture expressing intensification*
together with verbiage a bit more. At the exact same time that the speaker says a bit more, her facial expression changes, she beats her hand and the stress in the clause is on the three words a bit more. The semiotic spanning between the gesture and the linguistic intensifies the meaning. We find at a later point when the speaker is highlighting the development of the nominal group through another example (e.g. it was beautiful, it was large and beautiful, it was remarkably large and beautiful) that when the new lexis is introduced (e.g. large and remarkably), these words are accompanied with a particular hand movement, a beating of the hand as these words are spoken. Together, the word and the hand gesture realize an increase in the intensification. There are numerous other linguistic resources that contribute to audience engagement: the choices of process types, logical semantic relations of the clauses, the choice of the organization of information, the sequencing of the knowledge, etc., are all important areas that impact the engagement and are areas for further research.

The second dimension in our framework is involvement where the speaker uses gaze, body direction, pointing, movement, etc., to engage the audience’s attention and interest. The co-deployment of the frontal oblique angle of body direction and gaze toward the audience characterizes the speaker’s engagement with the audience, and at the same time saying we did this with some kids, we modeled it, we came up with these premodifiers, etc. The speaker follows this by shifting the gaze from the audience to the slides, and using the laser presenter to highlight specific words on the slide. Directing the audience’s attention to the slide pulls them in to a point where they are encouraged to engage with the ideational features presented through the combined linguistic, audio and visual resources.

Although speech/sign function (i.e. exchanging information) and involvement are discussed separately in our framework (see Figure 32.1), it should be noted that these
two aspects of engagement functions are always correlated and co-constructed through multimodal resources. As shown in Figure 32.4, the presenter initiated engagement by standing in a frontal angle, with her gaze directly towards the audience in anticipation of the question *are they any use to people?*. This question, accompanied by the gesture of opening her arms horizontally and a rising tone, invites the audience to be involved and to assess the value of the information being offered. At this point, the speaker is not expecting a response of *yes this is very useful or no it’s not useful*, but the onus for engaging with the object of study is on the audience. Or, perhaps the speaker was adopting a knower stance and using this interrogative more like an interpersonal metaphor where although a question is being raised, the speaker implicitly means *this should be useful to you*. However, in the nature of the speaker, we would argue that this is a true interrogative and that the speaker is actually saying *you may or may not find the information relevant*. By combining the linguistic with the visual body language and an additional pause for time to think, the speaker encourages participation and engagement within the talk.

The third salient feature about the talk is multimodal construction of humor to elicit brief moments of happiness from the audience. As shown in Figure 32.3, the speaker demonstrates the value of focusing on grammar when teaching poetry, and she highlights in her talk the student example *‘an elegant, spotted, fluffy-bottomed, tootsie-feet, shiny-eyed, smelly-ankled, sley-pulling police Huski who was highly toilet trained’*. In the highly conscious selection of the example of the student’s nominal group included on the slide, *smelly-ankled* was explicitly chosen in order to amuse and entertain the audience (i.e. invoking affect in Figure 32.1). At this point, there is a relative shift in voice quality and a noticeable difference when the speaker states *smelly-ankled* with a rise and fall tone, raising an ironic question or doubt about the choice of a dog having ‘smelly ankles’. The irony of the *smelly ankles* is accompanied by specific body language, the presenter opens her arms and shrugs her shoulders – mimicking the semantics of ‘why did the student choose smelly-ankled? What a strange choice? Wouldn’t you agree this is a strange choice?’ At this point, there is communal laughter heard from the audience, who appreciate the entertainment value of *smelly ankles* and support the presenter’s suggestion that it was an odd choice. Humor is an element of a presentation that brings the audience together, and generates solidarity through the shared understanding of laughter (e.g. Knight, 2010; Reershemius, 2012). Summarizing our discussion of the short video clip, we can conclude that it is the multi-semiotic resources used by the presenter and the semiotic spanning of the linguistic, visual and audio resources that combine with each other to construe the meaning potential and the solidarity between the speaker and the audience during an AP. The combination of the linguistic, visual and audio choices made can realize textual, ideational and the interpersonal meanings in an AP, which can be further investigated. The semiotic choices can be used to construct inter-semiotic cohesion, that is, referring to certain words or images on the slide, links between stages in the talk, or other features related to the mode and organization of the talk. The ideational elements help to develop the content: the object of study. For example, in the short extract, when the speaker refers to *building up*, she opened and raised her arms to symbolically represent and reinforce the meaning. All these resources can be used as engagement resources to construct the intended relationship between the speaker and the audience.

**Conclusion and future directions**

To conclude, in classes focusing on improving presentation techniques, we often hear advice such as ‘you have to connect with your audience’, ‘you have to use more body language
to engage your audience’, etc. (www.TED.com; Duarte, 2010; Gallo, 2014). However, exactly how this is achieved is unclear. We need to have a more explicit ‘metalanguage’ so that speakers know exactly how to design and combine their speech, gestures, movement, etc., to clarify ideas, to engage the audience and to organize the AP. Through an analysis of a small extract from the middle of an AP we have identified how these meanings combine and demonstrated the value of a multimodal investigation into APs. We would like to highlight possible directions for future multimodal research related to APs:

1. As the semiotic theories that are developed to model meaning-making in multimodal resources are often criticized for being based on single illustrative cases, which lack an empirical basis, developing a rigorously annotated multimodal corpus is an important step towards empirical multimodal research (Bateman, 2014, p. 238). The potential for future research would involve collecting a wider range of APs, and analyzing through digital software the range of multimodal features used for meaning-making. The greater the number of APs analyzed, the greater the understanding of common linguistic, visual and audio features found in these texts.

2. The theoretical descriptions of nonlinguistic resources should be further refined based on the empirical data. The huge benefit of using software analytical tools is that they allow the researcher to archive, compare and generate statistical data focusing on and developing a shared framework. The development of a shared framework will enable studies to provide a more convincing picture of the meaning potential found within APs.

3. As Holsanova (2012, p. 251) observes, ‘given the dominant role that multimodality plays, there is still a lack of empirical studies on how recipients interact with multimodal messages’. In future research, we can use questionnaire surveys to ask the audience to comment on the multimodal features of different presentations, or use a focus group after a presentation to discuss the audience’s reaction to the use of gesture, gaze, etc.

We pose a question for further thought: ‘if the content is excellent, but the speaker does not engage with the audience – does this matter?’ We believe it does. We believe that EAP teachers are committed to teaching interactive, interpersonal features that help speakers engage with their audience in a classroom; that in the workplace and in academia, if a speaker can engage with the audience when presenting, the warrantability of his/her proposition will be strengthened. We also believe that if a speaker has excellent content plus effective presentation skills, his/her status within the academic community can be enhanced. Through research into a comprehensive understanding of the semiotic range and resources used to construe meaning in APs, we can better scaffold improvement in students, colleagues and others involved in giving presentations.

Further reading

Kress & van Leeuwen (2006)

Related chapters

15 Systemic functional linguistics and EAP
20 Multimodal approaches to English for academic purposes
26 Seminars
References


