Genre innovation and multimodal expression in scholarly communication: Video methods articles in experimental biology

Christoph A. Hafner
City University of Hong Kong (Hong Kong)
c.hafner@cityu.edu.hk

Abstract

Digital media have the potential to foster genre innovation and change in scholarly communication, by 1) opening up new, diversified audiences to academics, and 2) facilitating the use of a range of multimodal semiotic resources, that combine word and image, in academic writing. However, relatively little applied linguistic research has investigated innovation in research genres, especially high stakes genres like the research article. In addition, some of the existing literature on the topic has concluded that the addition of digital elements to research articles (for example, embedded video) fails to add significant meaning to the genre, perhaps indicating a poor match between the affordances of digital media and the communicative purposes of the academic writers. This exploratory study provides a multi-dimensional genre analysis of a new research genre in the field of science: the video methods article (VMA), published online by the Journal of Visualized Experiments (JoVE). In order to understand the intertextual communicative context of this genre, community documents were reviewed and two specialist informants were interviewed. A sample of eleven VMAs from JoVE, one per year of publication from 2006-2016, was examined. The multimodal analysis shows how the VMA genre draws on the affordances of digital video in order to meet genuine needs of academic writers. The findings also show some innovation and development in the genre over time, which moves through an initial period of experimentation before settling on a stable generic structure.

Keywords: genre innovation, digital media, digital literacies, multimodality, English for academic purposes, English for research publication, digital video.
Resumen

Innovación de género y expresión multimodal en la comunicación académica: El vídeo-artículo de métodos en biología experimental

Los medios digitales pueden propiciar la innovación de géneros textuales y cambios en la comunicación académica por dos vías: 1) proporcionando a la comunidad científica públicos nuevos, más amplios y diversos, y 2) facilitando el uso de una gama de recursos semióticos multimodales que combinen palabra e imagen. La Lingüística Aplicada, sin embargo, ha estudiado poco la innovación en los géneros textuales empleados en la investigación, especialmente en aquellos que, como el artículo científico, revisten una importancia fundamental. Parte de la literatura especializada, además, ha llegado a la conclusión de que añadir elementos digitales a los artículos científicos (insertando vídeos, por ejemplo) no aporta ningún significado adicional de interés, lo cual tal vez sea indicativo de un desfase entre el potencial tecnológico del medio digital y el propósito comunicativo del autor. El presente artículo analiza, desde una perspectiva multidimensional y exploratoria, un nuevo género de investigación científica: el vídeo-artículo de métodos (VAM), que publica en Internet la revista especializada Journal of Visualized Experiments (JoVE). Para comprender el contexto comunicativo intercultural del género, se han revisado los documentos de una comunidad disciplinaria, se ha entrevistado a dos informadores especialistas y se ha examinado una muestra de once VAMs de JoVE, uno por cada año de publicación de 2006 a 2016. El análisis multimodal muestra cómo el género del VAM recurre al potencial del vídeo digital para satisfacer las necesidades comunicativas genuinas de los autores académicos. Los hallazgos del estudio evidencian asimismo cierta innovación y desarrollo del género a través del tiempo, desde un período experimental inicial hasta la consolidación de una estructura retórica estable.

Palabras clave: innovación de género, medios digitales, competencias digitales, multimodalidad, inglés para fines académicos, inglés para la publicación científica, vídeo digital.

Introduction

There have, for some time now, been calls for language and literacy educators to engage with emerging digital and multimodal forms of expression (e.g. New London Group, 1996; Kress, 2000; Belcher, 2004; Hafner, Chik & Jones, 2013, 2015; Early, Kendrick & Potts, 2015). Although such new forms of expression have emerged at a dizzying pace in popular contexts, they have been much slower to take hold in the academic context,
where traditional research genres using traditional forms of expression have tended to predominate. Relatively little attention has been paid to the ways in which digital media potentially affect academic genres. As Pérez-Llantada (2016: 24) puts it, “There is therefore a need for stimulating scholarly conversation on issues of genre innovation and change so as to better understand the new forms of research communication in today’s academic settings”. In this article, I hope to contribute to this conversation by exploring an important innovation in scientific research genres, namely, the development of the video methods article.

Existing applied linguistic research into digital research genres includes work on academic homepages (Hyland, 2011), academic blogs (Blanchard, 2011; Luzón, 2012, 2013), and innovative digital forms of expression in the research article (RA) (Pérez-Llantada, 2013). Given the central importance of the RA in academic genre systems, it is an especially interesting genre to consider. Pérez-Llantada (2013) investigates the so-called Article of the future introduced by the academic publisher Elsevier in order to provide academic writers with a number of enhanced elements for expression: an interactive web interface with navigation sidebars, research highlights, graphical abstracts, embedded video and interactive figures and graphs. According to the Elsevier website, “The Article of the Future initiative aims to revolutionise the traditional linear format of the academic paper to make it more dynamic and user-friendly” (Zudilova-Seinstr, 2013: n.p.).

Experimenting with the form of scholarly communication in RAs is not unusual: for example, a recent special issue of the *European Journal of Information Systems* examines six alternative representational forms in that discipline (Avital, Mathiassen & Schultze, 2017). However, one possible criticism of initiatives like Elsevier’s Article of the Future is that the simple introduction of new digital elements, for all their potential, fails to add significant new meanings that address the genuine communicative needs of academics. As one critic puts it, the new features are “just lipstick” (Pérez-Llantada, 2013: 222). Comparing a corpus of Article of the Future “prototypes” (i.e. commissioned articles that used the new online elements) with a larger journal article corpus, Pérez-Llantada examined macro structural and linguistic features and found that the new online elements in the Article of the Future did not affect “actual text-composing and reception practices” of the RA genre (page 233). However, that study focused on the main text of RAs and did not consider the new part genres (for example, embedded videos), where discursive innovations might well be observable. It
remains unclear how the RA genre might be developing as a result of innovations in expression resulting from the affordances of digital media. More research that examines the way that academics recruit these affordances in order to serve their communicative purposes is needed.

One important affordance of digital media is to allow academics to reach out to and interact with a much wider audience than would be possible through traditional means. Academic blogs provide an interesting example of scholarly communication that engages a wider, “diversified” audience in this way (see Luzón, 2012, 2013). Such blogs are publicly available and allow readers to “write back” to the blog author by posting comments. They likely engage a range of different audiences simultaneously: disciplinary specialists, specialists from related disciplines, as well as non-specialists like policy makers and a popular audience of interested amateurs. The discussions that take place in academic blogs can have a profound impact on research writing. For example, Blanchard (2011) describes how the philosopher John S. Wilkins acknowledged his “blog readers for comment and discussion” in an article published in the academic journal *Studies in History and Philosophy of the Life Sciences*. Similarly, access to this diversified audience could have an effect on the way that academics go about seeking funding. Mehlenbacher (2017) examines science-focused crowdfunding proposals on the website Kickstarter.com, suggesting the development of a new, hybrid genre of funding proposals aimed at a wider audience.

A second important affordance of digital media is to increase the range of semiotic resources available to academic writers, especially by facilitating multimodal forms of expression that combine words, images and other modes. Considerable applied linguistic work has gone into this area, with the emergence of the field of “multimodality”, a social semiotic approach to communication (Jewitt & Kress, 2003; Kress, 2010). Within this tradition, the work of Kress and van Leeuwen (2006), which examines the design of visuals, has been especially influential. In the field of ESP, we have also seen considerable interest in the way that visuals are recruited in order to communicate, establish identities, and learn disciplinary practices (Johns, 1998; Rowley-Jolivet, 2002, 2004; Tardy, 2005; Prior, 2013). This work has tended to focus on the use of static images in conference presentations and student assignments and there is therefore a need to extend existing research by examining multimodal expression in academic video genres.
In view of this background, the aim of this article is to explore how the affordances of digital media may serve to enhance scholarly communication in a meaningful way and, in so doing, contribute to genre innovation and change. In the article, I provide a description of the emerging scientific genre of video methods article: its intertextual and interdiscursive communicative context, its generic structure, its multimodal semiotic features, as well as some observations about how the genre has developed over time. In light of this analysis, I consider how the affordances of the web-based digital video medium potentially affect the generic features observed. I also evaluate the extent to which the features observed can be seen either as innovative or as building on existing features of academic discourse.

The study

The Journal of Visualised Experiments (JoVE)

This study focuses on a scientific video journal called the Journal of Visualised Experiments, or JoVE for short. JoVE was established in December 2006 in order to address a serious problem in science: the difficulty of replicating scientific procedures reported in written protocols in RAs. To do so, it employs “video methods articles” (VMAs) that draw upon the affordances of digital video in order to visually demonstrate experimental procedures. The use of digital video is intended to enhance the communication of procedures by filling in the gaps encountered in a written protocol with visual information. A video demonstration has the potential to precisely convey minute details of, for example, how researchers must handle equipment in laboratory or surgical procedures. Access to such a visual record can in turn save other researchers time and money, as they spend less effort on perfecting techniques to use in their own research.

JoVE therefore presents an interesting example of specialised disciplinary communication that has adopted an innovative form of representation, i.e. digital video, in an attempt to solve a serious communication problem. Note that the journal is well respected in the scientific community, being listed in indexes like Scopus and PubMed Database, with a 2016 impact factor of 1.232 according to Journal Citation Reports. The journal focuses on multiple disciplines in the physical and life sciences, with thirteen areas of interest: behaviour, chemistry, genetics, biochemistry, developmental biology, immunology and infection, bioengineering, engineering, medicine, biology,
environment, neuroscience, cancer research. The journal also includes a science education section, which provides a range of video resources to teach the basics of science, from the fundamentals of the biology lab to clinical skills needed by medical doctors.

**Data sources and analytical methods**

This exploratory study adopts a multi-perspective approach to genre analysis (Bhatia, 1993, 2004; Hafner, 2010) drawing on multiple sources of data to understand the video methods article (VMA) genre. This approach aims to go beyond a formal description of texts to develop an understanding of the genre as a social practice. This involves understanding the intertextual and interdiscursive communicative context within which the genre is situated, including processes of text construction and the social relationships of writers and readers as members of a discourse community that employs the genre in order to achieve strategic communicative goals. In order to develop this more detailed understanding of the communicative context of the genre, researchers may draw on secondary data sources to complement their analysis of text. These include: 1) existing community texts that describe and explain the genre, e.g. handbooks and guidelines; 2) interviews with specialist informants, community insiders, who draw on their expert knowledge in order to explain features observed in the analysis.

The present study involved the collection of community texts, namely relevant pages from the *JoVE* web site (e.g. “About *JoVE*”, “Publishing Process”, “Editorial Policies”) and “Instructions for authors”. In addition, two specialist informants were consulted and interviewed (for about 30-60 minutes) on questions relating to the production process, authoring and reception of the genre. One was Moshe Pritsker, the CEO and co-founder of *JoVE*. The other was Raymond Wong (a pseudonym), a biologist at the City University of Hong Kong, who, as well as answering general questions about the genre, provided some feedback on initial analysis. While Wong does not have first-hand experience of creating VMAs, he is an occasional user of such videos, and *JoVE* has approached him in the past to suggest that he might like to author a video about his own research methods. Pritsker was interviewed over Skype while Wong was interviewed in person. Both gave informed consent to participate in the study. The interviews were audio recorded and transcribed for later thematic analysis.
The primary source of data is a set of eleven VMAs, one per year from 2006 to 2016 (see Appendix for complete citations). In order to sample videos that had achieved clear community acceptance, articles with high numbers of citations in the Scopus database were selected. Also, selecting one highly cited article per year allows for the observation of potential developments over time in this innovative genre. The selection was limited to articles in the biology subsection of JoVE in order to limit possible confounding effects of discipline on the genre analysis. Biology was a convenient choice for two main reasons: first, many of the highly cited articles in JoVE appear to come from biology; second, the specialist informant was a biologist and well placed to comment on these articles. VMAs appear embedded in a dedicated web page, which, as well as the video, includes a written protocol and a comments section, where readers/viewers can leave comments for authors. Therefore, after downloading the selected videos, the related written protocol was downloaded as a pdf file, and comments were also copied for later analysis. Table 1 summarises key characteristics of the eleven VMAs finally selected for analysis.

The VMAs selected have received between 14 and 172 citations (81 on average). They are between 8:01 and 16:11 minutes long (11:35 on average). According to the view counter on the JoVE website they have been viewed between 5,196 and 65,098 times (32,062 on average). They have received between zero and 71 comments (22 on average). The subject matter includes a behavioural test for mice, two different surgical procedures for mice, and a range of lab procedures that aim to detect, identify, map, isolate, quantify, and measure the size of different kinds of biological samples (descriptions in the table are based on the stated aims recorded in the videos). Seven out of the eleven samples are available open access, while the other four require a subscription.

Once identified, the eleven videos were viewed one by one and initial observations were made about the multimodal presentation strategies, including potential genre moves and their multimodal semiotic expression. The spoken narration of the videos was transcribed and selective multimodal transcriptions were made. This multimodal transcription followed a similar procedure to that reported in Hafner (2014, 2015), using ELAN (Max Planck Institute for Psycholinguistics, 2018), software that allows the user to create a multidimensional transcription of video. In this procedure, the video was first divided according to the shots observed (typically a robust unit of analysis), with shots further subdivided into visual...
frames if more than one visual frame was presented in the shot (following Baldry & Thibault, 2006). The following dimensions were transcribed for each shot: 1) speech, i.e. what was said in the video; 2) action, gesture, gaze performed by visual participants; 3) soundtrack; 4) camera position and perspective, i.e. the vertical angle of the camera, the horizontal angle, and the distance of the shot (close up, medium, long); 5) setting and participants, i.e. where the shot took place and who/what was in it; 6) text on screen.

<table>
<thead>
<tr>
<th>#</th>
<th>Year</th>
<th>Cites</th>
<th>Comments</th>
<th>Views</th>
<th>Length</th>
<th>Open access?</th>
<th>Field</th>
<th>Subject matter</th>
</tr>
</thead>
<tbody>
<tr>
<td>VMA1</td>
<td>2006</td>
<td>58</td>
<td>3</td>
<td>20,829</td>
<td>10 min</td>
<td>Yes</td>
<td>Biology</td>
<td>A behavioural test for mice</td>
</tr>
<tr>
<td>VMA2</td>
<td>2007</td>
<td>61</td>
<td>58</td>
<td>24,598</td>
<td>15 min</td>
<td>No</td>
<td>Biology</td>
<td>A surgical technique on mice</td>
</tr>
<tr>
<td>VMA3</td>
<td>2008</td>
<td>107</td>
<td>17</td>
<td>55,913</td>
<td>8 min</td>
<td>No</td>
<td>Biology</td>
<td>A surgical technique on mice</td>
</tr>
<tr>
<td>VMA4</td>
<td>2009</td>
<td>121</td>
<td>25</td>
<td>19,271</td>
<td>10 min</td>
<td>No</td>
<td>Biology</td>
<td>A lab procedure for detecting proteins</td>
</tr>
<tr>
<td>VMA5</td>
<td>2010</td>
<td>107</td>
<td>39</td>
<td>34,048</td>
<td>12 min</td>
<td>Yes</td>
<td>Biology</td>
<td>A lab procedure to identify binding sites of RNA binding proteins</td>
</tr>
<tr>
<td>VMA6</td>
<td>2011</td>
<td>103</td>
<td>71</td>
<td>34,292</td>
<td>10 min</td>
<td>Yes</td>
<td>Biology</td>
<td>A lab procedure to map binding sites of RNA binding proteins</td>
</tr>
<tr>
<td>VMA7</td>
<td>2012</td>
<td>172</td>
<td>27</td>
<td>65,098</td>
<td>9 min</td>
<td>Yes</td>
<td>Biology</td>
<td>A lab procedure to isolate exosomes</td>
</tr>
<tr>
<td>VMA8</td>
<td>2013</td>
<td>53</td>
<td>1</td>
<td>14,703</td>
<td>16 min</td>
<td>Yes</td>
<td>Biology</td>
<td>A lab procedure to create a cell free expression system</td>
</tr>
<tr>
<td>VMA9</td>
<td>2014</td>
<td>67</td>
<td>1</td>
<td>63,595</td>
<td>8 min</td>
<td>No</td>
<td>Biology</td>
<td>A lab procedure to quantify cellular anchorage</td>
</tr>
<tr>
<td>VMA10</td>
<td>2015</td>
<td>23</td>
<td>0</td>
<td>15,141</td>
<td>11 min</td>
<td>Yes</td>
<td>Biology</td>
<td>A lab procedure to quantify and measure the size of exosomes</td>
</tr>
<tr>
<td>VMA11</td>
<td>2016</td>
<td>14</td>
<td>0</td>
<td>5,196</td>
<td>11 min</td>
<td>Yes</td>
<td>Biology</td>
<td>A lab procedure to assess the relative abundance of histone post-translational modifications</td>
</tr>
</tbody>
</table>

Table 1. Video methods articles selected for analysis.
A qualitative, multimodal genre analysis was performed, attending to relevant visual and verbal cues to identify sections and moves in the video article genre. For example, title frames occupying the whole screen (similar to a heading in text) provided a useful visual cue that sometimes delineated a section of the genre. At the same time, the video narration provided a verbal cue to further identify genre sections and moves. Once sections and moves were identified in this way, the multimodal semiotic resources typically employed to realise these moves and sub-moves could be analysed in greater detail. The analysis of text was supported by a thematic analysis (Braun & Clarke, 2006) of community documents and transcribed interviews with specialist informants. The approach taken was primarily deductive in nature, with data coded according to the themes of communicative context (audience, purpose, and reader use), composing process (overview, production, script writing, filming, editing, proofing, challenges), and innovation. The entire analysis was facilitated by the use of MaxQDA (Verbi Software, 2018), qualitative data analysis software that allows the user to code a range of text types.

Findings and discussion

The intertextual communicative context

Information about the target audience(s), the purpose of communication, and the process of text construction, can be gleaned from information on the JoVE website as well as from comments by specialist informants. The following statement from the “About JoVE” page gives a sense of JoVE’s mission and intended audience:

By allowing scientists, educators and students to see the intricate details of cutting-edge experiments rather than read them in text articles, JoVE increases STEM research productivity and student learning, saving their institutions time and money.

Like written articles in traditional scientific methods journals, the purpose of the VMAs is to share advances in scientific technologies and techniques with other researchers. Such methods might be faster, more sensitive, more accurate, more user friendly and more cost effective than existing methods. The statement above also draws attention to the use of the digital video medium to fulfil an important aim of JoVE, to share methods in visual form,
so that the audience can “see... intricate details”. Both informants considered the use of this digital video medium as the main innovation of *JoVE*, with Wong (the biologist informant) confirming that “it saves a lot of time” (Interview).

Pritsker (*JoVE* CEO/co-founder) identified the audience(s) for *JoVE* articles as follows:

Audiences are practising researchers, practical scientists who work in the laboratory. So everywhere between professors to postdoc to graduate students to undergraduate, including technicians, basically everybody who does experiments with their own hands and everybody who sometimes has to guide these experiments. And that both in academia but also in biopharma industry. (Interview)

Although articles are sometimes made available in open access format on the Internet, neither informant felt that they would appeal to a wider audience of the general public, because of their technical nature. Pritsker commented: “It’s a professional resource for professionals” (Interview).

The process of text construction builds on traditional forms of scientific discourse and extends these into the presentation of content through video. The *JoVE* website (“Publishing process”) identifies five main steps:

1. Manuscript preparation;
2. Editorial and peer review;
3. Scripting and filming;
4. Video editing;
5. Proofing and publication.

Of these five steps, the first two are identical to those required for publication in a traditional scientific methods journal, with scientists submitting a written manuscript for review. Pritsker noted:

While we have a new media, a new production process, at the same time, we want to preserve the foundation of scientific publishing as we know it, which is peer review based. So we preserve that and then, finishing that we apply this new production process. (Interview)
The last three steps are where the process diverges from traditional publishing, involving the creation of the digital video methods article. Two different processes are possible, according to the website (“Editorial Policies”): 1) “Video produced by JoVE”; or 2) “Video produced by author”. The vast majority of submissions fall into the first category. Steps three and four of the process involve a JoVE production team developing a script based on the written manuscript, sending a film crew to the lab to record the procedure, recording a narration using professional voice actors, and finally editing the footage and narration. In the final proofing and publication step, the video is shared with the scientist authors, who may provide feedback and request changes.

According to Pritsker, JoVE has developed an extensive infrastructure to support the development of VMAs, one that “includes local filmmakers in 25 countries around the world, residing in cities with high concentration of academic institutions”. This makes it possible “to film… practically in every big university around the world” (Interview). Developing this infrastructure became necessary as scientists were generally not capable of producing high quality VMAs without assistance. Pritsker commented:

Our first idea was that eventually, after making a few videos and having them as an example, scientists would submit their own, would make and submit their own videos. And it didn’t work. And the reason it doesn’t work is because making a scientific video is much more complex than making a video for interview for example, right? …So we realise that we have to do it for them… if we want this information out of their heads, we have to do it. (Interview)

What we see here is a complex process of collaborative authorship, where the production of text is divided in numerous ways, which can be understood using Goffman’s notion of “production format”. Goffman (1981) points out that production of text can be divided between a principal, whose beliefs and views are represented by the words uttered; an author, who composes those words; and an animator, who acts as the “sounding box” to make the utterances.

In the typical case for JoVE VMAs (produced by JoVE), the principal is the team of scientist authors who produce the written manuscript with the backing of their laboratory and institution. Regarding the author, we need to consider not only words but also other semiotic resources, especially visual resources. In both cases, the semiotic resources deployed can be seen as
collaboratively authored by the scientist writers and the JoVE production team: after scientists produce the written manuscript and highlight important parts, the JoVE production team uses this to create a script. Similarly, scientist authors will need to consider how best to visually present laboratory procedures, working with the JoVE team to supply diagrams as the basis for conceptual animations, demonstrating procedures for a film crew, and providing representations of results to be incorporated into the video. Finally, a range of participants animate (i.e. in Goffman’s sense of making utterances in) the VMA. Scientists address the camera to explain procedures and feature in the video as demonstrators. Often, these are not the same people, with a more senior scientist explaining a procedure while lab technicians, graduate students, and research assistants take on the demonstrator role. Professional voice actors from the JoVE team animate (again, in Goffman’s sense) a narration of the procedures.

In order to properly understand the VMAs, it is important to note that the videos are just one part of a larger intertextual system, including the video itself, the written article that it is based on, and the website frame that houses both of these texts as well as readers’ comments. These three different kinds of texts may interact and complement one another, drawing on their unique affordances to contribute different kinds of meanings. Video provides visual resources capable of extending the meanings made in the written protocol. As Wong explained,

(...) how a protocol is being executed is being captured visually. So there is like no mistake on how it is being done, especially if, say, if we’re doing micro-injection of DNA into eggs, okay, how do you handle the pipette? How deep do you penetrate it? So there’s lots of finer details which are being sort of captured visually (...) (Interview)

The written protocol provides a complete record of the procedure, including all of the steps that need to be followed, as well as “recipes” that are needed, for example to create particular kinds of materials and solutions. Finally, the comments allow scientists to interact, asking for clarifications of procedures where necessary.

**Generic structure**

Analysis of the videos reveals a stable generic structure, summarised in Figure 1.
Figure 1 shows seven macro sections. All seven sections are obligatory but they may be realised through different moves. Section A, Video Intro, is a brief opening section, whose purpose is to state the title of the video article. It consists of several frames of text on screen: the JoVE logo with an accompanying jingle, a black on white title frame displaying the article title, authors and affiliations. These can be followed by another title frame stating that procedures have been approved by relevant ethics committees. Section B, Overview, provides a summary of the procedure in a conventionalised way: Stating the Goal, Summarising the Procedure and Explaining Significance are all obligatory moves in this section. The Overview section draws on a wide range of visuals: conceptual animations, video of procedures, images of raw data, charts and graphs to aggregate results. Section C, Researcher’s Introduction, provides an opportunity for the researcher to engage the audience and introduce the demonstration of procedures. While the section is obligatory, it is realised in a variety of ways and all of the moves identified are optional. Visually, the section is primarily realised through a shot of the researcher addressing the camera.
The longest section is section D, Demonstration, whose purpose is to demonstrate the procedures, key to the VMA's communicative purpose of sharing techniques. Move D2, Describing Steps, is obligatory and is optionally supported by the other three moves: Introducing Apparatus/Materials, Elaborating on Steps, Explaining/Evaluating Steps, all of which occur very commonly. The spoken narration is combined with video of demonstrators in relevant lab settings, who perform the procedures. In addition, text may appear on screen, especially where additional information is required, for example the ingredients of a solution that must be mixed as part of the procedure. Section E, Representative Results, aims to provide examples of the kind of results that can be obtained from the procedure. Visually, these are represented through images of raw data, and charts and graphs that aggregate results. These are sometimes evaluated by comparing with other techniques in terms of cost-effectiveness, sensitivity, speed and so on. Section F, Researcher's Conclusion, allows the researcher to appear on camera again and re-engage with the audience through the moves listed in Figure 1. Section G, Closing Credits, includes title frames with author and institution information, followed by credits for the production team.

In this paper, I will focus on sections B, C, and D. Section B (Overview) because of the range of visuals that it draws on: conceptual animations, footage of the procedure, visual displays of data and results. Section C (Researcher's Introduction) because in it the researcher addresses the camera directly, engaging with the audience in a novel way that illustrates the hybrid nature of this genre. Section D (Demonstration) because this is the core of the VMA, the move considered most innovative by the expert informants. As the researcher demonstrates the procedure, we can observe an interesting division of labour between different modes like image, spoken narration, and text on screen, which all perform slightly different roles in terms of the kinds of information that they convey. Finally, it is interesting to consider how the genre has stabilised over time and this can be observed in the development of a more and more conventionalised generic structure for section B, Overview.

**Multimodality in the Demonstration section**

Both informants saw the principal innovation of the VMA as the use of video to visually demonstrate novel procedures. An analysis of the Demonstration section therefore has the potential to show how the affordances of digital video are recruited in order to clearly convey experimental techniques and
procedures. Constraints of space do not permit a full multimodal analysis of every move. Instead I will examine how multimodal resources are combined in a short extract from VMA8, shown in Figure 2.

Figure 2 shows two shots, the first (shot 37) being an over-the-shoulder view of a researcher standing at a lab bench with an ice bucket on the left and a centrifuge on the right. This is a Describing Steps move, which uses the imperative form in the narration: “Centrifuge the filter apparatuses…” The video shows the researcher placing samples in a centrifuge. Text on screen summarises key settings: “6,000 x g | 5 min | 4 °C”, reinforcing the narration. Here, the written mode provides viewers with a simultaneous list, introducing the information all at once and preserving it on screen for some time. In contrast, the spoken mode must introduce each element sequentially. This strategy is frequently used with long lists of ingredients that sometimes must be mixed to obtain necessary experimental solutions. Shot 38 begins with a Describing Steps move (“verify that…” and continues with Elaborating on Steps (“Properly beaten extract will not be turbid…”). This move serves to provide additional information and detail. The video here features a still image of two samples, one good one and one poor one, which are described and compared in the narration. Here the visual mode again reinforces the narration as well as elaborating on it by adding visual detail to exemplify the good and bad samples.

Figure 2. Extract from Demonstration section (VMA8, shots 37-38). Images used with permission.
Figure 2 shows how the modes of spoken narration, still and moving image, and written text on screen work together in a complementary way. The visual mode clearly plays an important role in this, as images are capable of answering some of the questions that spoken or written texts leave open. In shot 38, the images make it possible to see precisely what is meant by “distinct layers”. At the same time, there are certain kinds of meanings that are typically made through spoken narration, being difficult to capture visually. These include certain kinds of explanations that detail why a particular step has to be carried out in a particular way. Elsewhere in VMA8, we have an example that depends on a counterfactual statement:

(...) add a very small drop of bead cell solution onto the inside of a bead beating tube cap, being careful to not put solution into the outside lip of the cap. Otherwise, the bead beating tube will not close sufficiently. (VMA8, narration, Demonstration)

Here we see the “functional specialisation” (Kress & Jewitt, 2003) of different modes, which tend to play slightly different roles: the visual mode can be used to illustrate in detail, while the spoken mode can be used to consider hypothetical situations and provide reasons not demonstrated visually. Furthermore, the video text itself is more constrained in terms of time/space than is the accompanying written protocol, so that researchers often refer to the protocol as the most complete account, for example: “Refer to the written article for processing steps after dialysis” (VMA8, narration, Demonstration). Thus, the different texts in the JoVE genre system, with their unique affordances and constraints, also perform different roles and functions: while the video both shows and explains (through affordances of moving image and spoken narration), the written protocol affords the space to provide a complete account.

Hybridity in the Researcher’s Introduction section

In the Researcher’s Introduction section, the researcher appears on screen like a host on a television show (for example a cooking programme). The purpose of the section is to engage the audience and introduce the demonstration of procedures by highlighting key points. A number of optional moves, like Introducing Self, Introducing Other Researchers, and Inviting the Audience, are clearly designed to engage with viewers and sometimes welcome them into the lab where the demonstration is conducted. Of the eleven videos, nine included at least one of these
engagement-oriented moves in the Researcher’s Introduction. In this section, the researcher may draw on informal stance and engagement features to construct a professional identity not available in written RAs. The section likely resembles other speech events that scientists regularly participate in, like conference presentations for their colleagues or lab demonstrations for their students. The section is therefore replete with hybridity, as it draws both on the discourse of hosted popular media programmes and the discourse of spoken scientific instruction.

An analysis of the Researcher’s Introduction in VMA3 demonstrates the hybridity that comes from explicit forms of engagement in this genre and shows also how such engagement is constructed through multiple modes. Figure 3 shows the beginning and end of the Researcher’s Introduction in VMA3. Shot 13 realised an Introducing Self move, consisting of a greeting (“Hi”) and introduction (“my name is”). The researcher is seated at a computer in an office, facing the camera in a medium length demand shot, with gaze on the audience, a highly engaging visual composition. Shot 14 realises a Forecasting the Demonstration move and is accompanied by diagrams related to the surgical procedure. The camera returns to the researcher as he explains key applications of the procedure, culminating in shot 18, which realises the Inviting the Audience move. The researcher’s invitation (“So, let’s get started”) is accompanied by embodied action as he raises his lab coat (folded over one arm), stands up and leaves the shot.

Hi. My name is [name omitted] from [name omitted]’s lab at the [institution name omitted].

Today I will show you how to collect cerebral...

...shots 15-17 omitted...
Compared to the written article that VMAs are based on, the engagement observed in the researcher's Introduction can be seen as an innovative feature of the genre, likely linked to the visual affordances of the video medium. This is to suggest that, because researchers are using video, they potentially engage with their audience in an innovative way. While the section is obligatory, moves that explicitly engage the audience are not. Nevertheless, the section provides an opportunity for researchers to engage the audience in a range of ways. It is worth noting that the section sometimes includes an introduction of other researchers, graduate students or research assistants, who will be doing the actual demonstration of the procedures. In this way, the section sometimes highlights the collaborative, team-based nature of scientific study, drawing attention to the different roles adopted by different authors, again in a way that is not seen in written RAs.

**Development over time: Emergence of a conventionalised Overview section**

Comparing early videos with later ones reveals a noticeable development over time, with later VMAs showing greater sophistication in terms of the range of multimodal resources used and also showing a more conventionalised and stable generic structure. Asked about this, Pritsker noted:

(...) when we really started, we didn't know how exactly everything is going to work. And we have to trial our own to make it with these video articles. And actually the first video articles were filmed by me, in laboratory. So my
friends who agreed to do it. It was very unusual at that time, video journals, something unheard of.

We see here the gradual development of the video methods article genre. Table 2 presents the sections observed in the first five video methods articles in the corpus.

<table>
<thead>
<tr>
<th>VMA1</th>
<th>VMA2</th>
<th>VMA3</th>
<th>VMA4</th>
<th>VMA5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Video intro</td>
<td>Video intro</td>
<td>Video intro</td>
<td>Video intro</td>
<td>Video intro</td>
</tr>
<tr>
<td>Demonstration</td>
<td>Researcher’s Introduction</td>
<td>Overview</td>
<td>Overview</td>
<td>Overview</td>
</tr>
<tr>
<td>Representative Results</td>
<td>Demonstration</td>
<td>Researcher’s Introduction</td>
<td>Researcher’s Introduction</td>
<td>Researcher’s Introduction</td>
</tr>
<tr>
<td>Credits</td>
<td>Researcher’s Conclusion</td>
<td>Demonstration</td>
<td>Demonstration</td>
<td>Demonstration</td>
</tr>
<tr>
<td>Credits</td>
<td>Representative Results</td>
<td>Researcher’s Conclusion</td>
<td>Researcher’s Conclusion</td>
<td>Researcher’s Conclusion</td>
</tr>
<tr>
<td>Credits</td>
<td>Credits</td>
<td>Credits</td>
<td>Credits</td>
<td>Credits</td>
</tr>
</tbody>
</table>

Table 2. Sections observed in VMA1 to VMA5 (2006-2010).

VMA1 (2006) consists of only two sections of substance: Demonstration and Representative Results. A narrower range of multimodal resources is used in this video: there is no soundtrack and no oral narration of any kind, with text on screen performing the narrative role. In VMA2 (2007), the Researcher’s Introduction and Conclusion are added. Throughout this video, the researcher verbally narrates the procedures describing what we see happening in the lab, with frequent long pauses as the demonstration catches up with this narration. This narration is at most semi-scripted and the delivery is perhaps similar to what one would observe in an actual lab demonstration with the researcher. In VMA3 (2008), the Overview section is added, serving to summarise the procedure before it is demonstrated in detail. This is highly scripted and draws on a wide range of visuals, including both still and moving images of the (surgical) procedure being performed in the lab, as well as drawings/diagrams of the procedure with accompanying text on screen. The Overview in VMA3 is accompanied by a professionally produced narration, performed by a voice actor and added in post production. There is an unobtrusive soundtrack. Finally, VMA4 (2009) and VMA5 (2010) adopt the conventional generic structure with all seven sections present. Both present a wide range of multimodal resources, with a sophisticated combination of spoken narration, still and moving images, animated diagrams, text on screen, and soundtrack.
Considering only the Overview section in VMA3, VMA4, VMA5, we also see some change over time (recall also that there was no Overview present in VMA1 and VMA2). Table 3 shows the moves observed.

<table>
<thead>
<tr>
<th>VMA3</th>
<th>VMA4</th>
<th>VMA5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summarising the procedure</td>
<td>Forecasting the demonstration</td>
<td>Stating the goal</td>
</tr>
<tr>
<td>Summarising the procedure</td>
<td>Summarising the procedure</td>
<td>Explaining significance</td>
</tr>
</tbody>
</table>

Table 3. Moves observed in the Overview section of VMA3 to VMA5 (2008-2010).

All three articles contain the Summarising the Procedure move in their Overview section, but again this becomes more complex over time. VMA3 (2008) is a simple summary, with narration beginning as follows:

The procedure begins with anaesthetising the mouse and the preparation of the surgical site. The mouse is then placed… (VMA3, narration, Overview).

VMA4 (2009) is similar but the narration begins by explicitly forecasting the demonstration:

In this video we will demonstrate the power of a modified coomassie staining protocol for detecting several nanograms of protein in polyacrylamide gels, using the example of two dimensional gel electrophoresis. (VMA4, narration, Overview)

VMA5 (2010) adopts the highly stable and conventional structure that is observed in all subsequent articles in the corpus: Stating the goal, Summarising the procedure, Explaining significance. Key elements of the narration are:

The overall goal of this procedure is to determine the transcriptome wide binding sites of R N A binding proteins or R B Ps. This is accomplished by first using photoreactive nucleocide analogues for in vivo labelling of nascent transcripts. The second step of the procedure… Ultimately, results can be obtained that show transcriptome wide binding sites of the R V Ps and enable the definition of the R N A recognition element and its binding preferences. (VMA5, narration, Overview)

Note the highly formulaic language, variations of which occur in almost all of the subsequent VMAs in the corpus: “The overall goal of this procedure
is to…”; “This is achieved by…”; “results can be obtained that show…”. A full multimodal analysis of this section would also be interesting but given constraints of space, such an analysis is not possible here.

Discussion and conclusions

While this exploratory study is based on a limited amount of data, it is nevertheless possible to generate some valuable insights into the VMA genre. The analysis shows that this new and innovative genre draws heavily on a range of existing types of scientific text types and interactions, including traditional written methods articles, lab demonstrations and conference presentations. The interview with Pritsker, JoVE’s CEO and co-founder, showed how the genre was conceived of as an extension of a traditional written methods article. The genre, which makes use of digital video with an accompanying spoken narration and opportunities for scientists to “speak to” their audience, is replete with hybridity as the VMA combines features of written articles with those of spoken scientific interactions as well as popular media text types.

Compared to written RAs, a number of innovations can be observed in the VMA genre. One such innovation is the way that the researcher intrudes into the VMA in the Researcher’s Introduction and Researcher’s Conclusion sections, often explicitly interacting with and engaging the audience, wishing viewers well in their research, thanking viewers and so on. These strategies are reminiscent of conference presentations, which place importance on the interactional metafunction: getting the talk off to a good start or ending on a warmer note to get question time off to a good start (Rowley-Jolivet, 2002). In one VMA, one even observes a wry sense of humour as a researcher concludes with the comment that:

While attempting this procedure it is important to remember that each of the sixty-four steps is critical and has to be performed with hundred per cent accuracy. (VMA6, narration, Researcher’s Conclusion)

This observation is interpreted as humorous by a viewer contributing to the comment section, who responds:

First I would like to say this latest method is really neat. I also like Julian’s comment at the end of the video when he said with a big smirk, “You have
to perform each of the 64 steps with 100% accuracy”. :D That is epic. (VMA6, comments).

Such a use of humour is a feature of popularisations (Giannoni, 2008) and also present in academic blogs (Luzón, 2013), where it is used to “construct solidarity and reinforce common assumptions” (page 448).

A second innovation – highlighted by specialist informants – is the reliance on the video medium to convey essential information about procedures, especially in the Demonstration section. This makes it possible to use a wide range (and a large quantity) of visuals, something that is not possible to the same extent in written RAs. Drawing on Rowley-Jolivet’s (2002, 2004) classification of images in scientific conference presentations, we see the use of: 1) figurative images (e.g. photos/film of the research setting, photos/film of raw data like X-rays or gel photos) to visually communicate procedures and results; 2) graphical images (e.g. charts, graphs, diagrams) to aggregate data and present concepts; 3) ‘scriptural’ images (e.g. section headings, flowcharts); and 4) ‘numerical’ images (e.g. mathematical script). Based on this exploratory study, further research is necessary to understand the precise nature of the visual elements of this genre.

The analysis of VMAs thus suggests some clear genre innovations in terms of the choice of semiotic resources used as well as how these resources are used, including the use of language. In particular, the analysis shows that discoursal identities not available in RAs are made possible by the video medium, with much greater engagement (realised through visual and linguistic strategies) unlikely to be found in traditional RAs (e.g. the use of “Let’s get started” to invite viewers in the Researcher’s Introduction). This finding contrasts with Pérez-Llantada’s (2013) study of Elsevier’s Article of the Future, the digitally enhanced RA. That study showed a similar lexical and grammatical profile for the Article of the Future and contemporary written RAs. More research is necessary to precisely identify the nature and extent of lexico-grammatical innovations in VMAs. Nevertheless, this exploratory study shows how, when digital media is recruited to meet the specific communicative needs of a discourse community, its use can lead to the development of an innovative, hybrid genre that breaks from convention in certain ways.

Commenting on the kinds of digital enhancements that are beginning to appear in RAs, Pérez-Llantada (2016: 31) notes that
This type of abstract is visually attractive and explicit at a glimpse, yet, as Sancho Guinda (2015) notes, it does not extend the information of the standard verbal abstract. The same applies to the interactive figures/tables and charts in the article text, which add very little to the article content.

In contrast, the visuals provided in VMAs make significant additional meanings possible. The difference here seems to be a genuine match between the purpose of the genre (precisely describing a protocol) and the affordances of the semiotic resources used, especially those of moving images. In VMAs, the use of digital media is not just an add-on: it is a crucial feature to deliver essential meanings which frequently get lost in a traditional methods article.

In common with some other digital research genres like academic blogs, VMAs tend to highlight the process of research, a feature that is of course by design, considering the purpose of the genre to share advances in scientific technologies and techniques. A move that is occasionally seen in the Researcher’s Introduction section and which would be unlikely to occur in written methods articles is Introducing Additional Researchers, typically laboratory technicians, research assistants, or graduate students, whose specialist expertise lies in conducting the experimental procedures. This move highlights the distributed expertise of scientific teams, a feature that is obscured in written RAs, which tend to suppress agency of individual authors and co-authors.

Unlike many studies of digital research genres, the design of this exploratory study explicitly focused on the dimension of time, by selecting one VMA per year for analysis from the inception of the journal in 2006 to 2016. By considering this dimension, the study was able to identify an initial exploratory period, during which proponents of the genre experimented with forms of expression before the genre became relatively stabilised. This is seen especially clearly in the development of the Overview section, which was not observed at all in initial videos and which developed a highly conventionalised structure within a relatively short period of time (in this corpus it was observed to emerge in 2008 and become highly conventionalised by 2010). This perspective neatly illustrates the way that recruiting digital media, while providing additional affordances for expression, can also provide challenges. It is a reminder that genres are not static but rather evolve in response to changes in the communicative context, including changes in medium.
In terms of its pedagogical implications, this study supports the notion that an important component of communicative competence in the academic context is the ability to understand and construct digital multimodal genres. Effective academic communication extends beyond the written mode to encompass the integration of various forms of visuals in both written and spoken texts. At the same time, the study also highlights the way that VMAs are collaboratively constructed, with experts in scientific communication (employed by the journal) facilitating the construction of the videos, especially more technically demanding aspects such as filming, editing, and the creation of diagrammatic animations. In spite of this support, those scientists who are able to visually design necessary elements of such VMAs are likely to be at an advantage in constructing this genre. Given the importance of visual design in academic communication, it would therefore be of great benefit to learners if courses in ESP embraced this aspect of academic communication. In order to do so effectively and systematically, ESP teachers need an understanding of the affordances of different modes for communicating different kinds of information and the way that these modes can interact. I hope that this article provides a useful step in this direction.

Acknowledgements

I am grateful to the two specialist informants for the time that they devoted to this study. I would also like to thank the editors for their valuable comments on drafts of the manuscript and for giving me the opportunity to contribute this article.

Article history:
Received 22 July 2018
Received in revised form 13 August 2018
Accepted 27 August 2018

References


Bhatia, V. K. (2004). Worlds of Written Discourse:


Christoph A. Hafner is Associate Professor in the Department of English, City University of Hong Kong. He has published widely in the areas of English for specific purposes, digital literacies, and language learning and technology. His latest book, English in the Disciplines: A multidimensional model for ESP course design (with Lindsay Miller) is published with Routledge.

Appendix: Corpus of VMAs


