Analyzing layering in textual design: a multimodal approach for examining cultural, linguistic, and social migrations in digital video

Myrrh Domingo

Steinhardt School of Culture, Education, and Human Development, New York University, New York, USA

Online publication date: 11 May 2011
Analyzing layering in textual design: a multimodal approach for examining cultural, linguistic, and social migrations in digital video

Myrrh Domingo*

Steinhardt School of Culture, Education, and Human Development, New York University, New York, USA

(Received 1 September 2010; final version received 22 March 2011)

This paper explores the nature of layering noisy and moving features in video to better understand how the movement of people, their ideas, and their texts are shifting functions of design and circulation within digital contexts. Drawing from video collected as part of an ethnography of digital literacies in urban contexts, features of textual design in a music video will be examined to demonstrate how youth today produce and circulate digital products as an expression of their diverse social identities. I sketch an approach for analyzing design as embedded in the layering of image, language, gesture, color, visual, and sound effects in a participant’s music video.

Keywords: literacy; language; text; migration; digital video design; multimodal analysis

Introduction

As made visible by the prolific use of videos, mobile devices, computers, and other technological tools in everyday settings, participating socially in digital contexts is no longer exceptional but part of daily urban life in most countries (Appadurai, 1996; Hull, 2003; Kellner, 2001). Given this seemingly ubiquitous trend in vast digital and global migration, participation within digital contexts has been studied in terms of expanding what counts as texts and how such texts are circulated both locally and globally (Kirkland, 2008; Lam & Rosario-Ramos, 2009; Lankshear & Knobel, 2003). Other studies have explored how digital connectivity opens spaces for exploring and expressing social identities and cultural knowledge (Alvermann, 2008; Lewis & Fabos, 2005; West, 2008). There is shared recognition that while paper-based texts are necessary, it is no longer adequate for communicating across the various digital platforms in our global world (Andrews, Borg, Boyd, Domingo, & England, forthcoming; Domingo, 2010; Kress, 2010). As Suarez-Orozco (2007) writes, ‘New global realities increasingly define the contexts in which youth growing up live, learn, love, and work … Theirs is a world in flux where the rate of change is of an order never seen before’ (p. 11). The ways in which people experience belonging are increasingly destabilized as fixed categorical memberships are disrupted by the constant flux of people, their ideas, and their texts across digitally and physically mediated boundaries (Banks, 2008; Bauman, 1998).

*Email: myrrh@nyu.edu
Speaking directly to this concern, this paper offers an analytical approach for digital video that can be applied to arrive at salient migration patterns that consider how youth today, their ideas, and their texts circulate given their prolific participation in digital textual production. I assert that the design and circulation of video texts permitted the Filipino British youth I studied (who call themselves the ‘Pinoys’) to express their diverse social identities as an inherent feature of their digital practices.

**Textual design and dialogics in digital video**

In our digitally connected world, textual products abound that display new configurations for interactive meaning-making. These digital texts include music videos, websites, and blogs that permit layering of modes such as sounds, images, colors, among other moving and noisy features of design. This layering expands what counts as knowledge to include non-linear texts (Jewitt, 2002; Mills, 2009; Pahl, 2007). Such an evolved notion of text includes more dialogic participation in which new media technologies and the interactive nature of online platforms allow youth to reproduce, remix, and reconstruct texts by engaging literacy practices that resemble ‘bricolage’ and ‘collage’ rather than traditional linear print (Williams, 2009, p. 8). For Williams, these participatory reading and composing opportunities provide youth with rhetorical devices for both expressing identity and building community in digital contexts.

Resonating with Bakhtin’s (1981) dialogics, this shift in textual design practices more readily takes into account that people’s relation with language is a pliable art form. Images, sounds, among other features once considered ancillary to written language, now comprise layering capabilities for designing video texts that display the inextricable interrelationships among modes such as language, oral and written; images, still and moving; and sound, voice and music (Cope & Kalantzis, 2000; Jewitt, 2008; Kress & van Leeuwen, 2001; New London Group, 1996). Layering of modes as a feature of digital video design enables new configurations of data to circulate that express meaning in ways not previously possible with only page-bound texts (Hull & Nelson, 2009; Kress, 2003). For example, Wingstedt, Brändström, and Berg (2010) write about meaning in narrative media music – music used for narrative purposes in film, computer games, television programs, etc. – as achieved through the interplay of music and image, whereby vision is guided by the auditory sense and vice versa. The authors posit that audiences are able to ‘read’ the filmic narrative because the interactivity among modes, such as music with visuals, dialogue, sound effects, actively contributes to the telling of the multimodal story (Wingstedt et al., 2010, p. 197). Further, they recognize that while specific meanings may be realized differently for individuals, established social and cultural conventions more concretely shape reception of the narrative filmic experience. Aligned with Howes’ (2009) discussion of recent research in cognitive science and sensory anthropology, the networked relation among senses disrupts the conventional sensory perception as having disparate purposes (e.g., sight is concerned with visuals or hearing is for sound); instead, the interactivity among them is defined as a form of adaptive processing. He identifies these networked and ‘conjugated’ (combined or separated) relations among senses as a cultivated practice, whereby the degrees to which different sensory fields are elaborated vary by culture (p. 225). Similarly, Kress (2010) articulates this human engagement with the social and semiotic world as culturally variable; as he writes, ‘What may be done by speech in one culture may be handled by gesture in another; what may be well done in image in one culture may be better done in 3D
forms in another’ (p. 81, emphasis in original). He describes the ‘reach’ of modes as contextually specific, thus it cannot be assumed that modes carry the same translation across cultures (p. 83).

Given that layering of moving and noisy features in digital video design provides social and cultural insights, it is problematic to analyze recorded speech, image, music, and gestures, among other modes, in isolation from one another. Further, accounting for configurations of modes without considering their social and cultural reach would also yield liminal understanding of the cross-cultural textual design practices embedded within the digital video. In contrast, viewing digital texts with a dialogical frame opens possibilities for layering of modes not only as a video resource but also as a medium for expressing one’s diverse social identities and linguistic repertoire in digital communities. As Lemke (2009) states, ‘In some very basic sense the use or function of every media work is not just to link a producer and a user, but to link across the timescales of production, circulation, and use’ (p. 143). Aligned with Bakhtin’s assertion that all utterances are not encapsulated in a vacuum but reside in living interaction with the social world (Landay, 2004), the moving and noisy features of video enable designers of digital texts to manipulate language as a pliable art form. It is through this artistic reworking – the tension-filled practice of listening to the social cadence of language and interjecting one’s own sound to its historical orchestration – where I locate the opportunities for youth to express their diverse social identities as an inherent feature of their digital video design.

An ethnographic study of the ‘Pinoys’

The digital video data to be analyzed are among a collection of artifacts from my ethnography of Filipino British youth in London. I have spent the past three years studying the Pinoys and their distinctive ways of practicing languages and literacies across social spaces. The research questions informing the ethnography are: How do the Pinoys assert their social identities as they migrate across physical and online spaces? How do they engage in literacies across spaces navigable by the body and mediated online? In what ways do their migrations across physical and imagined spaces – ideas and identities, texts and contexts, cultures and communities – extend their language and literacy practices? Because the Pinoys constantly traversed social contexts that included both physical and digital spaces, it became problematic to study their literacies as only rooted in place-based structures. Thus, data collection and analysis methods required that I also partake in migratory practices. My fieldwork traced the movement of my participants and their ideas across contexts – performance halls, homes, online social networks, among others – to explore the ways in which their languages and literacies migrated across social spaces. I gathered the cultural artifacts that the Pinoys produced and noted how they circulated these dynamic texts. To account for the vast movement of the Pinoys, their ideas, and their texts, the study involved ethnographic field notes, semi-structured interviews as well as online discussion threads, videos, and profiles of participants from sites such as MySpace and Facebook. In addition, analytic memos to track emergent patterns and theoretical assumptions during each field visit were also maintained (Dyson & Genishi, 2005).

In light of the Internet, the ‘field’ in research is no longer merely identifiable as physically bound but now also includes digital communities navigable online. Ethnographers conducting fieldwork in virtual sites can remain physically invisible without the use of video chats and still manage to collect data (Denzin, 1999; Schaap, 2002).
However, Murthy (2008) points to a balanced combination of physical and digital ethnography for providing researchers with a broader array of methods, but also enabling them to ‘demarginalize the voice of respondents’ (p. 837). Such an approach was taken with the large-scale Digital Youth Project by the MacArthur Foundation, which examined the ways in which digital media are shifting the learning of young people as they socialize and participate in civic life across varied social contexts (Ito et al., 2008). Similarly, I understand the digital video production of the Pinoys not as isolated events but as an intricate process of meaning-making that unfolded over time, across social spaces, and as embedded in the lived realities of my participants. This ethnography therefore focused not only on the finished digital video products but the entire process of the Pinoys’ generative practices of textual design.

The research necessitated working with participants as they interacted with both local and global audiences. The Pinoys featured in this study associate with one another as members of a hip hop production group that includes affiliation with members all over the world. The group is drawn together by their affinity for their Filipino heritage and hip hop as a platform for expressing their cultural knowledge and social identities. Recruitment of the Pinoys involved reaching out to various community centers in England to gain access into working with youth whose reading and writing practices encompass a digital, multimodal, and socially networked dimension. A member of a Philippine Community Center responded to my visitation request and introduced me to several youth groups. The Pinoys were among these groups, and the five participants voluntarily opted to partake in this study after they were informed of the project statement and research purpose.

Given the scope of the research, it was necessary to work with participants who are ardent readers and writers, avid users of digital technologies, and prolific producers of multimodal texts. Kyd exhibited these literacies in his everyday practices. Of all the participants, Kyd is the most versed in designing video texts and he prominently shaped the layering configurations during the video-editing process. To this end, examining features of Kyd’s textual layering lends insight into the ways in which he actively manipulated language as a means for expressing social identities and affirming cultural affiliations as a video design feature. The hip hop video to be analyzed is among the numerous multimodal texts that Kyd designed to assert his social identities to a wider audience.

Analyzing layers in video data
To display how participation in hip hop culture and prolific use of digital video extended the everyday functions of video design and circulation within digital communities, six modes were purposely accounted for in Kyd’s music video: landscape, gestures, language, visual effects, sound effects, and color. Further, the interrelations among these modes were also examined. As evidenced by previous conversations, observations, and interviews, Kyd purposefully layered these modes when designing his video texts. Choosing to focus only on two or three of the modes would have disentangled the linked meaning of Kyd’s dynamic textual design. Whereby linear and paper-bound text often privileges the visual sense, digital video saturates multiple senses and necessitates their collaborative activity. This interaction among senses promotes a layered reading of the spatial and temporal configurations inherent in noisy and moving texts. Thus, I account for the ways in which he layered modes throughout my analysis of his music video, aptly titled in Taglish¹ as ‘Flow
While my research required close examination of these six modes, others may opt to explore stratal couplings (Kress & van Leeuwen, 2001) or combinations with regard to their rationale and with consideration to the data to be explored. Below, I outline my approach for analyzing Kyd’s video text.

**Developing a unifying frame**

The transcription frame allowed for both a linear/temporal and a layered/spatial reading of the video data. Each transcription frame created for the music video comprised three logics of organization. First, a **title bar** was assigned to the clip (title of the video), theme (main topic), and segment (time frame) of study. Second, the **body frame** was filled out, which accounts for each mode utilized in the particular segment and displays still shots of images from the video. Third, a **narrative description** was written about the segment that links the transcription with related reflexive notes, interviews, and observations. This last step proved critical for linking each segment of video analysis to the overall textual product, as well as to previously collected data. The multimodal analyses key, which identifies each mode by color, was used not only in the body frame but also in the narrative description.

As made visible in my transcription frame, the total playback time of this particular video segment is three seconds out of the total running time of two minutes and 10 seconds. If played in real time, keeping pace with the dynamic medium of Kyd’s music video fails to reveal the compressed complexity of his design. Goldman and McDermott (2007) refer to time in video analysis as more ‘reticular and reflexive’ rather than a simplistic reduction of then and now (p. 112). In their study of a middle-school math curriculum, multiple viewings and slow-down options in their video records ‘added layers of complexity’ that opened new ways for their analysis to evolve beyond initial understanding to more nuanced exploration of the classroom experience as embedded in a social structure (Goldman, Knudsen, & Latvala, 1998). In other words, applying a linear reading path to understand video segments would not enable an in-depth account of the spatial configurations displayed within the short time span. For example, the various modes that Kyd employed in designing this video text, and how these modes are interconnected, cannot be analyzed without a layered reading approach. Among the key findings made visible by reading this frame spatially is how Kyd merges written and spoken language by using his body as a canvas and his gestures as inscription tools. In the video, when he begins to perform the first two words of the lyrics, ‘Lirikong Supremo,’ he opens his jacket like a curtain and gazes directly at these two words visually depicted on his self-designed shirt. As a viewer, Kyd instructs my eyes to read the inscription of his gestures as he lyrically echoes the writing to my ears. He synchronizes his spoken and written language, whereby the words uttered echo the visual text.

Visually, the multimodal transcription frame required rows and columns that could flexibly accommodate the simultaneous and co-dependent nature of the modes (see Figure 1). Again, this was achieved simply by adding color. For example, to account for Kyd’s coupling of sound effects and visual effects, the corresponding graphic display incorporated both the colors, orange and pink (color codes used to represent sound effects and visual effects), to demonstrate the inseparability of the two modes. Sometimes this graphic representation was achieved simply by adding a background fill and other times fonts were depicted in respective colors. An alternative solution might be assigning symbols to each mode examined instead of color-coding the
Identifying discourse communities

Among the patterns distinguished early in this study is how the participants used layering in video texts as a form of migratory practice. For example, Kyd used the six modes previously discussed to identify and navigate his belonging in four discourse communities: Filipino, British, hip hop, and youth pop culture. For the Pinoys, participation in these discourse communities is not an isolating but rather an overlapping experience with the use of digital video. While it is not possible to display how belonging in each discourse community is constantly shifting using still image, it is nonetheless helpful to view how a new social space opens given their distinctive migratory practices within these communities (see Figure 2).

Distinguishing layering functions

Having identified the discourse communities in which the participants’ belonged, the next step was to track their linguistic patterns and social practices within these spaces. The research involved studying the Pinoys as they navigated physical and digital social spaces; however, this paper will focus only on Kyd’s discourse practices within digital communities. In light of layering capabilities in digital video, Kyd’s design approach extended Bakhtin’s conceptualization of language as a living artistic
resource. Not only were the written and spoken words in the music video figuratively brought to life but Kyd also personified the language by interactively layering modes. According to Bakhtin (1981):

Language is not a neutral medium that passes freely and easily into the private property of speaker’s intentions; it is populated – overpopulated – with the intentions of others. Expropriating it, forcing it to submit to one’s own intentions and accents, is a difficult and complicated process. (p. 294)

For Kyd, layering in digital video functioned not as an additive form of meaning-making (Hull & Nelson, 2005; Jewitt, 2002); rather, Kyd layered modes to deftly navigate his diverse social and cultural connections (see Figure 3). This dialogic discourse practice functioned as a pliable art form for Kyd to artistically remake ideas, thoughts,
and texts into his own voice. As this interactive view of language and literacy considers how meaning-making processes are influenced by social realities, it offers insight into reading and writing practices that do not yield a finite production of cultural artifacts but become part of ‘interanimating relationship with new contexts’ (Bakhtin, p. 346).

**Representing dynamic categories**

Given that Kyd’s digital video data is a moving and noisy text form, mining for its layered meanings also required searching for a dynamic way to represent the emergent categories. As made visible by the sample multimodal transcription frame (see Figure 1), assigning codes was most challenging when working with this type of interactive data. The dialogics of Kyd’s textual design cannot be neatly and plainly categorized as the modes are inextricably linked in his artistic process. Whereas images, sound, and other modes were once considered additive or ancillary to written text, layering in digital video enabled Kyd to display meaning without having to sever the linkages among the modes.

Data were color-coded to represent the dynamic text in a form that will further illuminate the relations among modes, and to analyze more deeply the layered textual configuration. This analytic phase enabled the logical patterning behind Kyd’s video text to be made more discernable. By color-coding the data, it was possible to display Kyd’s use of each mode without isolating them during the transcription process. For example, the sound effect of a bass-heavy beat is synchronized with the visual effect of the screen shaking (see Figure 1). With each shake of the screen, the camera panned in or panned out to reveal a different landscape. The same pattern was repeated throughout the video, which illustrates how Kyd coupled image and sound as inextricably linked in his textual design.

**Analysis of ‘Flow Ko’**

Kyd’s music video ‘Flow Ko’ is two minutes and 10 seconds in length. It evolved from a live performance to a recorded rap song, which was later remixed to a different beat and eventually made into a music video. It consists of footage from three videos compressed into one, and rhythmically orchestrates its constituent parts. A merging of poetic lyricism with hip hop and moving image, Kyd juxtaposed words, gestures, sounds, and color throughout his video text. Situated within three local London landscapes, Kyd is depicted in the video as traversing the streets armed with his supreme lyrical ‘skillz’ in rhyme and rhythm. As if heading out into a battle, the ‘intro’ presented Kyd leaving the home base when he stepped out of a London flat with his friend Aziatik. ‘Ingat,’ which means ‘be careful’ in translation, they greeted one another before parting in opposite directions. Armored in hip hop gear, Kyd and Aziatik wore an array of black and white clothing inscribed with their group name ‘Lirikong Supremo.’ The only other inscription was on a lanyard they both wore that read, ‘Pilipinas.’ Taken together ‘Lyrical Supremacy’ and ‘Philippines,’ in translation from Tagalog to English, represent their social and cultural affiliations to hip hop and their native country. The video verbally and visually portrayed his battle verses, lyrics composed that aim to challenge opposing hip hop artists (Alim, 2006; Smitherman, 1997). Aligned with Bakthin’s dialogics, Kyd’s battle verses are in direct response to what has been said before and in anticipation of others’ response to his music. For Kyd, Aziatik, and the other participants in my study, forming Lirikong Supremo was
a way to engage Filipino British youth to learn about their heritage, Philippine culture, history, and current events, preserving their language and cultural ties to Philippines. They wanted to create a space where their social identities as Filipino, British, hip hop artists, and designers of digital music videos can merge and evolve.

Analyzing the layers of the video using the framework outlined above shows how living words as cross-cultural interactivity are layered and how layering in digital video productions can represent a form of linguistic and social migration. The video design *layers living words as cross-cultural interactivity* using a number of multimodal resources – the alignment of spoken rhythm to the beat of the music in atypical patterns, a synchronized flow by merging vision and sound, for example, coordinating gaze and hand gestures with the drop of the beat or a significant change in rhythm, or shaking of the screen to mimic the visual effect of an earthquake with the heavy bass of the music. Kyd’s layering of beats, rhyme, and visual cues resonates with Howes’ (2009) assertion of interrelations among senses as a form of adaptive processing. Kyd used modes as rhythmic resources for his audience to stay connected to the narrative that unfolds in his music. The overall piece also extends Bakhtin’s notion of language as contradiction-ridden given Kyd’s display of words as a living and evolving art form in this video.

This multi-sensory video text also reflects a Bakhtinian perspective as it illustrates Kyd’s engagement of languages and literacies as historically fraught with social and cultural narratives. Smitherman (1997) writes about hip hop as a discourse rooted in Black resistance culture. While Kyd’s identifies his cultural affiliations as rooted in being Filipino and being British, participating in hip hop discourse for Kyd nonetheless still functions as a form of resistance culture.

**Layering as a form of linguistic and social migration**

The textual design in Kyd’s ‘Flow Ko’ video visually displayed his migration within the discourse communities of Filipino, British, youth pop culture, and hip hop through his avid use of color. Thus, colors functioned as a mode in Kyd’s digital video by communicating a cultural discourse (Kress & van Leeuwen, 2001, p. 57). His selection of colors signals his merging of his two cultural memberships. The colors red, white, blue, and yellow are emblematic of his cultural heritage as it represents the colors of the Philippine flag. With the exception of the color yellow, the same colors comprise the British flag. His palette choice functioned as a form of migratory practice as the colors integrated indicated his belonging and movement within the varied discourse communities of Filipino, British, youth pop culture, and hip hop.

In addition, Kyd’s video design depicted a migration by displaying a range of backgrounds, from the streets of London to only a black backdrop. Kyd described his use of landscapes in video as a movement from the city of London into his inner mind. Migrating across different landscapes, as a layered design element in video, enabled audiences to travel vicariously with Kyd and to explore the various discourse communities in which he belonged. A dialogic practice is thus apparent in digital videos produced for public circulation, whereby layering of modes functioned as a living artistic resource for expressing social and cultural identities within digital communities.

**Conclusion**

Continued examination of the interactivity in youth’s prolific digital textual production has potential for contributing to evolved notions of texts and hybrid textual
designs that promote cultural knowledge and linguistic practices. Such insights are increasingly necessary given the vast movement of people, ideas, and texts in our global and digital world. By shifting the study of literacies from bounded spaces to examining digital migratory patterns, this paper aimed to display how digital video can be analyzed to understand how youth today cultivate a global voice and digital dexterity to participate as designers of cultural texts. As evidenced by the analysis of ‘Flow Ko,’ briefly discussed in this paper, attending to the layering of modes in video texts makes visible how *dialogics* (Bakhtin, 1981) in digital video design functioned as a linguistic passport for youth to migrate within digital communities.

Aligned with the notion of *timescales* (Lemke, 2010), the multimodal analysis was designed to understand the meaning-making behind the digital video as embedded in the design and circulation process. Because layering of features in digital video design lends social and cultural insights, it was problematic to analyze recorded speech, image, music, and gestures, among other modes, in isolation from one another. Thus, the analytic approach was vested to arrive not only at a comprehensive coding practice of temporal and spatial configurations but to also represent the analytic codes in a way that echoes the dynamic nature of a video text. The multimodal transcription frame enabled both a linear and layered reading of the video data, revealing relational patterns among modes. While connecting the codes in this video to other data collected balanced the micro-approach for analyzing layering, further inquiry into multi-sensory approaches for handling a larger collection of moving and noisy data would offer insights to expand the micro-analysis presented. Lastly, though this paper displayed the inextricable linkages among modes using multimodal techniques, there is a need for developing a transcription frame that can more flexibly handle interactive texts that refuse to lay flat on the page.

**Note**

1. Kyd sometimes combines Tagalog and English when speaking, rapping, or writing. The title of the music video to be analyzed is a prime example of his hybrid linguistic mixing. In translation, ‘Flow Ko’ means my flow.

**Notes on contributor**

Myrrh Domingo is a doctoral candidate in English Education at New York University. Her work explores literacy development and multimodal textual production in the context of digital technologies and global migration. She is the recipient of the National Academy of Education Adolescent Literacy Fellowship from the Carnegie Corporation of New York.

**References**


