

Multimodal communication and learning in on-line instruction videos

Andrew Jocuns, Thammasat University

Abstract

This study analyzes the multimodal communication patterns within on-line instruction videos on such social media platforms as Youtube and Instagram raising such questions as: what does it mean to learn on-line and how can multimodal communication and analysis be used as a tool for learning? First, the focus of the analysis is drawing attention to the different modes (language, gesture, space) that are used to perform instruction. Second, the different modes that are foregrounded and backgrounded in terms of mode saliency are analyzed in accordance with Norris' theory of modal density and multimodal interaction analysis. The role of mediated action – an individual social actor performing a mediated action with a mediational means – in instruction videos draws attention to the types of learning that take place when social actors engage in learning on-line. Findings indicate that the relationship between action, modes, and instructional texts forge instructional identities and/or expert status.

Keywords: learning on-line; mediated discourse analysis; multimodal discourse analysis; multimodal identity; on-line instruction

1. Introduction

This paper reports on a pilot study of instruction videos found on YouTube. The impetus for the study was to examine the interaction that takes place between the instructor(s) and their respective audience in an effort to answer the question: what role does multimodality play in learning on line? To that end the salient modes that make up these videos were analyzed as well as the conversational analytic structures such as openings and closing in addition to such interactional features, namely intersemiosis and resemioticization. I examined six videos from instructional videos for food. Three of these videos were on the same topic of healthy meal choices. The other three videos were chosen to analyze from the same YouTube host focusing upon how her presentation of self changed over time.

2. Vlogs and vlogging

There has been a good amount of recent research that has explored the phenomenon of digital literacies. For me, a good starting point for the basis of this research is the recent work by Jones and Hafner (Jones & Hafner, 2012), entitled *Understanding Digital Literacies*, which explores the various digital literacies that one can engage from a Facebook page to on-line gaming to video blogging, more popularly known as vlogging. Vlogging is the social practice in which an ordinary person can assume the role of a news reporter or an expert on something, shoot, edit, and post a video that can resemble a newscast, or in the present case, a cable cooking show. In many ways, the vlog has allowed the novice to enter a role of expert purely through the construction and posting of a video. This move to vlogging has prompted shifts in research as well for example Michael Wesch's digital ethnography, which includes an anthropological introduction to YouTube.

As Benson (2016) noted, YouTube pages employ the use of multiple semiotic modes from moving images, text, icons, clickable items such as a like button, as well as links to videos and pages off site. Benson (2016) noticed that there are two ways to think about how YouTube and sites are in fact, alike in texts. One way is to examine how the pages of such social media are organized in a way in which samples of different media are laid out in a text-like fashion. But for YouTube, the focal point of one's visit to YouTube is the video. Another way in which one can think of YouTube as a text is to consider it in terms of multimodal discourse analysis and interaction. However, Benson (Beard, 2004; Benson, 2016) added that a tendency has been

to focus more so on the written literacy qualities of such texts. Benson stated three characteristics of YouTube videos: 1) they deploy multiple semiotic modes, 2) YouTube pages are sites for multiple authorship where much of the texts on the page is generated by user comments, and 3) YouTube pages are highly dynamic in that they are constantly changing in terms of how they oriented between user texts and other generated content.

One of the points that Benson (2015) made is that videos on YouTube, especially his focus upon language and intercultural learning, follow a similar pattern to the initiation-response-follow-up, IRF, framework often found in classroom interaction, which has been extended to use in everyday English conversation. This is quite similar to Mehan's (1979, 1998) finding of the structure of classroom interaction around the initiation-response-evaluation, IRE, framework. In this structure the teacher, or in this class the vlogger, initiates an interaction in the classroom usually through a question. The question/initiation is responded to and then an evaluation or follow-up is made by the instructor.

Some features of the multimodal framework of YouTube instructional videos that are important in the present context is that, as Benson noted, they can be analyzed using the framework for conversational interaction. For my purposes, I am interested just in openings in instructional cooking videos. In part because of context collapse, we have an assumed audience as opposed to a known audience. These assumptions are based on the multimodal nature of instructional videos. For example, intonation and pitch in voice quality may index a reference to an audience that is a specific gender. High pitch, raising intonation and vocal fry may index an audience that is female. In addition, how the vlogger addresses the audience may also index a particular gender identity of an audience. One example of this is from the YouTube page for The Domestic Geek who usually addresses her audience with the greeting "hello my lovelies!" An informal survey from my students suggested that the audience would be women, teenagers, or a group of people younger than The Domestic Geek.

The Domestic Geek uses a number of semiotic resources in the delivery of her instructional vlog. These include: spoken language in an informal style; kinetic modes which include gaze, gesture, posture, and head movement; and cinematography to include camera angle and frame size. Her videos are very professionally done and it appears that she is in a studio kitchen. In addition, The Domestic Geek uses a variety of overlexicalization in the form of evaluation (Machin & Mayr, 2012). This is exemplified in her video for 5 Healthy MEAL PREP Ideas. Table 1 notes this practice of evaluation and overlexicalization.

Table 1: Overlexicalization and evaluation

Data
1. And this deliciousness is ready to be enjoyed.
2. This box is such a gorgeous combination of tasty flavors, textures and colors.
3. Uber good!
4. This really flavorful , colorful salad that you will just devour.
5. You just start with a nice hot frying plan.
6. What can be easier than that ?
7. You just need to mix all this goodness up.
8. This meal prep has beautiful , complementary flavors that I think you guys are going to love.

Table 1, cont.

9. And I honestly say, if you've never made Chinese fried quinoa, it'll seriously change your life .
10. I've got some fun Mexican inspired flavors for you.
11. Once all that goodness is in the pan.
12. This beautifully spiced chicken mixture is really well complemented .
13. Guys this is a really tasty way to get a ton of nutrition with a ton flavor.
14. I top that with some gorgeous steamed vegetables.
15. This incredible lemon dill salmon.
16. Is really beautifully complemented by those gorgeous green veggies, and of course that
17. Amazing that lemon dill salmon.
18. As far as I'm concerned, this is heart healthy lunch at it's finest

The Domestic Geek's text in bold in table 1 are references to the overlexicalization practices that take place during her instructional videos on cooking. The specific video is on healthy recipes, and through using the mode of spoken language and overlexicalizing the evaluation in her talk, she is able to construct herself as someone who is an expert at healthy meal prep.

YouTube instructional videos vary in their scope as well as interactional sequences; not all vloggers are interactive in the same way. In fact, for some, this is a process that evolves over time. To that end, I note the vlogger Maangchi, a Korean woman living in the United States who offers a variety of recipes for cooking a variety of Korean cuisine. Her earliest vlog on her YouTube page was uploaded in 2007, and to illustrate the point of how interactivity within vlogging can evolve over time, I choose 3 videos to compare 3 different timescales along her vlog. If we examine her interactivity in 2007, we note that she was very stoic. In 2011, a middle time scale for her vlog, we already begin to see her taking on more of a performative stance. By 2017, she was a very comfortable vlogger who not only had excellent rapport with her audience, but also had the stance that she takes towards her vlogging. Earlier, her language was not that nuanced, however, later, she used more complex linguistic practices such as anthropomorphization and reduplication.

Table 2: Maangchi's language shift over-time

2007	2011	2017
<ul style="list-style-type: none"> • English • Informal • Cooking terms 	<ul style="list-style-type: none"> • English • Informal • Cooking terms • Anthropomorphization • Intonation • Describing words 	<ul style="list-style-type: none"> • English • Informal • Cooking terms • Anthropomorphization of things • Reduplication • Describing words

Table 2, cont.

Keep stirring rice	With this guy (holding up a pumpkin	First I like to give these guys a really good shower
		Crazy crazy hot hot

Lastly, it is worthy to note how Maangchi’s shifts are not only on her language, but other modes are also manipulated in her vlog, namely camera angle shifts that focus on her and shifts to point of view angle. These shifts are noted in Table 3.

Table 3: Maangchi’s camera angle shift over-time

2007	2011	2017
		
		
		

3. Conclusion

YouTube has created a context in which vloggers can have an interactive stance to an audience, which is context collapsed, that is to say the audience is unknown. Through the manipulation of different modes as noted in the vlogs of The Domestic Geek and Maangchi, we note how vloggers can emerge over time as competent instructors. The stances that they take up to their audience and what they are teaching index how they use modes to display identities as experts. The next step in the process of my analysis of such instructional videos is look at the other end of the vlogs, focusing upon its interactive features and what people learn from such interaction.

References

Beard, A. 2004. *Language change*. London: New York: Routledge. Available at <http://site.ebrary.com/id/10099964>.

Benson, P. 2015. *YouTube as text: spoken interaction analysis and digital discourse*. New York: Routledge, 81–96.

Benson, P. 2016. *The discourse of youtube: multimodal text in a global context*. New York: Routledge.

Healthy MEAL PREP ideas. Available at <https://youtube.com/SffWUBxXQ9E>. (Accessed on June 5, 2018).

Jones, R. H., & Hafner, C. A. 2012. *Understanding digital literacies: a practical introduction*. New York: Routledge.

Maangchi. Available at <https://www.youtube.com/user/Maangchi>. (Accessed June 5, 2018).

Machin, D., & Mayr, A. 2012. *How to do critical discourse analysis: a multimodal introduction*. Los Angeles: SAGE.

- Mehan, H. 1979. *Learning lessons: social organization in the classroom*. Cambridge: Harvard University Press.
- Mehan, H. 1998. The study of social interaction in educational settings: accomplishments and unresolved issues. *Human Development*, 41, 245–269.
- The Domestic Geek*. Available at https://www.youtube.com/channel/UCQQQ3RxX_o-B-68wSKdcfMQ. (Accessed on June 5, 2018).