

Locating the Semiotic Power of Multimodality

GLYNDA A. HULL
MARK EVAN NELSON
University of California, Berkeley

This article reports research that attempts to characterize what is powerful about digital multimodal texts. Building from recent theoretical work on understanding the workings and implications of multimodal communication, the authors call for a continuing empirical investigation into the roles that digital multimodal texts play in real-world contexts, and they offer one example of how such investigations might be approached. Drawing on data from the practice of multimedia digital storytelling, specifically a piece titled "Lyfe-N-Rhyme," created by Oakland, California, artist Randy Young (accessible at <http://www.oaklanddusty.org/videos.php>), the authors detail the method and results of a fine-grained multimodal analysis, revealing semiotic relationships between and among different, copresent modes. It is in these relationships, the authors argue, that the expressive power of multimodality resides.

Keywords: *multimodal texts; semiotics; multiliteracies; New London Group; multimedia; digital stories*

All about us, there are unmistakable signs that what counts as a text and what constitutes reading and writing are changing—indeed, have already changed and radically so—in this age of digitally afforded multimodality. To rehearse the obvious, it is possible now to easily integrate words with images, sound, music, and movement to create digital artifacts that do not necessarily privilege linguistic forms of signification but rather that draw on a variety of modalities—speech, writing, image, gesture, and sound—to create different

Authors' Note: We appreciate the opportunity to develop an article for this special issue honoring Steve Witte, a friend and colleague whose keen intellect shed the light that opened paths for the rest to follow. So it is with this article, and the tradition of work that it represents, growing as it does from Steve's early example and his vision. Our sincere thanks to Randy Young, the multimodal composer whose work we featured in this article, and to Mike Rose, Chris Haas, and Mira-Lisa Katz for their helpful commentary.

WRITTEN COMMUNICATION, Vol. 22 No. 2, April 2005 224-261
DOI: 10.1177/0741088304274170
© 2005 Sage Publications

forms of meaning. There are now Web-based scholarly journals that illustrate and explore these possibilities (e.g., *Kairos* and *Born Magazine*¹), there are community-based media organizations that promote a variety of forms of multimodal composing (cf. Lambert, 2002), and there are beginning to be empirical studies that examine multimodal practices in context (Stein, 2004). And as we will shortly review, of late, helpful theorizing about multimodality has begun (Kress, 2003). Some scholars, it is true, recognized the advent and importance of multimodality as an aspect of literacy a long time ago, taking heed, for example, of the importance of multiple forms of representation (Witte, 1992, 1993). Yet the full import of this sea change in semiotic systems has, for most people, just begun to be felt.

In this article, we want to pay homage to the range of work that is beginning to explore new literacies or multiliteracies (New London Group, 1996).² We hope also to extend this work by making and supporting a still radical claim. We will argue that multimodal composing of the sort to be illustrated here is not simply an additive art whereby images, words, and music, by virtue of being juxtaposed, increase the meaning-making potential of a text. Rather, we plan to demonstrate that through a process of braiding (Mitchell, 2004) or orchestration (Kress & van Leeuwen, 2001), a multimodal text can create a different system of signification, one that transcends the collective contribution of its constituent parts. More simply put, multimodality can afford, not just a new way to make meaning, but a different kind of meaning. Thus, in this article, within the constraints and affordances of a primarily linguistic text, we aim to illustrate and offer an initial framework for analyzing a particular and increasingly popular form of multimodality.

Despite growing interest in, research about, and examples of multimodality, we feel a certain urgency about our project and like-minded work. It is no exaggeration to say that most Western societies remain print dominated, even as pictures push words off the page and even as the Internet and the World Wide Web become virtually ubiquitous. This is especially true of schools and universities, which are staunchly logocentric, book centered, and essay driven, invested as are most educators in the versions of meaning making whose value they know best and committed as are many educators to sharing the languages and modes of power (cf. Delpit, 1995). Several recent and important policy documents illustrate how easy it is to exclude digital multimodality when addressing issues of literacy and literature and

how natural it is to think in terms of print-based, unimodal texts. The National Commission on Writing for America's Families, Schools and Colleges (2003) recently produced a booklet rightly calling on all to recognize anew the importance of writing yet did not make mention of multimodality as a potential type of digital composing and gave just a nod to the mediation of writing via computer technologies. During 2004, a national adult literacy survey was administered in the United States in which participants were asked to perform everyday literacy and numeracy tasks—to fill out forms, to read a paragraph, to calculate sums, and so forth—yet these tasks excluded digital texts, the Internet, and even computers (National Center for Education Statistics, 2003). And recently, the National Endowment for the Arts (NEA, 2004) released the results of a study entitled "Reading at Risk," which decried the significant decline in the consumption of literature among Americans during the past two decades, all within the framework of a traditional canon. The report implied that Americans who are less likely to be avid readers of literature are more likely to engage with other media—television, video, and the Internet—to their detriment and to that of an educated citizenry.³

The value of the NEA (2004) report, and other similar calls to action around valued forms of literacy, is that it underscores the importance of developing and maintaining the literacy practices of print-based reading and writing, which, notwithstanding the proliferation of other new media technologies, remain paramount for individual growth and meaningful participation in the broader society. However, given the range of semiotic tools available for literate practices at this particular historical moment, we find it worrisome to exclude the new forms of reading and composing from mention, and it concerns us as well to assume a hierarchy of value among them. We believe that the increasingly multiplex ways by which people can make meaning in the world, both productively and receptively, can potentially represent a democratizing force whereby the views and values of more people than ever before can be incorporated into the ever-changing design of our world. We hope to suggest, then, that the new media that afford multimodal composing might helpfully be viewed not as a threat to or impoverishment of the print-based canon or traditional means of composing, but rather as an opportunity to contribute a newly invigorated literate tradition and to enrich our available means of signification.

RELATED LITERATURE

A burgeoning body of theory and research has broadly addressed systems of signification other than the verbal and has explored the interplay between systems, especially visual images and print. The field of visual culture, for example, combines critical and social theory with an analysis of all types of visual media—painting and art, to be sure, the traditional focus of art theory and practice, but including as well television, photography, advertising, and architecture (cf. Mirzoeff, 1999). Visual methods have made their way as well into the social sciences through approaches such as visual anthropology and visual sociology and an ever-increasing interest in documentary research across a range of fields (cf. Coles, 1997; Stanczak, 2004). Literacy studies, on the other hand, have until recently been positioned somewhat peripherally as far as visual things go, mostly eschewing the pictorial in favor of the verbal. This, of course, is notwithstanding multitudinous examples, historical and modern, of the inclusion of the visual as part of the written: Illuminated manuscripts, such as the *Book of Kells*; illustrated story books for children; experimental novels; scientific treatises that depend heavily on detailed illustration; and everyday workplace documents that juxtapose diagrams, drawings, pictures, and words are a few examples that come to mind (cf. Finnegan, 2002). Yet the commonplace assumption has been, especially in school and university settings, that the affordances of written verbal texts far outstrip what can be offered by or offered in conjunction with other modalities. As we will shortly discuss, such conceptions have begun to shift, and quickly so, in the field of literacy studies, in large part because of the concept of multiliteracies, an idea itself influenced by the increasing prevalence of digital technologies and their potential as a mediational means.

Just as it's possible to look back historically or around us at the current moment and find various examples of the integration of the verbal and the visual within a single text, it is also important to recognize that multimodality too has ancient and deep roots in cultural practices the world over (despite what some would view as its neglect in the West), and that multimodality is in fact what distinguishes human communication. Anthropologist Ruth Finnegan (2002) terms communicating "a multiplex and versatile process" and describes humans' communicative resources as encompassing

their powers of eye and ear and movement, their embodied interactions in and with the external environment, their capacities to interconnect along auditory, visual, tactile and perhaps olfactory modalities, and their ability to create and manipulate objects in the world. (p. 243)

It is certainly the case that educators regularly rediscover the power that students experience when released to communicate and learn multimodally. Stein (2004), in fact, on the basis of her work as a language educator and literacy researcher in South Africa, advocates for multimodal pedagogies; these, she writes, “allow for the expression of a much fuller range of human emotion and experience; they acknowledge the limits of language, [and] admit the integrity of silence” (p. 95).

The idea of multiple literacies has a long and helpful tradition in literacy studies (Cole & Scribner, 1981; Cook-Gumperz, 1986; Gee, 1996; Lemke, 1998; Street, 1984). But it was the New London Group (1996) that provided an especially useful expansion of literacy through the term *multiliteracies*. This team of scholars from various fields—education, linguistics, and sociology, among others—first met in 1994 in New London, New Hampshire, to discuss the big picture of present and future literacy pedagogy (New London Group, 1996), and the group subsequently continued its work in a range of international contexts. In brief, the focus was on “the changing word and the new demands placed on people as makers of meaning in changing workplaces, as citizens in changing public spaces and in the changing dimensions of our community lives—our lifeworlds” (Cope & Kalantzis, 2000, p. 4). The group’s manifesto called for literacy pedagogy to account for the following:

1. “the context of our culturally and linguistically diverse and increasingly globalised societies,”
2. “the multifarious cultures that interrelate and the plurality of texts that circulate,” and
3. “the burgeoning variety of text forms associated with information and multimedia technologies” (Cope & Kalantzis, 2000, p. 9).

At the center of the group’s argument for how literacy pedagogy might take into account such issues, and germane to our article as well, is the concept of design. This notion assumes semiotic activity to be a “creative application and combination of conventions that, in the process of design, transforms at the same time that it reproduces these

conventions" (New London Group, 1996, p. 74, citing Fairclough, 1992, 1995). In the view of the New London Group (1996), it is through an informed, intentional process of design on the part of the individuals, making creative use of available preexisting designs and resources, that meanings, selves, and communities are powerfully made and remade. It is important to note, as does the New London Group, that the process of design in our digital age draws widely on multimodal materials and resources. And in thinking of multimodal texts, it is obvious how useful the notion of design can become as a way to conceptualize the suddenly increased array of choices about semiotic features that an author confronts.

To conceptualize the nature of these choices, we have found the notion of the affordances that are associated with each semiotic modality helpful. Adapting this term from Gibson (1979), who applied it within a scientific, ecological milieu, Kress (1997, 1998, 2003; Kress & van Leeuwen, 1996, 2001) uses it to reference the fit between a semiotic resource, with its inherent properties of organization, and the meaning-making purpose at hand. Pictures, for instance, do not convey meaning in the same way that language does, and as such, their respective meaning-making affordances are different. As Kress (2003) notes, "'the world narrated' is a different world to 'the world depicted and displayed'" (pp. 1-4), making the point that although different semiotic modes may seem to encode the same content, they are nonetheless conveyors of qualitatively different kinds of messages. More specifically, the meaning in images is apprehended by the viewer in accordance with an ordering principle that is spatial and simultaneous, whereas language, particularly oral language, is organized and apprehended temporally and sequentially. It is important to note, however, that despite the particular affordances associated with each semiotic mode, the same kind of meaning can in fact be conveyed in quite different modalities, as Kress (1997, 1998, 2003) also emphasizes. The point is that images, written text, music, and so on each respectively impart certain kinds of meanings more easily and naturally than others. We believe that this idea is the most crucial conceptual tool that one must bring to bear in understanding the workings and meanings of multimodal texts.

The big challenge yet to be taken up within the study of multimodality is how to locate and define the deeper aesthetic power of multimodal texts. Given what individual modes are and do, how might the unique potential of these modes to aesthetically transcend themselves in multimodal composition be conceptualized and

described? This is not to say, of course, that researchers have not offered insights into the effects and implications of acts of multimodal communication. The work of Tufte (1983, 1990, 1997), for example, has been of immense value in understanding how the success or failure of an act of communication often pivots, crucially, on the way in which words, images, and quantitative information are coordinated. As well, Kress (2003, p. 36) discusses the accordant, complementary processes of transformation and transduction (the reshaping of semiotic resources and the migration of semiotic material across modes, respectively) as the locus of creativity in multimodal communication. However, what has yet to be fully conceived and adequately demonstrated, in our estimation, is an approach to understanding how these processes of transformation and transduction actually play out and to what effect. This is our project: to locate and characterize the ephemeral yet aesthetically powerful properties of multimodal text design.

CONTEXT FOR THE STUDY

Our theorizing about multimodality has not been done just in the abstract; we are fortunate to have been able to study digital multimodal texts created by children and adults and to have these often innovative artifacts push and challenge our conceptualizations. More specifically, for the past 4 years, we have been involved in helping to found, fund, and operate a community technology center located in the urban neighborhood of West Oakland, California, a local bus ride away from the University of California, Berkeley. Called DUSTY, or Digital Underground Storytelling for You(th), this center was conceptualized from the outset as a mechanism for making powerful forms of signification (tools for and practices of digital multimodal composing) available to children and adults who did not otherwise have such access at home or at school.⁴ A university and community partnership, it draws professors, undergraduates, and graduate students together with youth and children from the community to study, learn, play, and create. As we have described elsewhere (Hull & James, *in press*), West Oakland is an isolated community that has fallen on very hard times, with high rates of joblessness and crime, a deteriorating infrastructure, struggling schools, and few of the ordinary resources that most communities take for granted, such as supermarkets, bookstores, restaurants, and banks. Many of its grand old Victorians, once

summer homes for the San Francisco wealthy, have been renovated and occupied by outsiders as gentrification intensifies. Yet the West Oakland population, mostly long-time African American residents joined by recent immigrants from Southeast Asia, Mexico, and South America, are finding ways to reclaim its community. With a rich history on which to build, including a significant role during the civil rights movement in the 1960s and a thriving economy related to ship building around midcentury (Rhombert, 2004), residents are currently alert to and working toward safer, healthier, more equitable, and stable futures. DUSTY is but a small piece in this much larger fabric of community growth and change.

DUSTY started as a center to teach digital storytelling, a form of multimedia composing that consists of images and segments of video combined with background music and a voice-over narrative.⁵ Digital stories are, in effect, brief movies distinctive in featuring the digitized voice of the author who narrates a personally composed story and an assemblage of visual artifacts (photographs old and new, images found on the Internet, snippets of video, and anything that one can convert to digital form). In our experience, digital stories have wide appeal among children, youth, and adults, in part simply because they are multimodal and digital, and thereby allow individuals those compositional means and rights that used to be associated just with the world of mass media. They are popular too because they typically privilege a personal voice and allow participants to draw on popular culture and local knowledge. Our youth sometimes create stories that feature their own original digital beats as background music in lieu of commercial hits. Thus, one of the natural expansions of DUSTY has been to teach digital music making, especially because Oakland, California, is known as the birthplace of many famous rappers. At DUSTY, aspiring wordsmiths as young as 9 and 10 can be seen writing their lyrics, practicing their freestyles, and deeply and undistractably engaged in sophisticated software that allows for the creation of digital beats.

A culminating activity at DUSTY is viewing participants' digital stories on the big screen of a local theater or other public venues. On these occasions, we invite authors to answer questions from the audience at the end of the showing. This is one example of how, in the design of our curriculum and participant structures, we attempt to position our digital storytellers as authors, composers, and designers who are expert and powerful communicators, people with things to say that the community and the world should hear.⁶ Given that some

DUSTY youth have not always developed this sense of an authorial self in other settings, including school, the opportunity to do so in an alternative educational site becomes all the more important.⁷ Thus, DUSTY is organized as an after-school program and a summer camp; in addition, we periodically hold workshops for adults and seniors.

Simultaneous with the creation and operation of DUSTY, we have engaged in research roughly within the tradition of design experiments, whereby program development is intertwined with continual attempts to assess and improve our efforts and document what participants have learned (cf. Design Based Research Collective, 2003; Shavelson, Phillips, & Feuer, 2003). Throughout the years, with our colleagues, we have collected a range of ethnographic and qualitative data, principally field notes from participant observations and interviews. We have also videotaped and audiotaped teaching activities, workshops, and community events, including showings of digital stories. Our data also include pre-post inventories and surveys as we attempt to assess not only what kids and adults learn but also how their notions of self as authors and communicators develop. Of late, especially because of the requirements of our funders, we have begun to collect test scores, attendance records, and grades from school in anticipation of comparative studies that will allow us to estimate whether being in DUSTY appears to affect school-based measures. Last, we archive the digital stories and other artifacts that participants create.

In this article, we feature a digital story created by a young man whose particular mix of talents, interests, and predilections seemed to precisely match the available multimodal mediational means. A musician, rapper, poet, writer, photographer, videographer, and a clear-eyed social critic, Randy joined DUSTY in its first year and has thus far produced half a dozen digital stories, many of them featuring his own music. Below, we analyze what we think is his most impressive work, "Lyfe-N-Rhyme," by attempting to articulate the power it derives from multimodality. After a detailed, fine-grained analysis, we end the article by thinking more globally about the affordances and challenges of multimodal composing.

ANALYZING DIGITAL STORIES

Our first pass at analyzing digital stories was simply to categorize their genres and purposes. By watching the approximately 200 stories

created by children, youth, and adults at DUSTY that now reside in our archive, we inductively devised the following broad category scheme:

Genres: autobiographical narratives; poems and raps; social critique and public service announcements; reenactments or extensions of stories, cartoons, and movies; animations; reports; and biographies and interviews

Purposes: Offer a tribute to family members or friends; recount or interpret a pivotal moment or key event; represent place, space, or community; preserve history; create art or an artifact; play or fantasize; heal, grieve, or reflect; and reach, inform, or influence a wider audience

Of course, many authors had multiple purposes, and digital storytelling is an internally diverse and necessarily dynamic and evolving genre. We do not make any claims about the relative frequency or stability of the categories; we simply offer this rough cut as a starting description. Although general, this category system has been useful in pointing to directions for more fine-grained analyses. For example, we were initially surprised by the number of stories by children and adults that centered on space, place, and landscape. But as we reflected, we realized first that the visual nature of digital stories invites authors to situate themselves in places; and furthermore, many of our storytellers made strong identity statements through valences of alignment and distancing in relation to particular locales and neighborhoods (cf. Hull & James, 2005).

It also became clear through the perusal of the stories in our archive that certain ones stood out as especially evocative not only for us but for the wider audiences with whom they were shared. And thus, we began to puzzle over how to account for or where to locate their power. This was an interesting question for us in a theoretical sense, for it stretched our analytic competencies considerably, but it also has great practical import. If digital storytelling becomes widespread, and certainly if it is incorporated into school-based literacy activities, there will need be a way of saying what is powerful about such compositions. Undoubtedly, images, music, and language are each significant conveyors of meaning and sentiment. However, for especially potent stories, such as Randy's "Lyfe-N-Rhyme," it did not seem to us that any one of these modalities was preponderant over the others. Nor is it satisfying simply to believe that increasing the number of semiotic modalities present in a single composition has an accordant

multiplicative effect on the semiotic efficacy of the piece (although this may well be a popular view). The real task, as we saw it, was to understand both the individual and combinatory semiotic contributions made to the synesthetic whole by its material components. To grasp and articulate the emergent qualities of true multimodal design (i.e., design that actually fulfills the promise of expression, which Lemke [1997] describes as multiplicatively powerful), we need to understand the particular logics of organization and respective meaning-making affordances of different modalities. We next describe the methods we devised for this purpose.⁸

Selection of Multimodal Artifact

Given the size of our archive of digital stories and the rich range of multimodal work represented, settling on a story for a fine-grained analysis might have been a complex task. Yet “Lyfe-N-Rhyme” was to us an obvious choice. Of the digital stories from DUSTY that have been viewed by multiple audiences during the past 3 years, Randy’s work has received the most acclaim, its expressive power has been regularly commented on, and its emotional and intellectual impact has been frequently noted. The second factor that influenced our choice of text was analyzability. As this was our first attempt at this kind of analysis, we felt it necessary to choose a piece that was not only an exemplar of powerful expression but that also had an economical design (i.e., an artifact that was not so complex as to make it overly difficult to deconstruct). In comparison to many other digital stories produced at DUSTY, “Lyfe-N-Rhyme” is what we might call manageably multimodal. It does not feature any animation or slick transition effects, but rather, it presents a series of different still images that are coordinated with music and the spoken word.

Identification and Representation of Modalities

When choosing a multimodal text, it is necessary to identify which modes, in relation to each other, will be the focus of the analysis. In an ideal world, one would take into account all of the modes—spoken words, images, music, written text, and movement and transitions—but such complexity quickly overwhelms. We chose to focus primarily on the conjunction of images and words and regretfully gave short shrift to music.⁹ One could imagine, however, depending on one’s

analytic rationale and given the particular features and emphases of a multimodal text, different foci and starting points for analysis.

The next step is to visually represent these selected, simultaneously apprehended modes to transcribe the text in such a way as to clearly illustrate the copresence of each focal mode within the boundaries of appropriate units of analysis. Put another way, one must invent a way to graphically depict the words, pictures, and so forth that are copresented in the piece at any given moment. The form that this transcription scheme takes will be dictated to a great extent by the respective materialities and affordances of the focal modes. Consideration must be given to the principles of temporality, segmentability, and so forth respective to each mode, and a common denominator, so to speak, must be found to parse the piece into analyzable multimodal units. This may be the most challenging juncture of the process.

Our solution to this problem was to transcribe “Lyfe-N-Rhyme” in such a way as to make apparent its couplings of images and language. An important semiotic particularity of the variety of multimodal texts we are working with is that they unfold in time.¹⁰ Therefore, in examining Randy’s “Lyfe-N-Rhyme,” we adopted the momentary conjunction of image-spoken language as the basic unit of analysis. The obvious rationale for this choice of minimal unit (this *multieme*, if you will) is that the independent units into which spoken language and images may each be parsed are fundamentally dissimilar, and as such, they cannot be usefully compared if regarded discretely. However, the trick here was to preserve the temporal relation between the different channels at each moment in the piece but simultaneously set aside the flow of temporality. As a resolution to this problem, we adapted the graphic interface structure used to create the narrative in the first place: a timeline. Adobe Premiere, and virtually every other video editing program like it, entails multiple time-coded tracks in which the various components of the story are rolled out in parallel, so to speak.

With this method of parallel presentation as an inspiration, we conceived of and created a multitracked, horizontal, time-coded transcription format, as shown in Figure 1. Notice that in the specific case of this analysis, the timeline is structured in half-second increments, which was the minimum amount of time that any one image continuously appeared on the screen in “Lyfe-N-Rhyme.” In sum, the temporal flow of the piece is frozen into half-second chunks, in accordance with the shortest on-screen duration of any visual image, rendering a

	18.5	19.0	19.5	20.0	20.5	21.0
IMAGE						
WORD	Truth	Trust	[Silence]	Tribu-	-lation	That's what's up
INTERVIEW NOTES	Randy: That's just something I like to do in my writing sometimes but I really thought it was cool that I did it like that, like in the first one. I was like "love, truth, trust." you know. At first I did it in twos, and then when I moved on it was, "murder, money, mis-education," all those threes, and it's just relating, it just, you know, it's cool like that. (interview, 1/11/02)					

Figure 1. Transcript Format Featuring Time Code and Multiple Tracks (Image, Word, and Interview Notes)

minimal analytic unit of a still image and its corresponding half-second of spoken language. Again, it should be emphasized that there is no one formula for transcribing multimodal texts; the timescale (if there is any at all), segmentation scheme, and so on, must be created in direct relation and response to the modes and questions with which one is concerned.

Identification of Semiotic Patterns Across Modes

Once an appropriate transcript format has been decided on and the work of transcription has been completed, the next stage involves carefully examining the transcript for salient patterns, as with any qualitative investigation. What is different and difficult about this kind of examination in relation to multimodal texts, of course, is that one must not only be cognizant of emerging patterns of various types (thematic, visual, etc.) in each singular mode, but one must also look for identifiable patterns of relation between modes.

In this task, we were aided by two conceptual frameworks, the first helping us think about more local multimodal relationships (e.g., the nature of the pairing of word and image) and the other positioning us to think about more global multimodal relationships (e.g., the functioning of one segment of a digital story in relation to another or the whole). For the first, we drew on the distinctions offered by the American pragmatist philosopher C. S. Peirce (1992, 1998) between icon,

index, and symbol. For purposes of illustrating these concepts, consider any pictorial image of a donkey. According to Peirce's (1992, 1998) formulation, if this picture were to mimetically stand for the idea of donkey, it would be considered an icon. If the image stood for the concept of stubbornness, for example, the conjunction would constitute an index, where the form does not directly express but rather points to (hence the name) a meaning. Last, in Peircian nomenclature, if the donkey picture represented the concept of the Democratic Party, it would be regarded as a symbol, a strictly social, conventional sign. This system of sign classification not only provides a language to articulate different forms of representation, but it also speaks directly to issues of representational power and effect, as will be illustrated in some detail later. But in brief, we looked to see whether an image, when paired with language and music, functioned as an icon, index, or symbol and whether there were discernable patterns of these functions.

For a framework to aid us in understanding multimodal relationships of meaning on the macrolevel of the organization of the composition as a whole, we drew on Labov's (1997) continuing work on narrative theory, which offers, in effect, a kind of conceptual x-ray device, a way of seeing into and talking about narratives and the connections between their meanings and structure. Thus, in the analysis that follows, we apply Labov's notion of the semantic and structural roles of orientation, abstract, and coda in narratives.¹¹ Because most of the digital stories in our archive began as narratives of personal experiences, it made conceptual common sense to us to determine whether the dimensions and aspects of narratives identified by linguists, such as Labov (cf. Ochs & Capps, 2001), were present and, if so, how multimodality functioned to amplify meaning and functions.

Rerepresenting Semiotic Patterns

After identifying salient patterns within, between, and among modes, it is useful to return to the drawing board, quite literally, and rerepresent the text in a form that reveals and juxtaposes these patterns (see Appendix A). Further patterning may become evident and a graphic representation of patterns that emerge from the coding of the transcript may make those *patterns among patterns* more easily discernable. In fact, this kind of richly embedded semiotic patterning may well be the defining feature of powerful multimodal design.

“LIFE-N-RHYME”

Randy’s “Lyfe-N-Rhyme” combined poetry with rap and an autobiography with social critique and was situated verbally and visually within local landscapes and neighborhoods, even though it was aimed toward a broad audience.¹² Two minutes and 11 seconds in duration, this digital story presents a simple, coordinated series of mappings between 79 different still images, 4 of which are each repeated once; the lyrics of an original spoken-word poem (see Appendix B); and the fused driving bass line and somber melody of a classic jazz composition by Miles Davis. The video editing software that Randy used to author his piece, Adobe Premiere (Version 6.5), certainly afforded the panoply of pans, zooms, fades, and spins that are characteristic, for instance, of professionally produced content on the order of MTV. Furthermore, Randy could also easily have incorporated video clips, sound effects, and other technological bells and whistles into his project, yet he did not. What he did do, and we believe to a powerfully transformative effect, was achieve an orchestration (Kress & van Leeuwen, 2001) or braiding (Mitchell, 2004) of language, image, and music into a whole. What we hope to demonstrate is that this whole, in the gestalt sense, transcends the collective contributions of its constituent parts. Although it may be taken for granted that something irreducible is brought into existence whenever parts are assembled, we will show how Randy’s “Lyfe-N-Rhyme,” through an analysis of its design, emerges as a potent symbolic unity, a deeply personal, consequential sign in and of itself.

In his multimedia piece, Randy lays bare intimate, troubling aspects of his life and world, inviting audiences to do the difficult work of reflecting on the intimate concerns they hold for themselves, for those they care about, and for the larger community. Nevertheless, we proceed with the assumption that the story does not tell *the whole story*. We suggest that it is not only the propositional content of what Randy says and shows that moves us, but it is also the composite form into which these elements are organized. Evident in Randy’s “Lyfe-N-Rhyme” is not just a powerful story but also a transcendent synthesis of form and meaning across a variety of semiotic modes. Yes, we are at once touched and disturbed by the words Randy speaks. Yes, the montage of images that he lays out is arranged to an arresting effect: now soothing, now shocking, and so on. And yes, we feel the pulse of the music and verse thump in our chests and minds. Crucially though, we emphasize that the power felt from this piece is not

tantamount to the simultaneous, additive experience of the aforementioned effects, as one might suppose. Again, the full import of the semiotic tapestry that Randy crafts is not merely in but also in between the warp and the weft, as we hope in the following paragraphs to demonstrate.

In the sections that follow, we examine the meaning-making affordances of multimodality, and we argue that Randy's composition evidences patterns within and between different modes that together constitute a multimodal whole. In particular, we illustrate respectively how (a) the visual pictorial mode can repurpose the written, linguistic mode; (b) iconic and indexical images can be rendered as symbols; (c) titles, iconic, and indexical images and thematic movement can animate each other cooperatively; and (d) modes can progressively become imbued with the associative meanings of each other. We proceed through the digital story chronologically, beginning with an account of how titles and subtitles function, and proceeding with a discussion of the opening sequence of the piece, which serves both an orientation and abstract function, in the sense of Labov, as earlier noted. Next comes a discussion of iterative thematic movement in the piece and its relation to the foregoing titles and abstract and orientation section. Finally, we deal with the ending section, the coda, which is constituted by, builds on, and in fact transcends all of the other sections. Throughout, and most important, we show how it is only in the multimodal laminate that these patterns become evident and that these narrative effects are accomplished. This is to say, each pattern is constituted by and constitutive of others in such a way that what is communicated is distinctive and different from what can be accomplished in one modality alone.

We offer one word to the wise before we begin: It is best to view Randy's "Lyfe-N-Rhyme"¹³ and then read the analysis below. We are aware that by attempting to rely primarily on words to explain meaning that is irreducibly multimodal, we are engaging, and are asking readers to join with us, in an uphill task. Please refer to the images and to the schematics provided in Appendices A, B, and C as aids in this process.

Titles as Punctuation

Randy's composition opens with a simple statement of its title. Against a matte black background, about one third of the way down from the upper limit of the screen space, bright scarlet type appears



Figure 2. Title Frame

spelling out “Lyfe-N-Rhyme” in all capitals (see Figure 2). The contrast of blood red on black is startling; and also striking is the juxtaposition of the perhaps most institutional of font styles, Times New Roman, with radically unconventional orthography. In these contrasts, an intangible tension seems to be set up from the beginning. No words are yet spoken. All we hear as we take in the title are the tones of three simple keyboard chords that overlap slightly at their edges.

In the following section, we analyze how titles and subtitles—or more accurately, their visual features of font, style, and color—function in “Lyfe-N-Rhyme.” Our argument will be that multimodality offers opportunities for the expectations set up by conventions of the visual mode to redirect the function of the written linguistic mode. Randy’s choice of linguistic symbols (i.e., written language) to bound the beginning of his piece is certainly not unusual. The genre conventions for a story prescribe that a title be given at the start, as the author well knew. However, what is noteworthy about Randy’s use of a title is that it is not relegated solely to the beginning of the story. In fact, Randy’s story is shot through with titles. At each of six different junctures throughout the piece, he inserts words that function as titles (e.g., “DAMN,” “WAIT,” and “A PAGE FULL OF RAGE”; see Appendices A and C). Such words are not, we would argue, simply word

images. In point of fact, Randy does on several occasions use word images in place of a pictorial image; for instance, when he vocally invokes certain abstract concepts, such as justice and worth, he shows images in which these words are featured prominently in written form. What distinguishes the recurrent use of titles in this piece from other word images is a precise parallelism of style. Subtitles are depicted in exactly the same font and red color and with the same black background as the opening title, "Lyfe-N-Rhyme," to which title status indisputably belongs. As a result of this process of partial transduction, we expect that these word images will be functionally title-like.¹⁴

Again, the dressing of certain word images in a title's clothing, so to speak, creates on some level of consciousness an expectation on the part of the viewer that the word image in question will behave in a title-like way. So in the case of "Lyfe-N-Rhyme," by interspersing title-like word images throughout his composition, Randy effectively delimits what we will call subnarratives within the main text. Thus, when we encounter a title-like image in "Lyfe-N-Rhyme," we are likely to expect that following each subtitle will be a story part that in some respect exhibits the qualities of a fully formed story in and of itself. We may also expect that this story part will end when the next title appears or when the larger story is over (which, by the way, was a defining convention of silent films). In a real sense, then, we see these title-like word images as punctuation of a sort, demarcating both pivotal moments in the larger story and boundaries between its constituent subnarratives. This, we argue, is precisely the kind of emergent structure and meaning that multimodal communication is uniquely able to bring about. We would suggest that here is a case where the influence of the logical organization and meaning-making affordances of the visual (pictorial) mode is serving to repurpose the written linguistic mode in context. The elegance and efficacy of this move are noteworthy, considering the contribution it makes to the power of the composition as a whole, as we hope to show as we proceed with the pieces of our analysis.

A brief methodological aside, we would underscore again the importance of employing the sort of visual transcription described above as an analytic aid. In the case of the subtitles, for example, if one were only to view "Lyfe-N-Rhyme," the connection between these word images—the common visual design features that encode their titleness—might go unnoticed. That is, the word images would be seen in isolation, as they sequentially appeared in the flow of the

piece, separated by a myriad of other images and sounds. Thus, the transcription, and more specifically its affordance of the simultaneous visual presentation and apprehension of multiple images in sequence, makes these patterns comparatively much more salient.

Orientation and Abstract

Nearly 4 seconds into the piece, we abruptly first meet the copresence of image and spoken word. Randy lyrically philosophizes as follows: "What's done through life echoes throughout time. It's an infinite chase to become what I was, but what was I? I don't remember. The only thing I know is I've seen it before in the mirrors of my mind."

Concurrently, in the 13-second period during which these four sentences are unhurriedly delivered, a series of five images occupies the visual field. These images, in order of succession, include a photo of a sunlit Sphinx and the Great Pyramid at Giza; a grainy yet photo-realistic illustration of Malcolm X, in his characteristic black suit and tie and thick-rimmed glasses; a swirling Dalí-esque painting of hip-hop-culture icon Tupac Shakur; the famous rosy-toned three-quarter-view portrait of Marcus Garvey in full military regalia; and a close-up frontal image of the face of rap artist Biggie Smalls¹⁵ partially obscured in shadow (see Figure 3). Also, with the start of this sequence comes a layering of the moody, almost plaintive, trumpet of Miles Davis onto the vibrations of the keyboard chords.

In our analysis, we suggest that each of the multimodal couplings in this 13-second slice of "Lyfe-N-Rhyme," by virtue of their semantic kinship, serve the dual purpose of an orientation section and an abstract section, again to borrow two conceptual tools from Labov's (1997) framework for narrative analysis. First, with these four sentences and five images, Randy orientates or anchors himself and his onlookers in the physical and, most important, cultural time and space of the narrative. His words poetically invite viewers to bear witness to "an infinite chase," his candid examination of society and self. In the imagery, we see a succession of symbols of African American struggles and Black masculinity, with the timelessly monumental accomplishments of ancient Egypt giving way to four Black men who struggled for political, social, and creative self-determination only to be cut down or publicly shamed.¹⁶

Through examining the conjunctions of images and language in these 13 seconds (i.e., what Randy says and what he shows), a unitary



Figure 3. The Opening Sequence of *Lyfe-N-Rhyme*

whole emerges. That is, by noticing the categoric thread that runs through this group of five images and by mapping that onto the meanings in the words Randy speaks, these copresent elements become semantically associated, and their linkage contributes to a magnification of the meaning of the whole: a young African American man's search to reconcile personal identity with culture and history. The piece begins with a statement by means of which Randy opens a window onto a universal quality of the human condition. "What's done through life echoes throughout time," he professes. Concomitantly, the image of the Sphinx and pyramid appears on the screen for 3.5 seconds. He next shifts from stating the universal to relating the personal. Somewhat cryptically, he describes his search for an elusive self that once was but is now lost. The pictures that accompany this description, the four male figures, add value and depth in several important ways. They communicate the image of almost Jungian archetypes, different aspects of possible selves: the outlaw, the moralist, the artist, the statesman, the savior, and so forth. They also represent exemplars of African American males who each sought for something larger than himself and approached an elusive greatness. It is important to notice that these meanings are conveyed without being said as such. There is no direct, iconic, or indexical correspondence between these images and the spoken word; as viewers, it is our cultural, conventional recognition of these figures that impels us on some level of consciousness to derive the aforementioned categorical meanings in the copresence of their portraits.

The semiotic consequence of this recognition is that these images are rendered as authentic symbols and, furthermore, that Randy's personal quest for a lost self is, through language, implicated and echoed in this symbolic whole. Still further, it seems significant that the transitions between the sequenced images of Malcolm X, Tupac, Garvey, and Smalls occur at time intervals that decrease very regularly. That is, the first image appears for 2.5 seconds, the second for 2.0

seconds, the third for 1.5, and the fourth for only 1.0. This regularity fortifies the semiotic bond between these images and also signals a larger level transition to come. In sum, we argue that the resultant multimodal whole both frames the narrative situation and aligns the respective footings of author and audience within that situation (Goffman, 1981). That is, it functions as an orientation section, as mentioned above. And perhaps more important, it presents us with the thesis of "Lyfe-N-Rhyme," what Labov (1997) might call an abstract section.

What is remarkable, moreover, and central to our argument is the irreducible multidimensionality of this sequence. Through each semiotic channel, distinct but related aspects of meaning are imputed to the thesis in such a way that each mode is doing what it does best while exceeding the comfortable, conventional limits of its own meaning-making affordances. This is accomplished by virtue of the multimodal orchestration. For example, we would suggest that the logical simultaneity and semiotic fullness (Kress, 2003) of the images of Malcolm X and the others flow into and fill in the ambiguities of Randy's words in such a way as to create emergent meaning, meaning that makes of each language-image pairing an orientation and abstract clause. And in the abstract section, the whole constitutes the clauses, and the fullness of the thesis positively reverberates. Randy's infinite search is rendered, in some sense, finite. His lost self is given a face. His words and his own destiny are fused with the historicity and aspirations that inhere in those five images, which is his thesis, not accidentally. This is quite an achievement in 13 seconds.

Again, our argument is that this powerfully organic connection of the universal themes that those five images symbolically communicate with Randy's life and personal identity could only have been accomplished within the multimodal laminate. Furthermore, as we will next argue and explain, the instances of emergence that we have pointed at in the story features discussed so far in turn beget additional layers of emergent meaning through higher level forms of multimodal integration.

The Next Level

Thus far, we have discussed two salient patterns in the multimodal design of "Lyfe-N-Rhyme": the title-like word images that meaningfully punctuate the narrative and the 13-second opening section,

which, we have asserted, serves both an orientation function and an abstract function with regard to the piece as a whole. The purpose of the next portion of the analysis will be to show how these two components, each itself characterized by emergent meaning derived from effective multimodal integration, fit within the larger organization of the piece and contribute to the further emergence of powerful, irreducible multimodal expression. We will next illustrate how the aforementioned features, by setting up a succession of self-contained yet deeply interconnected subnarratives, help form a consistent tenor and texture throughout the piece. Moreover, we will demonstrate that these subnarratives, individually and in concert, serve to establish and emblemize Randy's message.

Returning to the point in "Lyfe-N-Rhyme" where we stopped our previous discussion, after 16 seconds of running time, the screen goes suddenly, if only momentarily, black. We hear a single clear tone similar to the faint, lingering knell of a temple bell. As the regular decrease in screen time occupied by each of the first five images foretells, we arrive at a new juncture in the piece. The half-second curtain of darkness opens onto a radically different second act and the first of the subnarratives (see Figure 4): "Life. Love. Truth. Trust. Tribulation, that's what's up. The older we get, the harder a habit is to kick. Damn."

In the nonlinguistic auditory channel, peeling bells and a beseeching trumpet are enjoined by a driving bass-line rhythm, and from this point, the images change precisely in time with the musical beat; the infinite chase is on. The first image, in direct temporal correspondence to the utterance of the word *life*, is a hazy photographic illustration of a dark gray, right-facing silhouette, an androgynous figure represented from the waist up with his or her balled fist held to his or her chin perhaps in a somewhat more upright interpretation of Rodin's "Thinker." Behind the foregrounded figure is a distant green meadow, a cloudless blue sky, and a sweeping, dramatic rainbow that nearly bisects the picture plane diagonally. Five more images follow: a graphic representation of Robert Indiana's ubiquitous pop art sculptural piece "Love," positioned on a white-sand beach; the word *truth* in italicized, violet-colored type on a black background; a photograph of two male Caucasian hands clasped in a handshake against a black field; a photograph of one of New York's World Trade Towers exploding into flames on September 11, 2001; and a photo from the neck up of a well-dressed, middle-aged, bespectacled Black man lighting a cigarette. Significantly, all of these pictorial images evince

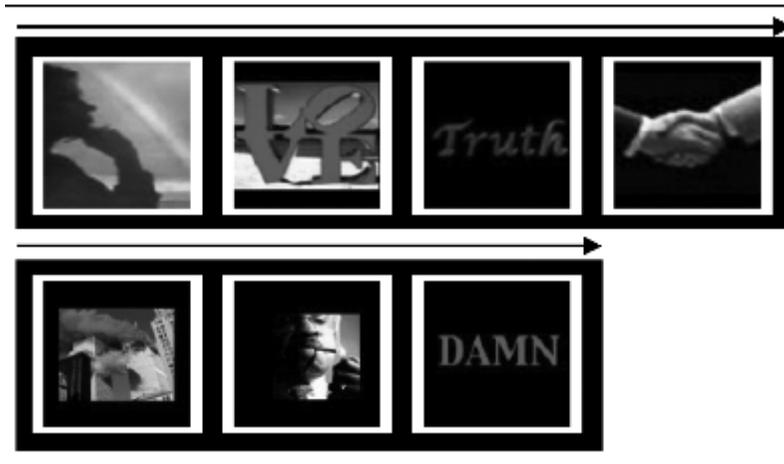


Figure 4. Thematic Movement From the Global to the Particular

iconic or indexical associations with the words they are paired with, their linguistic concomitants an important point that we will revisit. Finally, there is the word *damn* spelled out in capital letters in red Times New Roman letters on a black background.¹⁷

Let us return to the previous argument we offered regarding the boundedness of subnarratives. We see at the outset that Randy invokes universal, almost Kantian categories of human experience: life, love, truth, and trust. Then, he adds an evaluative statement, one that is expressed as truth but not a universal truth. "Tribulation, that's what's up." This statement represents a truth that is necessarily located within the temporal, geographical, historical, and political frames in which the author exists and with which he is concerned, although, the personal concern at this point is still somewhat implicit. Next, however, he states that "the older we get, the harder a habit is to kick." With this statement, he draws a smaller circle around himself, but one that includes all of us as well. By way of pronominal reference, Randy offers a truth that has personal implications for himself and for the viewer. In this subnarrative, and every other, we argue that he structures a participation of the worldview that he presents, a thematic shift from the global to the particular and from the universal to the personal. It is not just the repeated titles that identify discrete subnarratives within "Lyfe-N-Rhyme," but it is also this thematic cline, which marks the plot, to use the term loosely, or movement of each subnarrative and crucially emblemizes the broader thesis of

“Lyfe-N-Rhyme.” Recall the connection Randy makes at the beginning of the piece between broader cultural concerns and his own.

Here is a hybrid construction of visual boundary signals provided by the titles, iconic and indexical images, and universal-to-personal thematic movement. This construction is iterated throughout this main, middle part of the piece.¹⁸ Indeed, the melding of these features results in a multimodal presentation that echoes itself in both form and meaning, the implication of which is that the message is complementarily broadcast on multiple channels at once. To once again underscore our central argument, it is vital to recognize that this situation is quite different from one underpinned by the logic of saying and showing is better than saying or showing. Contrary to popular belief, saying and showing do not automatically amount to powerful expression, and when they do, it is often a matter of coincidence. Multimodal communication is powerful to the extent that the constituent modes are integrated in such a way that they each do what they do well and that these strengths are positioned so as to complement one another. As we have tried to demonstrate thus far, this is the kind of integration that “Lyfe-N-Rhyme” exemplifies.

Globalizing the Personal

The last part of our analysis of “Lyfe-N-Rhyme” focuses primarily on the ending sequence of the piece, the last 21 seconds. This may well be the most interesting part of Randy’s story and the part that affects viewers the most; although, in these 21 seconds, the only words that are spoken are a restatement of the title of the piece: “This is ‘Lyfe-N-Rhyme.’”

There are 10 images that compose this ending (see Figure 5), 7 of which make their first appearances and 3 of which are repeated. The sequence begins with an image of Randy wearing a black knit cap and dark bomber jacket and standing against the backdrop of a cream-colored Victorian house that is in a state of slight disrepair, a type of home quite characteristic of West Oakland. He looks in our direction, basically, but above and past us. Next, we see a frontal view of a different West Oakland home, a large white structure with royal blue trim. The windows are boarded up, and there are official City of Oakland orange safety stickers on the boards warning against trespassing. Following this, we see, once again, the image of a broken sidewalk strewn with litter, the photograph that was in a previous subnarrative to index poverty’s concrete. Next, there are two other repeated

images: a group of young African American men congregating on an Oakland street corner and a hazy view through a rusted chain-link fence onto the blacktop and brown brick buildings of an urban elementary school campus. These images appear earlier in indexical association with the lyrics “capitalism in my veins” (through linguistic implicature indicating that the young men mentioned above are dealing drugs and change, respectively). After this, in moderately rapid succession, several photos appear: images of a bearded olive-skinned man behind the counter of a liquor store, an African American man sitting on a porch, a trash-filled empty lot surrounded by a chain-link fence, and another abandoned West Oakland Victorian. The final image of the piece is another of Randy. This is a photograph that he took of himself, a view of his face and shoulders from beneath, again wearing a black knit cap, framed by the Corinthian columns and crusty architectural details of the same off-white Victorian house that he stood in front of 20 seconds before. Another salient feature of this final image is the bright rectangle of luminescent sky toward which Randy’s head points. This ending sequence of images is markedly different from the preceding two main sections already discussed, but it relates to and builds on them in a profoundly meaningful way.

Before examining the ways that this final section relates to the foregoing two, however, we will specifically discuss three distinctive features of the ending, the characteristics that we believe warrant regarding these 21 seconds as an analyzable whole. First, as mentioned above, no words are spoken throughout the sequence. It is bounded at its beginning by a repetition of the title. Notably, however, the mode of delivery for the repeated words is different. Recall that the only other time the actual title appeared in the piece was at the beginning, where it was typographically shown but not spoken. A further distinction between this section and those preceding it is that this section is bounded by pictures, not words. Instead of title-like word images, we here have two images of the author himself: the first a full-body image, which coincides with his restatement of the title, and the second a skyward head shot, which sees the piece to its end. Last, this part of “Lyfe-N-Rhyme” is distinguished from the other parts in that it does not feature any imagery collected from the Internet or other publicly accessible sources. Rather, all 10 images are photographs taken of a particular West Oakland neighborhood by Randy himself. For the first time, viewers see the undiluted photographic realism of Randy’s lifeworld as he sees it.



Figure 5. Coda or Ending Sequence

In interpreting a general pattern evident in these semiotic moves, we might first say that Randy sets up and trades on differences in the logical organization and meaning-making affordances of counterposed semiotic modes. For instance, the fact that he speaks his title and does not represent it in visual type, as was the case earlier, is suggestive of a transformation. It is ostensibly the same information with the same referent, which is the piece itself. Yet, and this is also key, the medium, in the McLuhanesque sense, conveys a message. When one communicates the same idea in different media, one does not say the same thing in each case (McLuhan, 1964). As such, we argue that in introducing this final section of his piece in oral language as well as in bounding the piece with images of himself instead of title-like expressions, Randy is signaling a metamorphosis of meaning.

Evident here is a transformation involving semiotic enrichment. At the start of the sequence, Randy uses voice in lieu of written text to identify the title of his piece. Attending to the basic organizational logics of each mode, the relative semiotic paucity of meaning inherent in written language—again, in the sense that Kress (2003) intends—is replaced with the semiotically fuller oral language. Moreover, in bounding the sequence with images instead of titles, Randy is again opting for a comparatively greater semiotic completeness over the generic reference of language. A crucial point we must make at this juncture, one that must be clearly understood, is that Randy does not avoid language or its expressive power; his piece is in large part a testimony to and celebration of the power of language. In fact,

counterintuitive though it at first may seem, it is his powerful use of language and his orchestration of it with image and music that permit, or perhaps even call for, the absence of language in the end.

To facilitate an explanation of this seeming contradiction, let us recapitulate key aspects of the analysis that have already been discussed. At the outset of "Lyfe-N-Rhyme," a sequence of several images appears that symbolically embodies the theme that will be pursued throughout the piece. Out of the multimodal mix emerges a unity of the universal struggles and hopes of the African American male with those of Randy himself. Next comes a series of segments, each delineated by title-like word images and comprising iconic and indexical language-image associations. In theme, the subnarrative within each segment in this body of the piece (to apply an essay metaphor) evinces a slide from the universal to the personal, echoing the thesis that was established in the preceding orientation and abstract section. Finally, then, the ending comprises a sequence of images that we have asserted draws import from the virtual absence of language. So is it our assertion, one might ask, that ultimately, this multimodal composition is at its most powerful when it is at its least multimodal? No, this is certainly not the case. What we would say instead, restating an earlier point, is that it is the multimodality of this piece that affords it, in the end, a transcendent unimodality, of a sort; but this *unimodality* is illusive. The photographs that appear at the end of "Lyfe-N-Rhyme" are steeped in the associative meanings that went before; therefore, we might say that the mode that is actually present is imbued in a real sense with the copresence of other nonpresent modes.

By way of explanation, consider the three repeated images in the final sequence: the broken sidewalk, the young men on the street corner, and the school yard fence. As was mentioned before, these images were set up within subnarratives in indexical association with "poverty's concrete," "capitalism in my veins," and "change," respectively, language that speaks directly to the main concerns of Randy's thesis. Necessarily, they carry these meanings with them into the final sequence. Also, each of these images is a photograph taken by Randy himself, as are the other images in the final sequence. The image quality and subject matter of these photographs is quite similar. It seems as if they were all taken on a walk through a particular neighborhood on one particular day (and indeed they were). Instinctively, we view these images as a set by virtue of their stylistic and thematic congruence. Accordingly, although only three of the images explicitly carry

forth meanings that were previously ascribed, viewers feel those meanings percolate out of the repeated three and saturate the set, in effect.

What happens is that this final section in its turn also takes on a wholeness. It becomes a coda in Labov's (1997) terms, a section that brings us back to the beginning in a sense and speaks thematically to the entire piece. In these simple neighborhood photos, the prior indexical link to language lingers unspoken, and the thematic movement resonates. All of these residual meanings coalesce and crystallize within the images in such a way that the global concerns of poverty, crime, desperation, hope, and change are powerfully emblemized. Similar to the portraits of symbolic figures of African American history with which the piece began, Randy's everyday images emerge as profoundly potent symbols in themselves: He conventionalizes the particular and creates a sublime symbolic unity of the global and the personal. And all the while, like the solemn rejoinders of a Greek chorus, the jazz trumpet alternately moans, shouts, and weeps.

CONCLUSION

The most striking claim that we want to make on the basis of our analysis is that Randy's "Lyfe-N-Rhyme" represents a different system of signification and a different kind of meaning. As an irreducibly multimodal composition, "Lyfe-N-Rhyme" is not just a good poem whose meaning is enhanced because it has been illustrated and set to music; rather, we would argue that the meaning that a viewer or listener experiences is qualitatively different, transcending what is possible via each mode separately. Through our analysis, we attempted to characterize the relationality between and among modalities and thereby demonstrate some of the semiotic dimensions and strategies that partly accounted for the emergent meaning of Randy's composition. We also believe our analysis and Randy's story offer a strong counterclaim to the argument that digital media simply facilitate the multimodal composing that could and does exist apart from computer technologies. If we are correct, the particular meanings and the experience of viewing and constructing these meanings via this form of multimodality are unique. Believing as we do that a culture and a time's mediational means, our psychological and material tools if you will (cf. Vygotsky, 1978), are intimately connected with our capacities

to think, represent, and communicate, it would seem hugely important to widen our definition of writing to include multimodal composing as a newly available means.

Having made this appeal, we are at once intrigued and daunted by the amount and range of research and theorizing that waits to be done. Many scholars have commented on the relationships between old and new technologies, noting for example that new technologies often serve old purposes before they come into their own (cf. McLuhan, 1964) or that new technologies can spawn new literate practices that are not necessarily beneficial (cf. Haas, 1999). One of the issues that our research raises for us is the relationship between multimodality and unimodal texts, such as academic essays. Some have argued that visually dominant meaning systems, such as film, can express all of the meanings that a written text can. For us, the more interesting analytic questions concern the blendings that we notice between new and old textual forms. For example, Randy's "Lyfe-N-Rhyme" and digital stories in general deal in linearity and temporality as their narrative-like compositions unfold on the screen. This makes the multimodality of digital stories closer kin with traditional narrative structure but distant cousins to the more associative potential of new digital forms, such as hypermedia. It may be the case that the power of digital stories for creators and viewers has to do with a happy melding of old and new genres and media. We are led to ask, then, what kinds of meanings do these meldings afford, what power do they have, and what constraints do they offer on what people can know, discover, and express (cf. Bennett, 1991)?

Looking back at our analysis, we are still intrigued by what we did not capture, especially around sound and music and the intersection of these modalities with language and image. To give an instance, we recognize that the layering of a hip hop aesthetic onto a classic jazz substrate simultaneously invokes two pivotal moments in African American history and culture, which very interestingly speaks directly to the theme that Randy sets up in imagery and language at the outset of his piece (recall the images of Egypt, Malcolm X, Tupac Shakur, etc.). This layer of meaning is one we did not take into account. Yet music is pivotal as a means of expression and identification, especially for youth. According to Hudak (1999), "so powerful is the desire to make music with others that one is tempted to conceive of music-making as an emergent, radical engagement with consciousness" (p. 447). Frith (1992) notes how important it is that music is directly experienced through the immediacy of the body. We simply

add that the ability to either compose music or merely to use music as a layer of meaning in one's composition; to cast a mood around one's story through a musical choice; to accent, punctuate, or emphasize spoken words through their connection to a beat—these are some of the pleasures attached to music in a multimodal composition. Music also adds an important emotional element to digital stories, and we believe that some of the satisfaction that viewers and creators experience around this form of multimodality derives from emotionality. Though we often ignore the interplay of emotion, cognition, and learning (for a notable exception, see Dipardo & Schnack, 2004), Vygotsky's (1934/1986) vision was for "a dynamic system of meaning" in which "the affective and the intellectual unite" (1934-1386, p. 10). Perhaps opportunities to create and learn through digital multimodality, such as the personal narratives offered through digital storytelling, could be a step toward that unity.

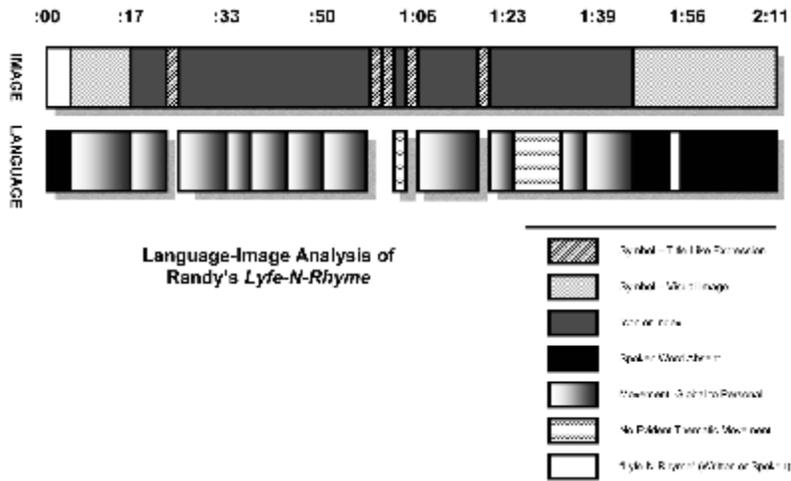
There are great challenges that accompany the incorporation of digital multimodality into classrooms, challenges that are at once technological, economic, and pedagogic. Nevertheless, we have suggested that there is much to be gained from the effort. We conclude by mentioning one last benefit that has to do with multimodality as a democratizing force, an opening up of what counts as valued communication, and a welcoming of varied channels of expression. It gives us great pleasure that "Lyfe-N-Rhyme," a supremely impressive multimodal composition, was created by a poet of the streets and that Randy's combination of abilities and predilections fit the mediational means of digital storytelling like a hand in a glove. We wonder how many other poets and storytellers there are for whom multimodality would offer unexpectedly powerful affordances.

Appendix A

An Analysis of Language and Image in "Lyfe-N-Rhyme"

The image below graphically depicts the concurrent unfolding of patterns of spoken language and image presentation throughout Randy's "Lyfe-N-Rhyme." Global-to-personal thematic movements are shown in gray-scale segments. In the two polka-dotted sections, no thematic movement is evident. The checkerboard areas in the image track represent symbolic use of images, and the solid gray areas show images with iconic or indexical significance. Striped bands represent title-like expressions that are both spoken and

seen in written form, and the white bands show statements of the actual title of the piece, which again is first only seen in writing and then, near the end, only spoken.



APPENDIX B

Script for "Lyfe-N-Rhyme"

What's done through life echoes throughout time
 It's an infinite chase to become what I was
 But what was I? I don't remember
 Life, love, truth, trust, tribulation
 That's what's up
 The only thing I know is I've seen it before in the mirrors of my mind
 The older we get, the harder a habit is to kick
 Damn! Pleasure, pain, purpose, prison
 Justice is a contradiction
 Living on a razor, fell into a felony
 And handled what was left of me
 Life is a lesson
 Groove with me
 Move like a millipede
 Thousands of lands controlled by one hand
 Yes, mama's only gun is mama's only son with a guillotine tongue

Murder, money, mis-education
 Mill gives an incarceration
 Urban voice, heart of the street
 Step by step on poverty's concrete
 Choice, change, crack cocaine
 Capitalism in my veins, yeah, that's what I'm talking about
 A page full of rage!
 Wait! How does a cage rehabilitate?
 Next, America's new war
 Billion dollar weapons don't feed the poor
 But then again, who cares?
 All we do is breathe what they put in the air, yeah
 I said it before, I'll say it again
 Contradiction, Section 8 living
 Society's rival, freedom of speech, who are we to teach
 Heart, body, mind, soul
 So many different worlds in one planet going on
 Youth neglected, expected to listen, born and raised on television
 Friction, failure, function, worth
 Me and Mom Deuce, family first
 Some rules are meant to be broken
 Some doors are meant to be opened
 And . . . regardless of race
 We all mostly come from the same place . . . Love
 This is life in rhyme.

Appendix C

Global to Personal Thematic Movements in "Lyfe-N-Rhyme"

Represented below is a schematic that illustrates each of the global-to-personal thematic movements in Randy's "Lyfe-N-Rhyme." In each white box, the lyrics from a particular section of the piece are transcribed. Lyrics that appear in *italics* evidence a meaning that is global or universal in scope, and those that appear in **bold type** reference more local, personal meanings. The terms that appear in the gray shaded areas in all capital letters are title-like expressions, and those in black shaded areas are statements of the title of the piece. The first title (in black) is displayed only in written language, and the last one is only spoken.

(LYFE-N-RHYME)

What's done through life echoes throughout time. It's an infinite chase to become what I was. But what was I? I don't remember. The only thing I know is I've seen it before in the mirrors of my mind.

Life. Love. Truth. Trust. Tribulation that's what's up. The older we get, the harder a habit is to kick.

DAMN.

Pleasure. Pain. Purpose. Prison. Justice is a contradiction. Livin' on a razor, fell into a felony, and handled what was left to me.

Life is a lesson. Groove with me. Move like a millipede.

Thousands of lands controlled by one hand. Yes. Mama's only gun is mama's only son, with a guillotine tongue.

Murder. Money. Miseducation mill gives an incarceration. Urban voice, heart of the street. Step out, step on poverty's concrete.

Choice. Change. Crack cocaine. Capitalism in my veins. Yeah. That's what I'm talkin' about.

PAGE FULL OF RAGE.

WAIT.

How does a cage rehabilitate?

NEXT.

America's new war. Billion dollar weapons don't feed the poor. But then again, who cares? All we do is breathe what they put in the air. Yeah. I said it before I'll say it again.

CONTRADICTION.

Section Eight living. Societies rival. Freedom of speech. Who are we to teach?

Heart. Body. Mind. Soul. So many different worlds in one planet going on. Youth neglected, expected to listen. Born and raised on television.

Friction. Failure. Function. Worth. Me and Mom Deuce, family first.

Some rules are meant to be broken. Some doors are meant to be opened. And regardless of race, we all mostly come from the same place. Love.

[...]

This is Life-N-Rhyme.

NOTES

1. *Kairos* can be accessed at www.english.ttu.edu/kairos/index.html, and *Born Magazine* can be accessed at www.bornmagazine.com.
2. There has been a great deal of valuable work on literacy and digital media that would fall under the label of new literacies that we do not review in this article because it is not directly related to the form of multimodality we feature here. In particular, there has been very helpful work on hypertextuality and hypermedia (Bolter, 1991), on uses of the Internet (Hawisher & Selfe, 2000), on critical literacy and new technologies

(Knobel & Lankshear, 2002), and on video games as sites for learning and identity development (Gee, 2003).

3. For instance, the report states, with particular reference to the findings of the 2002 Survey of Public Participation in the Arts, that

literary readers are nearly 3 times as likely to attend a performing arts event, almost 4 times as likely to visit an art museum, more than 2.5 times as likely to do volunteer or charity work, more than 1.5 times as likely to attend sporting events, and more than 1.5 times as likely to participate in sports activities. (p. 5)

4. DUSTY was cofounded by Glynda Hull and Michael James. We gratefully acknowledge support for DUSTY from the U.S. Department of Education's Community Technology Centers Grants Program, the Community Technology Foundation of California, the University of California's U.C. Links Program, the Robert F. Bowne Foundation, and the City of Oakland's Fund for Children and Youth. We also thank our university and community partners: University of California Berkeley's Graduate School of Education, the Prescott-Joseph Center for Community Enhancement, Allen Temple Baptist Church, Cole Middle School, Castlemont High Schools, and St. Martin de Porres Middle School.

5. Similar to most organizations that offer digital storytelling, DUSTY traces its beginnings to Joe Lambert and Nina Mullen, who founded the Center for Digital Storytelling (CDS) in Berkeley, California. In fact, we began DUSTY with a donation of old computers from CDS (Lambert, 2002).

6. We are partnering with colleagues in India and South Africa to enable youth to exchange digital stories and digital music.

7. For an account of the theoretical framework on identity that underpins the project, see Hull and Katz (2004).

8. Admittedly, the process by which we came to our interpretation of "Lyfe-N-Rhyme" was not quite as systematically predetermined as our account of our methods might imply. We had no clear preexistent model to follow and as such, intuitively felt our way through much of the analysis. To make it optimally useful and comprehensible to the reader, the procedure described here represents a somewhat cleaned up and pared down explication of the rather more iterative, recursive analytic method we actually applied.

9. By this choice of emphasis, we certainly do not mean to suggest that music and sound are unimportant or semiotically subordinate to the other modes. The musical component of "Lyfe-N-Rhyme" contributes undeniably and importantly to the emergent meaning of the whole, as music almost always does in the most evocative digital stories. It is simply the case that at this point in our own understanding of the entailments and implications of multimodal communication, we find that meaning in music and sound are comparatively more difficult to penetrate than language and image. Nonetheless, woven throughout the following paragraphs is a bit of music-related commentary, and we hope this limited discussion serves to round out our analysis somewhat, lopsided though it is in its privileging of language and image. For an excellent music- and sound-oriented treatment of multimodal communication, see Van Leeuwen (1999).

10. There are many forms of digital storytelling, although the distinctions between the varieties are usually not articulated.

11. We used one other conceptual frame. Taking a cue from de Saussure (1974), we examined the representational choices made along the paradigmatic axis, referring to

the selections of multimodal conjunctions made to fill each half-second semantic slot. Then, we considered how these units operate syntagmatically (i.e., how the individual language-image pairings themselves are interrelated in various ways). Parenthetically, according to this scheme, we might say that this kind of analysis of multiple, copresent modalities requires the addition of a third axis to Saussure's original two, one we might call (only for the sake of preserving parallelism) the *emblematic axis*, describing the semiotic relationality between sequenced elements in different modal channels. One caveat about our analytic methods is that it is not remotely our intent to offer an example of the full elaboration of the complex frameworks of Saussure, Peirce (1992, 1998), or Labov (1997). Rather, we simply and gratefully derive conceptual tools from the work of these scholars for purposes of elucidating only our specific concerns.

12. Randy wanted to reach a wider audience with his work, and in fact, he showed "Lyfe-N-Rhyme" and his other digital stories at a number of public events and forums. One of these was a conference held at the University of California–Berkeley in 2004, the National Council of Teachers of English Research Assembly. At a session featuring work done at DUSTY, Randy premiered a new digital story and answered questions from a rapt audience.

13. "Lyfe-N-Rhyme" can be viewed at <http://oaklanddusty.org/index.php>.

14. The question may arise as to what the semiotic implications are of imputing title-like characteristics to recurrent word images. First and foremost, a title plays an explanatory and referential role in relation to the object with which it is directly associated. The thematic relation between the title and the object may be opaque, as is the case with many examples of work in the plastic arts, and so the referential function is foregrounded. However, the concept of title in the genre of story, the one that primarily concerns us here, calls for a fairly substantial relationality between title and object with regard to both the explanatory and the referential functions. After all, the number of books and popular films we encounter that bear titles wholly unrelated to their narrative content are few indeed. A second critical assumption is that the title refers to only one object. By way of example, if there were a universe in which the films *Jaws* and *Jaws 2* are both naturally called *Jaws*, it would doubtlessly be a confusing place to live. A third assumption, one that is germane only to sequentially organized textual genres, is that the title signals the beginning, that the main content is to follow. One could probably come up with several other postulates for defining titleness, but these three are sufficient to make the present point.

15. Biggie Smalls, also known as Notorious B.I.G., was violently murdered, as was Tupac Shakur. There is popular speculation that the two murders are connected, the product of a personal grudge and a larger East Coast–West Coast rivalry (Smalls was a New York-based artist, and Shakur was based in Los Angeles).

16. Of these four figures, Marcus Garvey was the only one not to be assassinated. He was, however, imprisoned and ultimately deported from the United States after being found guilty of mail fraud. Not long after his deportation, the Universal Negro Improvement Association, a once two-million-strong progressive organization founded by Garvey, virtually collapsed.

17. The "DAMN" presented at the end of the preceding series is one of the title-like word images that, as was previously suggested, segment or punctuate the piece, which begs a brief note of qualification. It may rightly impress the reader that the presentation of this expression at the end of the sequence contradicts the foregoing explanation of the nature of titles: that titles are meant to introduce what follows. So to not be inconsistent, we would remark that the placement of the title at the end of the string above is

intended to illustrate and emphasize the boundedness of the elements between the titles. So to the extent that each of these internal titles forms a boundary between subnarratives, it belongs to both the preceding language-image string and the one that follows. This is not a contradiction but a situational particularization, one might say, of the definition given above.

18. Although there is insufficient space in the text of this article to fully explicate multiple examples of subnarratives, the others are amply illustrated in Appendix C.

REFERENCES

- Bennett, A. T. (1991). Discourses of power, the dialectics of understanding, the power of literacy. In C. Mitchell & K. Weiler (Eds.), *Rewriting literacy: Culture and the discourses of the other* (pp. 13-33). New York: Bergin & Garvey.
- Bolter, J. (1991). *Writing space: The computer, hypertext, and the history of writing*. Hillsdale, NJ: Lawrence Erlbaum.
- Cole, M., & Scribner, S. (1981). *The psychology of literacy*. Cambridge, MA: Harvard University Press.
- Coles, R. (1997). *Doing documentary work*. New York: Oxford University Press.
- Cook-Gumperz, J. (1986). Literacy and schooling: An unchanging equation? In J. Cook-Gumperz (Ed.), *The social construction of literacy* (pp. 16-44). Cambridge, UK: Cambridge University Press.
- Cope, B., & Kalantzis, M. (Eds.). (2000). *Multiliteracies*. London: Routledge.
- Delpit, L. (1995). *Other people's children*. New York: The New Press.
- de Saussure, F. (1974). *Course in general linguistics*. London: Peter Owen.
- Design Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, 32(1), 5-8.
- Dipardo, A., & Schnack, P. (2004). Expanding the web of meaning: Thought and emotion in an intergenerational reading and writing program. *Reading Research Quarterly*, 39(1), 14-37.
- Fairclough, N. (1992). *Discourse and social power*. London: Polity.
- Fairclough, N. (1995). *Critical discourse analysis*. London: Longman.
- Finnegan, R. (2002). *Communicating: The multiple modes of human interconnection*. London: Routledge.
- Frith, S. (1992). The cultural study of popular music. In L. Grossberg, C. Nelson, & P. Treichler (Eds.), *Cultural studies* (pp. 174-186). New York: Routledge.
- Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. New York: Palgrave Macmillan.
- Gibson, J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Goffman, E. (1981). *Forms of talk*. Oxford, UK: Oxford University Press.
- Haas, C. (1999). On the relationship between old and new technologies. *Computers and Composition*, 16, 209-228.
- Hawisher, G., & Selfe, C. (Eds.). (2000). *Global literacies and the World Wide Web*. London: Routledge.
- Hudak, G. M. (1999). The "sound" identity: Music-making and schooling. In C. McCarthy, G. Hudak, S. Miklaucic, & P. Saukko (Eds.), *Sound identities: Popular music and the cultural politics of education* (pp. 447-474). New York: Lang.

- Hull, G., & James, M. (2005). Geographies of hope: A study of urban landscapes and university-community collaborative. Manuscript submitted for publication in P. O'Neill (Ed.), *Blurring boundaries: Research and Teaching Beyond a Discipline*. Creskill, NJ: Hampton Press.
- Hull, G. A., & Katz, M.-L. (2005). *Crafting an agentive self: Case studies on digital storytelling*. Manuscript submitted for publication.
- Knobel, M., & Lankshear, C. (2002, February). *Critical cyberliteracies: What young people can teach us about reading and writing the world*. Keynote address at the National Council of Teachers of English Research Assembly, New York.
- Kress, G. (1997). *Before writing: Rethinking paths into literacy*. London: Routledge.
- Kress, G. (1998). Visual and verbal modes of representation in electronically mediated communication: The potentials of new forms of text. In I. Snyder & M. Joyce (Eds.), *Page to screen: Taking literacy in the electronic era* (pp. 53-79). Sydney, Australia: Allen and Unwin.
- Kress, G. (2003). *Literacy in the new media age*. London: Routledge.
- Kress, G., & van Leeuwen, T. (1996). *Reading images: A grammar of visual design*. London: Routledge.
- Kress, G., & van Leeuwen, T. (2001). *Multimodal discourse. The modes and media of contemporary communication*. London: Arnold.
- Labov, W. (1997). Some further steps in narrative analysis. *Journal of Narrative and Life History*, 7(1-4), 395-415.
- Lambert, J. (2002). *Digital storytelling: Capturing lives, creating community*. Berkeley, CA: Digital Diner.
- Lemke, J. (1997). Multiplying meaning: Visual and verbal semiotics in scientific text. In J. Martin & R. Veel (Eds.), *Reading science* (pp. 87-113). London: Routledge.
- Lemke, J. (1998). Metamedia literacy: Transforming meanings and media. In D. Reinking, M. McKenna, L. Labbo, & R. Kieffer (Eds.), *Handbook of literacy and technology: Transformations in a post-typographic world* (pp. 283-302). Hillsdale, NJ: Lawrence Erlbaum.
- McLuhan, M. (1964). *Understanding media: The extensions of man*. New York: McGraw-Hill.
- Mirzoeff, N. (1999). *An introduction to visual culture*. New York: Routledge.
- Mitchell, W. J. T. (2004, April). *Sounding the idols*. Paper presented at the Conference on Visual Culture, University of California, Berkeley.
- National Center for Education Statistics. (2003). *National assessment of adult literacy*. Retrieved September 14, 2004, from <http://www.nces.ed.gov/naal/design/about02asp>
- National Commission on Writing for America's Families, Schools and Colleges. (2003). *The neglected "R": The need for a writing revolution*. Princeton, NJ: The College Board.
- National Endowment for the Arts. (2004). *Reading at risk: A survey of literary reading in America*. Retrieved September 10, 2004, from <http://www.nea.gov/pub/ReadingAtRisk.pdf>
- New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review*, 66, 60-92.
- Ochs, E., & Capps, L. (2001). *Living narratives: Creating lives in everyday storytelling*. Cambridge, UK: Cambridge University Press.
- Peirce, C. (1992). *Selected philosophical writings, Vol. 1*. Bloomington: Indiana University Press.
- Peirce, C. (1998). *Selected philosophical writings, Vol. 2*. Bloomington: Indiana University Press.

- Rhomberg, C. (2004). *No there there: Race, class, and political community in Oakland*. Berkeley: University of California Press.
- Shavelson, D. C., Phillips, L. T., & Feuer, M. J. (2003). On the science of education design studies. *Educational Researcher*, 32(1), 25–28.
- Stanczak, G. (2004). Visual representation. *American Behavioral Scientist*, 47(12), 1471–1476.
- Stein, P. (2004). Representation, rights, and resources: Multimodal pedagogies in the language and literacy classroom. In B. Norton & K. Toohey (Eds.), *Critical pedagogies and language learning* (pp. 95–115). Cambridge, UK: Cambridge University Press.
- Street, B. (1984). *Literacy in theory and practice*. Cambridge, UK: Cambridge University Press.
- Tufte, E. (1983). *The visual display of quantitative information*. Cheshire, CT: Graphics Press.
- Tufte, E. (1990). *Envisioning information*. Cheshire, CT: Graphics Press.
- Tufte, E. (1997). *Visual explanations: Images and quantities, evidence and narrative*. Cheshire, CT: Graphics Press.
- van Leeuwen, T. (1999). *Speech, music, sound*. London: Macmillan.
- Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Vygotsky, L. (1986). *Thought and language*. Cambridge, UK: Cambridge University Press. (Original work published 1934)
- Witte, S. (1992). Context, text, intertext: Toward a constructivist semiotic of writing. *Written Communication*, 9, 237–308.
- Witte, S. (1993). *Semiotics and writing*. Paper presented at an invited lecture at the University of Wisconsin–Madison.

Glynda A. Hull is a professor of language, literacy, and culture in the graduate school of education at the University of California, Berkeley. Her research examines digital technologies and new literacies, adult literacy and changing contexts and requirements for work, writing and students at risk, and community, school, and university partnerships. Her books include Changing Work, Changing Workers: Critical Perspectives on Language, Literacy, and Skill (State University of New York Press), The New Work Order: Education and Literacy in the New Capitalism (Allen and Unwin; with James Gee and Colin Lankshear), and School's Out! Bridging Out-of-School Literacies With Classroom Practice (Teachers College; with Katherine Schultz).

Mark Evan Nelson is currently working toward a Ph.D. degree in language, literacy, and culture in the Graduate School of Education at the University of California, Berkeley. He has taught both art and design and the English language in Japan and the United States, and his research mainly investigates the social and semiotic implications of multimodal communication in the process of second-language literacy acquisition.