FROM BAKHTIN TO MEDIATED MULTIMODAL GENRE SYSTEMS

Paul PRIOR
(University of Illinois at Urbana-Champaign)

ABSTRACT: Voloshinov and Bakhtin’s expansive view of genres as concrete, historical phenomena, their linkage of dialogic semiotics (discourse) to the formation of individuals and societies (development), has been taken up in North American genre theory as an invitation to explore relations between genre and sociocultural theories (e.g., of Vygotsky, Schutz, Latour, Bourdieu), to see genres not as isolated texts/events but as forged within systems and chains of discourse woven into mediated activity (e.g., Bazerman; Berkenkotter; Prior; Russell), and to challenge the privileging of public texts by identifying genres that are occluded (Swales) or designed to mediate activity (Spinuzzi). Research has focused on semiotic dimensions of genres (e.g. Kress, Lemke), and situated analyses (e.g., Berkenkotter; Kamberelis; Prior) have investigated ways that literate activity involves laminated, multimodal chains of talk, visual representations, gestures, actions, artifacts, and writing. This presentation argues for the notion of mediated multimodal genre systems both theoretically and empirically.

KEYWORDS: genre; multimodality; genre system; activity theory; mediated activity

1. Introduction

Over the past 20 years, so much has been written on genre, so many astute analyses have been undertaken, so many important theoretical observations have been made (see, e.g., BAZERMAN, 1988; BERKENKOTTER & HUCKIN, 1995; COE, LINGARD, & TESLENKO, 2002; DEVITT, 2004; FREEDMAN & MEDWAY, 1994; HYLAND, 2004; SWALES, 2004), that it is challenging now to say something new that needs to be said, especially in the context of an international conference dedicated to genre studies. We have widely agreed for some time now that genres are not solely textual phenomena, that genres should be understood not as templates but as always partly prefabricated partly improvised or repurposed. Over the last 15 years, in different terms and with somewhat different emphases, but with increasing clarity, genre analysts have been moving from a focus on genres as isolated phenomena to a recognition of how specific types of texts are formed within, infused by, and constitutive of systems of genres. Genres have been described in terms of chains (SWALES, 2004; FAIRCLOUGH, 2004), colonies (BHATIA, 2002), repertoires (ORLIKOWSKI & YATES, 1994; DEVITT, 2004), sets and systems (BAZERMAN, 1994, 2004a; DEVITT, 1991, 2004), and ecologies (SPINUZZI, 2004). Theorists have also begun to highlight ways that genre theory has privileged public texts whose primary functions are informational, rhetorical or aesthetic. For example, Swales (1996, 2004) has identified the category of occluded genres, and Spinuzzi (2004) has highlighted the way many workplace genres are designed to mediate activity (e.g., as aids to thinking and action rather than as means of inter-office or external communication). Analyzing the topological and typological dimensions of scientific texts, Lemke (1998) argues that such texts represent multimedia genres, that such genres have long been a feature of writing in the sciences, and that the mix of modalities plays a crucial role in the construction of meaning. Situated genre analyses in specific sites (e.g., BAZERMAN, 1999; BERKENKOTTER, 2001; PRIOR, 1998) have also highlighted ways that literate activity involves multimodal chains of genres (e.g., from planning talk to a written text that is then reviewed orally and in writing), many of which are both occluded and more oriented to mediational or processual purposes of individuals or groups than to wider public exchange. More and more we understand that the rhizomatic threads of genre spread just about everywhere we might look into human societies. What is there to add to these insights, or more to the point, what do I have to add?
Without claiming a unique perspective, I will identify and elaborate on several points that do not seem to me widely shared and agreed to at present. The points I am identifying are related to where I am looking from, specifically from my participation in the field of writing studies, where attention to writing as a process was woven into the formation of the field and where the question of how writing relates to other modes has become a pressing concern. From this perspective, I will focus on four key points, points that lead in the end to a particular way of thinking about mediated multimodal genre systems. The first three points derive from theoretical and empirical attention to writing or more broadly literate activity, the final point from a consideration of disciplinary enculturation as a concrete, dispersed, historical, and thoroughly social and political process. I hope this paper makes some contribution to genre studies (the broad goal of the conference) and to the continuing development of the early insights of Voloshinov and Bakhtin (the focus of the panel in which this paper will be presented).

2. Recovering Voloshinov’s theory of the utterances from Bakhtin’s later definition

First, it is important to distinguish the work of Voloshinov from that of Bakhtin. (See MORSON & EMERSON, 1990, for an astute historical analysis of the authorship disputes and BAZERMAN, 2004b, for a substantive argument that articulates the close links between Voloshinov’s theories and Vygotsky’s, links that are much weaker in Bakhtin’s work.) Bakhtin’s (1986) essay, “The problem of speech genres,” offers an account of genres that has profoundly altered genre theory in the past decades. However, that seminal essay also displays how thoroughly Bakhtin’s approach to genre was grounded in literary issues, rather than the linguistic, semiotic, psychological, and sociological perspectives that Voloshinov (and evidently Medvedev) engaged. In fact, Bakhtin displays his limits in a prominent and repeated way: seriously undermining and confusing the fundamental unit of analysis in his theory, the utterance.

Let’s turn to three quotations that illustrate the problem. In the following passages, Bakhtin (1986) is defining utterances (spoken and written) as the real unit of speech communication (in contrast to the abstract sentences of linguistic analysis):

The boundaries of each concrete utterance as a unit of speech communication are determined by a change of speaking subjects, that is a change of speakers. Any utterance—from a short (single-word) rejoinder to the large novel or scientific treatise—has, so to speak, an absolute beginning and an absolute end…. (p. 71)

Complexly structured and specialized works of various scientific and artistic genres, in spite of all the ways in which they differ from rejoinders in dialogue, are by nature the same kind of units of speech communication. They, too, are clearly demarcated by a change of speaking subjects, and these boundaries, while retaining their external clarity, acquire here a special internal aspect because the speaking subject—in this case, the author of the work—manifests his own individuality in his style, his world view, and in all aspects of the design of the work. (p. 75)

The work is a link in the chain of speech communication. Like the rejoinder in dialogue, it is related to other work-utterances: both those to which it responds and those that respond to it. At the same time, like the rejoinder in dialogue, it is separated from them by the absolute boundaries of the utterance. (p. 76)

Bakhtin’s problem here is not subtle. In defining the utterance as the real unit of speech communication, he makes two claims that undermine the power of a dialogic approach. The first is the equation of utterances with externalized utterances. The second is the equation of spoken utterance (talk) with works (texts), his insistence that a “Hi!” spoken on the street to a
passing stranger and Tolstoy’s *War and Peace* each count equally, each one utterance, each a move in a sequence of dialogue. Bakhtin does seem to sense the oddness of this claim, but locates the difference in marks of individuality in texts, marks that index the vision and craft of the author (yet another indication of what a narrow literary canvas Bakhtin was painting on).

Bakhtin’s departure from the earlier theory articulated by Voloshinov could hardly be more plain:

The process of speech, broadly understood as the process of inner and outer verbal life, goes on continuously. It knows neither beginning nor end. The outwardly actualized utterance is an island arising from the boundless sea of inner speech, the dimensions and forms of the island are determined by the particular situation of the utterance and its audience (Voloshinov, 1973, p. 96).

Where Bakhtin (1986) boasts of the absolute beginning and end of utterances (which he believes gives them a hard scientific currency), Voloshinov argues that speech has neither beginning nor end, that utterance is an island rising from the sea of inner speech. Voloshinov (1973) had earlier articulated the point more generally as a semiotic rather than solely a linguistic issue (and Voloshinov does not mention signs only in passing).

We repeat: every outer ideological sign, of whatever kind, is engulfed in and washed over by inner signs—by the consciousness. The outer sign originates from this sea of inner signs and continues to abide there, since it’s life is a process of renewal as something to be understood, experienced, and assimilated, i.e., its life consists in its being engaged ever anew into the inner context. (p. 33)

Of course, Voloshinov’s (1973) attention to inner speech and consciousness needs to be placed in the broader context of a social (ideological) theory of the formation of consciousness: “Consciousness takes shape and being in the material of signs created by an organized group in the process of its social intercourse” (p. 13). In another early, disputed text, Bakhtin/Medvedev (1978) locates utterance and genre firmly within as well as outside of the individual:

It is the forms of the utterance, not the forms of language that play the most important role in consciousness and the comprehension of reality. …we do not think in words and sentences, and the stream of inner speech which flows within us is not a string of word and sentences. We think and conceptualize in utterances, complexes complete in themselves. As we know, the utterance cannot be understood as a linguistic whole, and its forms are not syntactic forms. These integral, materially expressed inner acts of man’s orientation to reality and the forms of these acts are very important. One might say that human consciousness possesses a series of inner genres for seeing and conceptualizing reality. (pp. 133-134).

Voloshinov clearly had a robust notion of utterance as inner speech and inner genre that Bakhtin only fleetingly affirms and easily abandons. (When Bakhtin writes of inner speech, he is typically writing of the *representation* of inner speech for characters in a novel.) From Bakhtin’s definition of the utterance as externalized utterance and his clear equation of talk and text flow the problems that I address in the next three sections.

3. Revisiting the problem of the text: The composed utterance

When Voloshinov and Bakhtin articulated their account of utterance and distinguished utterances from the specialized representation of “decontextualized” linguistic sentences, they
aimed to put the study of language and, especially for Voloshinov, signs, firmly in the lived world, in concrete space and time. However, locating signs-in-use also called for an initial recognition of the complex temporalities of semiosis. Utterances do not achieve their sense and function in a moment. Their relevance, production, interpretation, and use all require attention to temporal trajectories—to the histories that lead to an utterance, the unfolding events of its use, and the imagined projections of its future (of how the utterance might be understood, taken up, replayed and reused in near and perhaps more distant futures).

Writing Studies, which focused attention early (EMIG, 1971) on the acts of composing that lead to a text, suggests the need to see written utterances (the situated moment-to-moment production of texts) as historical acts exactly on a par with spoken utterances (the situated moment-to-moment production of talk). Collapsing years of written production across diverse events into the moment of publication (if such a moment ever arrives and for many, perhaps most, texts, it does not) is a high price to pay for “proving” that utterances are real units of communication.

However, the problem of the text, what I am calling composed utterances (for reasons that should become clear shortly), remains. If on-line production of written utterances is equated to on-line production of spoken utterances, how do we understand texts that emerge out of long histories of production, texts that are composed and often lengthy? Such utterances not only have a history, like “Hey, Sally, what’s up?”; they have a history of composition. Interestingly, the problem of the composed utterance is not limited to texts; it also applies to talk—to formally composed talk, repetition of memorized text, and even events that are worked out orally in advance. Judith Irvine’s (1996) analysis of insult poetry at Wolof wedding ceremonies makes this point clear as she examines how the insults are composed prior to the event by sponsors and others in the community with a griot (a low-ranking women bard), how the griot delivers and leads the insults in the event, and the way that what Irvine calls shadow conversations (those conversations that are not here-and-now but felt here-and-now) are critical to the production, uptake, and interpretation of the insults.

Composed signs (whether material artifacts, enacted performances, or both) are not unique in having a history, but are special in the ways that history is sedimented into and impinges on the present. Composed utterances (and genres) call on us to analyze the history of composition—the chain of utterances that are woven together, the various ways that a composed document/performance may overtly or covertly index those focused histories of composition, and the ways that production, reception, and use take that history into account. For example, to understand an utterance by an actor on a stage or in a film, the way the utterance is delivered, the way the audience interprets it, the way it is re-used and re-presented, it is critical to understand the shadow conversations, writings, and texts that are at play. The utterance indexes, not simply some authorial vision, but also the social identities and discourses represented, the interpretive work of the actor who is animating her lines, the influence of the director, stage crew, and others who have shaped both the occasion and contextualization of the utterance and the actors’ interpretation. Political speeches, film and stage drama, religious ceremonies, sales pitches, language drills, sermons—once we begin to look, a lot of talk fits into the category of composed utterance, sometimes with texts woven into the history but sometimes (as in Wolof insult poetry) without it.

Kevin Roozen’s (PRIOR, HENGST, ROOZEN, & SHIPKA, 2006) analysis of semiotic remediations in the historical trajectory of an amateur comedy skit offers us a detailed glimpse into the complexity of such composed performances, particularly the way compositional events can bring together multiple people who co-compose the text/performance in interaction. The presence of such interactions in the history bumps up against the second problem, the need to recognize writing not only as activity, but as activity that can be face-to-face. If we conceptualize genres as involving production, reception,
distribution, and representation, then it is important to not see these as separate stages, but as co-present dimensions of discourse with multiple and changing configurations over time.

4. Animating written utterances: Literate activity as co-present production

Even in some of the richest theoretical and empirical work, there remains a tendency to freeze writing, as though it entered the world from some other realm, to see writing as a noun rather than a verb, to specifically not study writing. For example, in what is otherwise a sophisticated account of dialogic theory and method, Linell (1998) devotes almost no space to the question of writing. When he does turn to writing, he touches briefly on the notion of writing as activity but clearly fills in the blanks with cultural assumptions rather than the kind of close research attention he offers talk:

Written texts, being permanent records, encourage the view that the meanings of texts “are there” “in the texts themselves.” But meanings are of course assignments and accomplishments by human beings, writers and readers. The production of meaning takes place in interactions, on the one hand in the writer’s struggle with thoughts and words in conceiving and formulating the text and in her interplay with the text so-far produced, and, on the other hand, in the reader’s efforts in assigning meaning to the text and in using the text as a vehicle, as a means for activating semantic potentials of words and text chunks, in the service of creating an understanding which somehow fits the contexts given and purposes which are relevant for him. (p. 268)

Linell usefully does invite us to consider text as a human product, to see writing and reading as acts, and also notes the role of in-progress text; however, he imagines a culturally prototypical scene of writing (see PRIOR, 1998, for analysis of such scenes) rather than studying actual scenes of writing. In Linell’s scene, the writer is always alone, the text is always permanent, the reader is always somewhere else, making meaning on her own. Scollon and Scollon (2003) also highlight this blind spot in current theorizing of discourse. Their approach to mediated discourse and what they call geosemiotics offers a theoretically rich and empirically rigorous examination of semiotic practices in material worlds. They pay close attention to ways that texts are handled, to the complex textures of texts, even to esoteric issues like text vectors in signage. Critically however, writing does not appear as activity on their expansive map. Consider the following quote:

…”there are three ways in which language can be located in the material world, the interaction order (including speech, movement, gesture), visual semiotics (including text and images), and place semiotics (all of the other non-linguistic symbols that directly or indirectly represent language). Geosemiotics analyzes the semiotic systems among which we take action in the world. (p. 13)

The point I want to draw attention to here is that the interaction order is glossed as “speech, movement, gesture” but not as writing. Writing (or at least its products) only appears in the next item, visual semiotics. It is true that this glossing of the interaction order is not presented as complete, yet writing does not come back in as action. Farther down in the paragraph, Scollon and Scollon indicate that their interest is in bringing together studies of the interaction order (talk, movement and gesture) and textual analysis (study of the structures of text).

Writing must be done in particular times and places and it can be done in face-to-face social interactions. Writing has begun to emerge as a face-to-face activity in studies that look at people working around boards and screens (see, e.g., HEATH & LUFF, 2000; HALL, STEVENS, & TORRALBA, 2002; PRIOR, 2007) and in situated studies of writing processes across varied settings (see, e.g., BAZERMAN, 1999; BEAUFORT, 1999; IEDEMA, 2003;
KAMBERELIS, 2001; PRIOR, 1991, 1994, 1998; PRIOR & SHIPKA, 2003). At first, examples of face-to-face writing and reading may seem esoteric, until we recognize that board writing in schools and workplaces routinely involves co-present writing and reading. Such texts are typically also temporary, not permanent records (as are many other texts, written on scraps of paper, windows, in the dirt or snow, and so on). It is also worth noting that many early literacy experiences seem to involve face-to-face reading and writing, something we would expect from a Vygotskyan perspective where practices move from the social to the (relatively) individual.

5. Voloshinov and Vygotsky: The currents of inner and outer semiotics as multimodality

That writing is a process also means that writing is a stream within the broader flows of semiotic activity. If we see genres as produced in processes, process that have histories, then genres are multimodal, not only in the sense that a particular genre will always be realized materially in multiple media but also in the sense that multimodality arises in the histories of chaining that mark processes and in the mix of inner and outer semiotics.

In “The problem of speech genres,” Bakhtin (1986) does reach into inner semiotics when he defines utterance, but only in the arenas of planning (by the speaker or writer) and reception (the inner responses of people). Voloshinov (1973), as I note above, articulates a much more robust and central notion of inner speech, inner genre, and inner semiotics. Extending the notion we saw above in quotation from Bakhtin/Medvedev (1978) that we possess inner genres to perceive and understand reality, Voloshinov (1973) suggests how ideological content, especially in the form of social evaluations, can be found even in inner feelings and emotions.

…not even the simplest, dimmest apprehension of a feeling, say, the feeling of hunger not outwardly expressed—can dispense with some kind of ideological form. Any apprehension, after all, must have inner speech, inner intonation and the rudiments of inner style: once can apprehend one’s hunger apologetically, irritably, angrily, indignantly, etc. (p. 87)

Prior and Shipka (2003; see also PRIOR, HENST, ROOZEN, & SHIPKA, 2006) argue that Vygotsky’s fundamental theory of human development and consciousness was very attuned to the semiotic transformations that link the inner semiotics of thought, perception, motivation and feeling to the outer semiotics of action (talk, writing, drawing, object production and manipulation, movement, stance). Vygotsky (1987) observed:

External speech is not inner speech plus sound any more than inner speech is external speech minus sound. The transition from inner to external speech is complex and dynamic…. (p. 280)

Hutchins’ (1995) account of distributed cognition as “the propagation of representational state across representational media” (p. 118), where one medium is the brain, begins to suggest how the boundaries of inner and outer are neither negated nor equated, but blurred and softened. Seeing genres as inner and outer; as semiotically remediated; as central to learning, the formation of the person, thought and motive, these are the kinds of issues that flow from Voloshinov’s boundless inner sea of signs fed by the ideological streams of cultural-historical practice.

In this light, multimodality is not some special feature of texts or certain kinds of utterance, and certainly not a consequence of technologies (cf. KRESS, 2003). Multimodality has always and everywhere been present as representations are propagated across multiple media and as any situated event is indexically fed by all the modes present, whether they are
focalized or backgrounded. In this sense, all genres are irremediably multimodal; again, the question becomes what particular configurations of multimodality are at work in a particular genre system.

6. Re-figuring the social: From social communities to flat actor-network flows

Conceptualizations of the social remain a critical area for genre studies. If genres are somehow both grounded in and constitutive of the social, if genres are learned in and formative of the ontogenetic trajectory of the individual person, then the question of how we understand the social becomes critical. An early and persistent tendency in the field has been to link genres to well-organized social entities. Genres might be owned by discourse or speech communities (SWALES, 1990) or perhaps by activity systems (RUSSELL, 1997). The notion of such systems can invoke a tree diagram or mapping of formal constituents. However, it is precisely that type of conceptualization that Voloshinov critiqued as an abstract objectivist view of semiosis. Systems can also be understood as historical and chaotic, as open and complex, as evolving. In this sense, we might begin to examine more seriously the consequences of seeing activity and genre systems as assemblages or actor-network rhizomes (LATOUR, 2005), mycorrhizal formations (ENGESTRÖM, 2006), as flow architectures (KNORR-CETINA, 2005). Prior (1998) makes this argument at some length, though the last three citations offer new ways of articulating a flat, dispersed, historical understanding of the social. Nevertheless, the gravitational pull of structuralist notions of the social remains powerful (for further discussion see also PRIOR, 2003)

7. Summary: Mediated multimodal genre systems

Voloshinov and later Bakhtin articulated an expansive view of genre as a concrete, historical phenomenon. Their historical orientation is key not only to a dialogic, non-structuralist understanding of language (more broadly signs), but also to the integration of semiotic mediation with a sociohistoric account of the formation of individuals and society. This dual emphasis on genre as discourse and development has led North American versions of genre theory in particular to explore relations between genre theory and sociocultural theories of mediated activity and agency (e.g., of Vygotsky, Engeström, Wertsch, Latour). I propose extending attention from activity systems and multimedia genres to mediated multimodal genre systems, where the mix of media and modes appear not only in specific texts, but also in their use (e.g., a text may be written to be read), in the chains of texts that make up the whole system (e.g., where a sequence of oral and embodied genres of discussion, inquiry, response, and presentation may mix with written and visual genres—or, more to the point, where a set of differently configured multimedia genres are linked together in locally situated ways), and in the consciousness of people as well as in artifacts and actions. This paper then argues for a multimodal or semiotic perspective on genre systems, considering the genre system as fundamentally constituted in the varied activities and artifacts involved in trajectories of mediated activity, in the whole ensemble of production, reception, representation, distribution, activity, and socialization (see, e.g., PRIOR ET AL., 2007).

References

BAKHTIN, M; MEDVEDEV, P. *The formal method in literary scholarship*: a critical introduction to sociological poetics (A. J. WEHRLE, Trans.). Baltimore: John Hopkins
BAZERMAN, C. *Shaping written knowledge*: the genre and activity of the experimental article in science. Madison: University of Wisconsin Press, 1988


BERKENKOTTER, C.; HUCKIN, T. *Genre knowledge in disciplinary communication*: cognition/culture/power. Mahwah: Lawrence Erlbaum, 1995


