Young People’s Everyday Literacies: The Language Features of Instant Messaging

Christina Haas

Pamela Takayoshi
Kent State University - Kent Campus, ptakayos@kent.edu

Brandon J. Carr

Kimberley Hudson

Ross Pollock

Follow this and additional works at: https://digitalcommons.kent.edu/engpubs

Part of the Communication Technology and New Media Commons, Discourse and Text Linguistics Commons, Interpersonal and Small Group Communication Commons, Reading and Language Commons, and the Social Media Commons

Recommended Citation

This Article is brought to you for free and open access by the Department of English at Digital Commons @ Kent State University Libraries. It has been accepted for inclusion in English Publications by an authorized administrator of Digital Commons @ Kent State University Libraries. For more information, please contact digitalcommons@kent.edu.
Young People’s Everyday Literacies: The Language Features of Instant Messaging

In this article, we examine writing in the context of new communication technologies as a kind of everyday literacy. Using an inductive approach developed from grounded theory, we analyzed a 32,000-word corpus of college students’ Instant Messaging (IM) exchanges. Through our analysis of this corpus, we identify a fifteen-item taxonomy of IM language features and frequency patterns which provide a detailed, data-rich picture of writers working within the technological and situational constraints of IM contexts to creatively inscribe into their written conversations important paralinguistic information. We argue that the written features of IM function paralinguistically to provide readers with cues as to how the writing is to be understood. By writing into the language paralinguistic cues, the participants in our study work to clarify, or more precisely disambiguate, meaning. Through a discussion of four of these features—eye dialect, slang, emoticons, and meta-markings—we suggest how the paralinguistic is inscribed in IM’s language features.

With digital technology . . . it seems, writers are *everywhere.*

—Yancey, 2009, p. 4

In “Writing in the 21st Century: A Report from the National Council of Teachers of English,” NCTE Past President Kathleen Blake Yancey (2009) issued a tripartite call to action for writing researchers and teachers to develop new models of composing, new models of curricula, and new models for teaching. Tracing a
history of school- and test-based writing and writing instruction in the last century, Yancey noted a parallel history, one that took place outside of school in the everyday lives of American citizens. “... [O]utside of school, people wrote” (pp. 3, 4) echoes throughout Yancey’s history, and the digitized writing practices ubiquitous in contemporary culture support her claim that “writers are everywhere” (p. 4). Specifically, the everyday worlds of college students are saturated with writing. Students use a rapidly evolving set of digital resources, not just instant messaging but Facebook, MySpace, and Twitter, to build and maintain social worlds that exist outside of school—a kind of parallel world (as Yancey would have it) that is composed of written texts, albeit ones that look distinctly different than the texts typically produced in academic and college settings.

In this article, we take up Yancey’s (2009) call to “articulate the new models of composing developing right in front of our eyes” (p. 7) by providing a detailed examination of the language features of instant messaging (IM), one form of interactive written discourse (Ferrara, Bruner, & Whittemore, 1991) popular with college-aged students. As one basis for developing new theories of writing, an analysis of the language features of IM—as it is practiced among the middle- and working-class college students whose writing we studied—can provide a glimpse of what writing looks like now, in the early years of the 21st Century. Especially with language forms mediated by new composing technologies, it is easy to overlook the actual writing (the letters, punctuation, spelling, words, and sentences), to focus instead on the interesting ways that literacy and writing are being placed on center stage in people’s everyday lives. It is exciting as writing researchers to explore the ways self-sponsored writing is being integrated into and made important in people’s lives. Indeed, the evolution of the mobile phone from a speaking to writing tool has been fascinating for us in what it suggests about the power of writing. Of course, to study writing can—and should—mean many things: to study the ways that texts are produced, by single writers, dyads, or groups; to study the functions of literacy in personal development and in social and institutional contexts; to study how writing is mediated and distributed; to study the practices of literacy in situ; and to study how written texts work within or against cultural discourses, among others. But we also believe it is important to understand what is happening to writing itself as a product in these new environments for writing. What shape does writing take? What innovative forms are being used? Are there patterns of usage?

MacDonald (2007) believes that as a profession, writing researchers have steadily moved away from seeing language-level concerns as interesting or worthy of attention: “I see a declining focus on language in the last three decades. ... an erasure of our attention to language ... We have spent ... not enough time either on basic understandings of the English language or on sorting out, evaluating, and revisiting the arguments that have ended in the current impasse and the neglect
of language” (p. 618). In a digital and global world where written communication is increasingly multimodal, an erasure of attention to language prevents writing studies scholars from understanding changes at the level of the language, and more broadly, limits understandings of contemporary literacy practices. Kress (2003) pointed attention to changes happening in the communicative contexts in which written language is situated, while at the same time recognizing that these changes are not merely occurring at the global level: “writing is undergoing changes of a profound kind: in grammar and syntax, particularly at the level of the sentence, and at the level of the text/message” (p. 21). As writing researchers, we are interested in what these global changes in the function of the written word mean for the forms writing takes. In the numerous digital locations where the written word remains dominant (email, chatrooms, blogs, instant messaging, text messaging), it seems pressing for writing researchers to examine just how writing is being shaped.

Although we are not linguists by training, our inclination has been to look at IM language features in the fine-grained, discourse analytic way that linguists do; further examinations of IM (and related forms of CMC) would certainly benefit by using the more precise tools of linguistic analysis to explore the specific orthographic and syntactic nature of new media writing forms. But such is beyond the scope of this article. Concerned with providing a general description of IM useful for writing teachers and researchers, our approach in this project has been to examine closely the written language features of IM. Specifically, we are interested in IM as one instance of students’ everyday literacy, the kind of writing that has long taken place outside the halls of academe. We begin by considering the ways non-academic writing has been understood; then we turn to an examination of a 32,000-word corpus of non-academic IM conversations conducted by college-aged writers. Through our analysis of this corpus, we identify a fifteen-item taxonomy of IM language features and frequency patterns which provide a detailed, data-rich picture of writers working within the technological and situational constraints of IM contexts to creatively inscribe into their written conversations important paralinguistic information. These IM features are situated within the lifeworlds of writers in a digital, mediated world of distanced, written communication. While the language features are quite distinct from academic writing, we believe they are valuable for understanding the creative possibilities for drawing together the oral and literate in a world where writing is increasingly being conducted in online spaces in which paralinguistic cues are necessary but difficult to embody.

**Parallel Writing Worlds beyond School**

Many writing studies scholars have devoted attention to the rich variety of writing that takes place beyond the walls of the classroom, in communities, homes, and other sites of everyday literacy. Witte (1992), for example, stipulated criteria for a
comprehensive theory of writing and presented a series of examples of school- and workplace-based writers, including a detailed and memorable analysis of the function of one everyday literacy, the “grocery list,” as a multi-semiotic and complex written artifact. Gere (1994, 1997) examined both contemporary and historical instances of writing groups that flourished outside of school settings. She argued compellingly for a consideration of the often transformative practices of writing in everyday contexts far beyond the walls of our classrooms. Gere’s is not a vision of colonization, in which writing scholars and teachers take over such sites, but rather one in which the amateur is valued, access is open, and power is shared. Studies of everyday literacy include, among many others, Brandt’s (2001) study of literacy sponsorship across lifetimes, Cushman’s (1998) work with inner-city residents negotiating institutional literacy, Dyson’s (1997, 2008) research on the literacies that young children bring to school, Farr’s (2005) studies of Latino literacy in Chicago, Flower and colleagues’ study of literacy practices at a Pittsburgh community center (Flower, Long, & Higgins, 2000; Peck, Flower, & Higgins, 1995), Hull and Schultz’s (2002) edited collection on out-of-school literacies, and Moss’s (2003) research on literacy practices in African American churches.

Significant theorizing of the literacies of the everyday is accomplished in two related projects: the New Literacy Studies and the work of the New London Group. Barton, Hamilton, and Ivanic (2000) brought together essays written primarily by participants of the New Literacies Research Study Group at Lancaster University. They noted that “literacies are situated... located in particular times and places [and] indicative of broader social practices” (p. 1). Significantly, their definition of literacy practices as “the general cultural ways of utilizing written language which people draw upon in their lives” (p. 7) emphasizes the everyday, and individual chapters in the volume discuss literacies in homes (Ormerod & Ivanic 2000), among Welsh farmers (Jones, 2000), and in prisons (Wilson, 2000).

The work of the New London Group (Cope & Kalantzis, 2000) takes up similar themes. While the NLG is interested in how literacy, through web technology, is becoming an increasingly global enterprise, they also emphasize the local and the everyday. The focus of the group’s initial meeting in New London, New Hampshire, was “the changing word and the new demands being placed upon 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). Their volume presented two related arguments, the first relating “to the increasing multiplicity and integration of significant modes of meaning making” (p. 5) and the second “to the realities of increasing local diversity and global connectedness” (p. 6). Directly relevant to our work, chapters in each of these volumes deal explicitly with literacy in new media contexts (de Pourbaix, 2000; Luke, 2000). Although adopting the general inclusive approach to literacy that characterizes the NLS and the work of the NLG, de Pourbaix (in her examination
of academic newsgroups) and Luke (in her analysis of factors impacting new media literacy) focus on schooling.

Another member of the NLG, James Paul Gee, has been particularly important in shaping our thinking (1999, 2000a, 2000b, 2003). Gee’s interest in and commitment to literacy in what he calls “lifeworld domains” (2003, p. 36) is directly relevant to our work with the everyday literacies of new media language. Specifically, Gee (2003) directly addresses an ubiquitous technology in the everyday lives of young people: video games. Although the video game is not conventionally thought of as a literacy technology, Gee argues compellingly that, by virtue of their multi-modal nature and the way they tap into and extend semiotic domains, video games can be powerful tools for making, shaping, and communicating meaning. The importance of Gee’s work for us has been in his understanding of humans as meaning-making creatures and of technologies—especially everyday technologies—as tools that humans use, in often unexpected ways, to build and convey meaning. A similar open spirit to the literacy practices of the everyday is evident as well in Lankshear and Knobel (2003), Selfe and Hawisher (2004), and in the volumes edited by Lewis, Enciso, and Moje (2007) and by Purcell-Gates (2007).

In this article, we are interested in IM as a literacy practiced by young people, particularly in North America. The body of scholarship on IM is growing, as is that on related technologies such as Facebook (Lampe, Ellison, & Steinfield, 2006), texting (Crystal, 2008; Baron, 2008), and online minute-by-minute commentary (Pérez-Sabater, Peña-Martínez, Turney, & Montero-Fleta, 2008). Most studies of IM have focused on the social functions it serves for users (Bryant, Sanders-Jackson, & Smallwood, 2006; Flanagan, 2005; Jacobs, 2006, 2004; Lewis & Fabos, 2005; Ling & Baron, 2007; Nardi, Whittaker, & Braden, 2005; Quan-Haase, 2008; Thurlow, 2006), and most of this research has relied on human subjects research (observations, interviews, and other forms of users’ self reporting of their behavior). For example, Bryant, Sanders-Jackson, and Smallwood (2006) compared participants’ online and face-to-face relationships through a survey of 40 teenage IM users. Using a similar methodology (a web-based survey of 274 college students), Flanagan (2005) drew conclusions about the social functions of IM, writing that “in its short lifespan, IM has become a central communication tool within this population, being used to satisfy a variety of needs, and utilized quite heavily relative to other forms of mediated communication” (p. 184). Lee (2006) analyzed IM users’ interviews and logbooks to show how perceived affordances influenced choice of writing system for bilingual Chinese writers in Hong Kong.

Some previous studies have, like ours, examined the language features of IM (Baron, 2004; Hult & Ritchins, 2006; Nastri, Pena, & Hancock, 2006), but our study differs in at least three ways. First, our study included a larger sample (103 participants) than did Baron (18 participants), Hult and Ritchins (17 participants), or Nastri and her colleagues (44 participants). Second, we studied “naturally
occurring” IM sessions; that is, participants in our study did not instant message expecting that their conversations would later be studied. Finally, the goal of our project is to describe and taxonomize the full range of features of IM. For other researchers, description was less comprehensive in that their research goals were comparative (Baron compared genders; Hult and Ritchins compared spoken discourse, written discourse, and IM) or part of a project to demonstrate the usefulness of theory (Nastri, et al., illustrated the use of Searle’s speech-act theory).

Methodology

Over the course of two years, our research team conducted a replicable, aggregable, and data-based analysis (Haswell, 2005) of the language features of a set of IM conversations produced by college students. We collected transcripts of 54 instant message sessions, involving 103 participants (some participants were involved in more than one IM session). The sessions ranged in length from 6 entries to 563 entries and averaged 107 entries per session, with an average of 7 words per entry. The total size of the corpus was approximately 32,000 words. Printouts of archived IM conversations were solicited from college students in two sections of first-year composition and a senior English seminar at a large, residential, public university in the Midwest where the student body is predominantly white (82% white, 8% African American, 4% unidentified, and less than 1% Asian or Hispanic). Sixty-eight percent of the student population is female, and the average age of students is 23. Most students come from middle- or working-class backgrounds, and a significant number are first-generation college students. Our data set, thus, is the writing of people of a certain age, socioeconomic context, educational background, and place in life. We asked students to share transcripts from conversations that had taken place prior to our invitation that they participate in our study, and we provided written instructions on printing archived IM conversations. This was important in order to avoid the possibility that participants would write IM transcripts with us as writing researchers in mind and provide us with what they guessed we wanted—either more “correct,” academic-like writing or more exaggeratedly idiosyncratic IM writing.

Our research team has been variously configured over the three years of this project: two core members of the team are tenured faculty, three have been advanced writing studies graduate students, and four have been undergraduate English majors. At all stages in the research project, from study design to drafting various manuscripts for publication, at least three research team members worked together. Our approach might be characterized as one based in some of the tenets of grounded theory, such as immersion in data, constant comparison, and collaboration via research teams (Glaser & Strauss, 1967; Strauss, 1987). An inductive, grounded theory-based approach begins with the language features present in the data itself, drawing from those features the categories which describe
them. We analyzed the entire corpus over several months’ time, moving iteratively from the developing taxonomy to the data and back again, reading and rereading the transcripts until we identified particular language features that we found with some frequency, giving those features names, and returning to the transcripts to code for those features. Once we had arrived at a working taxonomy, our research team moved systematically through each line of the corpus, identifying and coding every occurrence of each of the taxonomical features. In short, by repeatedly revisiting the data, we were able to develop a taxonomy of IM features that accounted for—and did justice to—the characteristic features of our corpus.

In our study, we identified 15 language features characterizing IM writing. We found IM writing to be systematic in that we were able to identify these 15 features with regularity across the corpus: we identified 5,502 occurrences of these IM language features within the 32,000-word corpus. The entire corpus was coded by at least two raters. Inter-rater reliability was checked on a subset of 16% of the data; there were 22 disagreements between two raters, for a simple reliability estimate of 98.5%. In addition, however, two or more coders examined every disagreement in the coding of the entire corpus and determined an agreed upon coding. In this way, we developed a data set of coded transcripts, wherein each IM feature had been coded the same way by two or more raters.

**Findings**

Our taxonomy of instant messaging features is presented in Table 1. In the following section, we briefly define each category and provide examples of the category from the corpus; we then discuss the frequency of these categories in order to establish how these language features contribute to the unique character of instant messaging.

The fifteen-item taxonomy provides a descriptor of every textual feature in the data corpus that is characteristic of IM. The taxonomy moves from the specific to the global—moving from the use of individual characters (punctuation and letters) to larger units such as words, dialect, and metadiscursive markers, which are not understandable apart from several lines or more of surrounding text. (Therefore, Table 1 does not include an example of meta-markings, although there are examples of these in the discussion below.)

**Punctuation**

The first two categories in the taxonomy concern punctuation. Specifically, repeated and non-conventional punctuation appear in IM as ways to indicate a kind of pausing and to indicate emphasis (as in, *what model??????*). Repeated punctuation indicating pausing and emphasis can be seen in: *goodnight...sweet dreams!!!* Punctuation was applied in order to emphasize or communicate an alternative meaning, as in this play on words: *that was the “SUCK”iest “COCK”tail party ive ever been to.* The quotation marks here (in combination with the capitalization) suggest the
sentence be read as: “not only was that was the worst party I’ve ever been to, but it sucked cock.” In addition, our data reveal some implicit “rules” for punctuation in IM. Specifically, ending punctuation, punctuation separating independent clauses, and apostrophes are virtually nonexistent across the data set.6

**Letters**

Conventions concerning letters in IM take three forms: dropped letters (e.g., *tho, eatin, bout*), replaced letters (as in *they try to force s**t on u* and *assumptions are a b**ch*), and repeated letters (i.e., *haha shuttt uppppp*). Dropped letters reduce

---

**Table 1: Taxonomy of Instant Messaging Features**

<table>
<thead>
<tr>
<th>Category</th>
<th>Language feature</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Punctuation</strong></td>
<td>1. punctuation indicating pausing</td>
<td>i have some liquor…and some beer..but.</td>
</tr>
<tr>
<td></td>
<td>2. punctuation indicating emphasis</td>
<td>How bout u let me know when u're ready?!; he just…LOOKED…at me.</td>
</tr>
<tr>
<td></td>
<td>3. non-conventional punctuation</td>
<td>it’ll be the “SUCK”iest “COCK”tail party</td>
</tr>
<tr>
<td><strong>Letters</strong></td>
<td>4. dropped letters</td>
<td>i’ll see u in a lil</td>
</tr>
<tr>
<td></td>
<td>5. repeated letters</td>
<td>Ummmmm no; riight</td>
</tr>
<tr>
<td></td>
<td>6. replaced letters</td>
<td>WOW wut r u a f**kin teacher lol</td>
</tr>
<tr>
<td><strong>Words</strong></td>
<td>7. numbers for words</td>
<td>b4 i knew mike; don’t lie 2 me</td>
</tr>
<tr>
<td></td>
<td>8. letters for words</td>
<td>yea that’s y im doing it over thanksgiving</td>
</tr>
<tr>
<td></td>
<td>9. symbols for words</td>
<td>@; ?? [in place of typing <em>at or what or why</em>]</td>
</tr>
<tr>
<td></td>
<td>10. abbreviations</td>
<td>wat’s up w/u seriously?</td>
</tr>
<tr>
<td></td>
<td>11. slang</td>
<td>Yeah dude</td>
</tr>
<tr>
<td><strong>Dialect</strong></td>
<td>12. eye dialect related to words</td>
<td>get outttttta here</td>
</tr>
<tr>
<td></td>
<td>13. eye dialect related to sounds</td>
<td>Ah man; hahahaha</td>
</tr>
<tr>
<td><strong>Metadiscursive markers</strong></td>
<td>14. emoticons</td>
<td>:-P what?</td>
</tr>
<tr>
<td></td>
<td>15. meta-markings</td>
<td>these items are so context-dependent that single instances cannot be meaningfully excerpted; see discussion for examples</td>
</tr>
</tbody>
</table>
the number of keystrokes required to produce the word, repeated letters increase the number of key strokes, and replaced letters exchange some keys for others, often leaving the total number of keystrokes the same but sometimes reducing the number (for example, in *thanx* where *x* replaces *ks* or *wut* where *u* replaces *ha*). Repeated letters were almost always double-coded as instances of eye dialect, a category we describe below to capture nonstandard spelling that draws attention to pronunciation and dialect. The repeated letters in an entry like *hahaha awww funnn!!!*, for example, were double-coded as repeated letters and as eye dialect related to words (*funnn*) or eye dialect related to sounds (*hahaha, awwww*). In addition, as with the punctuation category, our data reveals at least one implicit “rule” involving letters: initial capitals (to begin sentences and in proper names) are rarely used across the corpus.

**Words**
The “words” category captures slang and four ways that spellings of words are manipulated in IM: numbers for words, letters for words, symbols for words, and abbreviations. Numbers, letters, and symbols for words replace entire words with single typographic symbols (2 = to, too; 4 = for; u = you; c = see; or ? = who, what, when, or where) and additionally, numbers are used to replace clusters of letters and combine with letters to form words such as *l8r*. Some abbreviations appear to have genesis in the Internet (*ttyl* = talk to you later, *lol* = laugh out loud) while others predate the Internet (*b/c* = because, *hr* = hour). In the “words” category, individual characters are used to represent the whole word itself, thus representing whole words in the briefest of forms—with single letters, numbers, or typographic symbols. Slang in IM ranges from phatic words sprinkled throughout the language (i.e., *ok, yeah, cool, like*) to standard swear words (*damn, shit*, and the like) to references to others (*missy, babe, goof, hun*) to more idiosyncratic expressions (*yeehaw*).

**Eye Dialect**
This category includes two types of features similar to a literary device known as “eye dialect.” Mainly found in creative non-fiction and fictional writing, eye dialect is deliberate, non-standard spelling used to suggest pronunciation or, more precisely, to draw attention to associations between a character’s ways of pronouncing words and certain regional and cultural dialect (we discuss our use of “eye dialect” more fully in the discussion section below). Instances of eye dialect take two forms: eye dialect related to words and eye dialect related to sounds. Instances of eye dialect related to words include those that add, subtract, or replace letters and/or syllables to communicate a non-standard use of an existing word: as in *saweet* (sweet) and *partay* (party); *hidin* (hiding), *k* (ok), and *gonna* (going to); and *wanna* and *lil* (little). Occasionally, these appeared with apostrophes that suggest the dropped letters, as in ’k (ok) or *goin’* (going). Instances of eye dialect related to words were always double-coded, as both eye dialect related to words (taxonomy feature 12) and dropped letters
(feature 6), repeated letter (feature 5), or replaced letter (feature 6). A second form of eye dialect relates to sounds such as aw, ba-boom, ugh, and hmmm. Most of the items in this category were of a conventional type, and we found few sound words that were new or unfamiliar.

**Metadiscursive Markers**

Metadiscursive markers are used by writers in IM to step outside the conversation to make a comment on the writing—either to establish or suggest the tone of the utterance or conversation (through emoticons) or to make a comment on the writer’s writing (through meta-markings). Typical emoticons in the sample include:

- :-) (smile)
- :-( (frown)
- ;-) (wink)
- =:0 (surprise or fear)

They usually occurred at either the beginning or the end of an entry to indicate the tone of the utterance. Meta-markings include the use of a symbol (asterisks, quotation marks, or slashes) to indicate an “aside” to the previous comment, as in the example below in which Annie uses an asterisk to “step outside” the exchange to correct a spelling error.

**ANNIE:** I have to type up my academic awars

**RYAN:** lol

**ANNIE:** *awards

Meta-markers can also indicate revision of a slightly more global nature:

**SHERRY:** :-) too bad for u, u dont have another door /room to disguise behind ;-

**BRUCE:** 36 is barely ahead of 60 [refers to another conversational strand co-occurring]

**SHERRY:** *hide behind

With this meta-marking, Sherry indicates a word substitution that is presumably more precise than the meaning her original words suggested: hide behind in place of disguise behind. She has entered text, rethought it, rewritten it, and indicates the textual substitution through the use of a meta-marking, here again the asterisk.

**Frequency of IM Features**

In order to develop a rich understanding of IM, it is important to look not only at the types of writing features but also at frequency patterns of the features. It is here that we can see the usages that characterize IM writing. A closer examination of the most and least frequent language features of IM suggests both the unique nature of this writing while also revealing the representative nature of IM (and in this way, countering the popular media tendency to focus on infrequent but startling
features, of the writing to represent it as a whole). Although every conversation in our corpus included IM features, the frequency of features appearing in the 54 individual IM conversations ranged vastly: 37 conversations had between 4 and 100 features, 17 conversations had between 106 and 310 features, and one conversation had more than double any other conversation with 789 features. The briefest number of features occurring in any conversation was 4 features in one conversation of 36 words, while the most features appearing in any one conversation was 789 features occurring in one conversation of approximately 3,300 words (although we did not ask participants to include time stamps, conversational turns in this transcript were time stamped, so we know that this conversation took place over the course of four days with some lengthy breaks between turns). Frequency patterns for the 5,502 coded IM features across the 32,000-word corpus are shown in Table 2, including both the number of occurrences and the proportion those occurrences represent.

As Table 2 shows, the most commonly occurring feature of IM is slang (at 18% of the total IM features). The next three features (eye dialect related to words, abbreviations, and letters for words) occurred with equal frequency (each approximately 13%) and the fifth and sixth most commonly occurring features (repeated punctuation indicating pausing and eye dialect related to sounds) occurred with

<table>
<thead>
<tr>
<th>IM Feature</th>
<th>Occurrences</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>slang</td>
<td>1004</td>
<td>18%</td>
</tr>
<tr>
<td>eye dialect related to words</td>
<td>700</td>
<td>13%</td>
</tr>
<tr>
<td>abbreviations</td>
<td>695</td>
<td>13%</td>
</tr>
<tr>
<td>letters for words</td>
<td>658</td>
<td>12%</td>
</tr>
<tr>
<td>punctuation indicating pausing</td>
<td>550</td>
<td>10%</td>
</tr>
<tr>
<td>eye dialect related to sounds</td>
<td>503</td>
<td>9%</td>
</tr>
<tr>
<td>dropped letters</td>
<td>392</td>
<td>7%</td>
</tr>
<tr>
<td>punctuation indicating emphasis</td>
<td>350</td>
<td>7%</td>
</tr>
<tr>
<td>emoticons</td>
<td>299</td>
<td>5%</td>
</tr>
<tr>
<td>repeated letters</td>
<td>188</td>
<td>3%</td>
</tr>
<tr>
<td>numbers for words</td>
<td>52</td>
<td>&lt; 1%</td>
</tr>
<tr>
<td>symbols for words</td>
<td>46</td>
<td>&lt; 1%</td>
</tr>
<tr>
<td>meta-markings</td>
<td>35</td>
<td>&lt; 1%</td>
</tr>
<tr>
<td>non-conventional punctuation</td>
<td>16</td>
<td>&lt; 1%</td>
</tr>
<tr>
<td>replaced letters</td>
<td>14</td>
<td>&lt; 1%</td>
</tr>
</tbody>
</table>
equal frequency (approximately 10%). Six features of IM (slang, eye dialect related to words, abbreviations, letters for words, punctuation indicating pausing, and eye dialect related to sounds) represent three-quarters of the IM features. In other words, a small group of features give IM the majority of its character. At the other end of the frequency chart are five features—symbols for words, meta-markings, numbers for words, non-conventional punctuation, and replaced letters—that together comprise less than 5% of the total occurrence of features. It is interesting to note that two of these relatively rare features—numbers and symbols for words—are often used to characterize IM in popular discourses although in our study, these two features together appear in less than 1% of a 32,000-word corpus.

Another pattern revealed by frequency counts is a tension between abbreviated and elaborated forms. Instant messaging is seemingly made more efficient and briefer through several of the language features: replaced words with symbols, letters, and numbers; dropped letters; and abbreviations. 4 rather than four requires three fewer keystrokes, c rather than see requires two fewer keystrokes, and ? rather than what or why or where or when requires two to four fewer keystrokes. In these ways, then, IM can partly be characterized as an abbreviated form. However, IM is, paradoxically, an elaborated type of writing. Of the six most frequently occurring features (slang, eye dialect related to words, abbreviations, letters for words, punctuation indicating pausing, and eye dialect related to sounds), only two are shortened forms of writing: abbreviations and letters for words. The remaining four most frequently occurring features (slang, eye dialect related to words, punctuation indicating pausing, and eye dialect related to sound) as well as three other features (emoticons, repeated letters, and meta-markings) are all additive, increasing the length of IM entries: instead of simply no, six additional keystrokes are involved in ummmmmm no. In fact, while shortened forms (numbers, letters, and symbols for words; dropped letters; and abbreviations) accounted for 1,857 of the total features, fully twice as many features require more, rather than fewer, keystrokes (see Table 3).

In other words, 67% of IM features elaborate on the writing, while 33% abbreviate the writing. Interlocutors in IM abbreviate their writing half as often as they work to insert extra features into the writing. The frequency count in Table 3 counters common sense and popular notions that IM is simply an abbreviated type of writing or that writers using IM are motivated by a desire for brevity or speed of communication.

In fact, our analysis suggests that brevity and speed are not of primary importance for these writers but that instead, a focus on brevity and speed, in the case of our corpus, would focus erroneously on 33% of IM features (those identified in Table 3 which subtract keystrokes) rather than 67% of features (those in Table 3 which add keystrokes). Interestingly, the 67% of features in our corpus which add to, rather than subtract from, the writing are ones that inscribe what have
largely been oral features of language use into the written conversation. The ad-
dditive features we have identified in Table 3—slang, eye dialect related to words, punctuation indicating pausing, eye dialect related to sounds, repeated letters, emoticons, punctuation indicating emphasis, and meta-markings—are all extra-
linguistic in the sense that they serve purposes other than informational ones. That
is, the majority of IM features in our corpus (67%) are ones which are
related to paralinguistic inscription. Some of these features (slang, emoticons, and meta-markings) can provide paralinguistic information on how to “take” the meaning of the writing. Some of them (eye dialect related to words and sounds, punctuation related to pausing and emphasis, and repeated letters) provide paralinguistic signals related to how one might “hear” the writing—intonation, pronunciation, and pauses. These paralinguistic markers are not linguistic markers indicating literally how, for example, one pronounces the words, but rather, these paralinguistic signals seem related to how one might “read” the writer herself—as a creative, playful, innovative insider familiar with the characteristic uses of IM’s language features. Indeed, the features of IM identified in our taxonomy (including the subtractive features of IM: numbers, letters, and symbols for words; abbreviations; and dropped and replaced letters) often serve important paralinguistic functions involving how to read the writer, the context, or the message itself. In the following section, we discuss the paralinguistic features of IM, in particular focusing on four features: eye dialect, slang, emoticons, and meta-markings.

<table>
<thead>
<tr>
<th>Features which add keystrokes to the production of the writing</th>
<th>Features which subtract keystrokes from the production of the writing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1004 slang</td>
<td>695 abbreviations</td>
</tr>
<tr>
<td>700 eye dialect related to words</td>
<td>658 letters for words</td>
</tr>
<tr>
<td>550 punctuation indicating pausing</td>
<td>392 dropped letters</td>
</tr>
<tr>
<td>503 eye dialect related to sounds</td>
<td>52 numbers for words</td>
</tr>
<tr>
<td>326 repeated letters</td>
<td>46 symbols for words</td>
</tr>
<tr>
<td>299 emoticons</td>
<td>14 replaced letters</td>
</tr>
<tr>
<td>212 punctuation indicating emphasis</td>
<td></td>
</tr>
<tr>
<td>35 meta-markings</td>
<td></td>
</tr>
<tr>
<td>16 non-conventional punctuation</td>
<td></td>
</tr>
<tr>
<td><strong>3,645 TOTAL</strong></td>
<td><strong>1,857 TOTAL</strong></td>
</tr>
</tbody>
</table>
Discussion: The Paralinguistics of Instant Messaging

The paralinguistic may be appearing in written form (perhaps for the first time) in IM as a result of the unique convergence of technological and interpersonal factors characterizing IM: IM is *mediated* by a keyboard, screen, and network; it is *distanced*, in that participants are typically separated in space (or, if in a shared space, are working on separate computers); and it is *synchronous*, or nearly so. IM is *private*, in that it is not publicly available. IM is also typically a *one-to-one* form of communication. It is not one-to-many like an oration or a published article. IM participants are *known* to one another and usually have shared knowledge and shared experiences outside of the IM context. Finally, IM exchanges are *participant driven*; that is, topics are not pre-established like they would be in context-specific chat rooms, specialized publications, or academic writing, nor are participants’ contributions planned acontextually of the conversation as it develops.

Some of these features are ones that attach to writing, while others attach to speech. IM shares with oral communication several important characteristics: both are synchronous, they are generally private, between a small number of known participants, and the exchanges are participant driven. The immediacy of the communication and the shared time and (virtual) space of IM conversations are suggestive of speaking situations. But there are significant similarities between IM and written communication, as well: both are mediated by technology, distanced, recorded/archived, and rely on typographic symbols. Thus, participants in our study were using writing in a communicative context reminiscent of speaking. We believe that when faced with this communicative challenge, participants drew on resources familiar to them in oral contexts and gave those written form. Collins and Michaels (1984) argued that “…in the case of written versus spoken communication there are channel constraints, that is, conditions on communicative form which are derivable from the nature of the medium” (p. 247). The channel constraints faced in IM include little information about intonation in written language form, little information about the audience’s immediate reception, and no paralinguistic cues. Interlocutors in IM and TM must compensate for these channel constraints—thus the innovations and frequency of those innovations in our data corpus.

Consider the following excerpt from our corpus:

**Bill:** haha i dont know, it just sounds weird having you call me that
**Bill:** its too informal for my taste
**Bill:** haha

**Sharon:** WOW wut ru a fuckin teacher lol

**Bill:** hey we dont use the f bomb
**Bill:** i was just joking
In this conversational exchange very little informational content is shared. Rather, the entire exchange is a social exchange, and what content is imparted is primarily social content—how Sharon and Bill each want to be understood, what they each view as humorous, what the social rules of their relationship will be. In IM, Sharon and Bill cannot draw on the range of social cues so necessary and common in face-to-face spoken communication. But this conversational exchange might be seen as an attempt on Sharon and Bill’s parts to incorporate into writing (or to at least account for the loss of) the paralinguistics of face-to-face communication. In this case, the interlocutors draw on those familiar resources from another context (spoken communication) and (literally) write them into the conversation. Extra-linguistic features (such as *haha* and *WOW*) and the explictive, *fuckin*, are not essential to the literal meaning of the words, but they communicate how the words are meant to be interpreted in social context. Further, in this excerpt, slang words like *WOW*, *bitch*, and *fuckin* play a social role evidenced in slang usage across the corpus as a whole: most slang expressions can be removed without changing the literal meaning of an exchange. That is, across the corpus, slang served a phatic rather than an informational purpose. Largely understood in the context of oral speech acts, phatic expressions are speech acts that serve social purposes (such as establishing a mood or setting a tone to one’s interaction) as opposed to conveying information. The ubiquity and phatic quality of slang suggests that its role in signaling social affiliation and interpersonal bonds is quite important to IM participants.

Gumperz and Berenz (1993) have described paralanguage as “the nonlexical components of communication by speech, for example intonation, pitch and speed of speaking, hesitation noises, gestures, and facial expression” (p. 95, emphasis added). Our analysis of IM writing, however, suggests that the lexical/non-lexical distinction drawn in existing conceptions of paralanguage may break down in technologically mediated communications such as instant messaging. Our study shows evidence of writers appropriating paralinguistic features—the intonation, pitch and speed of speaking, hesitation noises, gestures, and facial expressions Gumperz and Berenz (1993) describe as non-lexical components of speech—and literally inscribing them into the written language of IM. Written features of IM such as punctuation indicating pausing or emphasis, repeated or dropped letters, slang, eye dialect, and meta-markings all function paralinguistically to provide readers with cues as to how the writing is to be understood. That is, the language features of IM most often serve to establish the tone of the exchange as a whole, to attach emotional meaning to utterances, and to smooth the communication. By writing into the language (in often innovative and creative ways) paralinguistic
cues, the participants in our study work to clarify, or more precisely disambigu-ate, meaning. Looking more closely at four of these features—eye dialect, slang, emoticons, and meta-markings—reveals how the paralinguistic is inscribed in IM’s language features.

**Eye Dialect**
Two features on our taxonomy involved “eye dialect,” a term we borrow from literary studies to suggest writing that is used to capture spoken dialogue (usually of characters in novels and other fictional work). In our corpus, eye dialect took two forms: eye dialect related to words and eye dialect related to sounds. In the first, identifiable words were spelled in non-traditional ways which were suggestive of intonation or pronunciation. Items in this category involved the combination of two existing words or the addition, deletion, or substitution of letters.

In the second form, eye dialect related to sounds, sounds that do not have a fixed traditional spelling, were written into the conversation. Perhaps because there is no traditional or standard spelling for the sounds in the corpus, such as laughter, the spelling varied in choice of letter arrangement and in repetition of letters.

As both Table 4 and Table 5 show, eye dialect features are very often related to the addition, subtraction, and substitution of letters and syllables in words, and thus, instances of eye dialect were often double-coded (as in *hidin*, for example, which was coded as an instance of eye dialect related to words and dropped letter). Eye dialect relating to words and eye dialect relating to sound comprise 22% of the total coded features.

Walpole (1974) described eye dialect as a writing task undertaken by a writer who “tries to distinguish dialogue from expository or narrative prose, to reproduce the rhythms of speech, to differentiate among his characters, and to do all this in as natural and authentic a way as the conventions of print will allow” (p. 191, emphasis)

**Table 4: Examples of Eye Dialect Related to Words**

<table>
<thead>
<tr>
<th>Two words combined</th>
<th>wanna, gonna, gotta, kinda, shoulda, sorta, dunno</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dropped letters &amp; abbreviations</td>
<td>sittin, cuz, prolly, gnite, lil, bout, tryin, kickin, keeping</td>
</tr>
<tr>
<td>Added letters</td>
<td>saweet, oooooookkk, aannnnytime, screeeeam, reaaaaallly?, sooo</td>
</tr>
</tbody>
</table>

**Table 5: Examples of Eye Dialect Related to Sounds**

| hahaha, hehe, ummm, awww, ooohh, oooo, eeeek, evww, ugh, phew, oops |
added). This understanding of eye dialect as a writing task suggests that IM is not best understood as a hybrid of speech and writing, but rather as writing itself. IM is writing in which paralinguistic features are inscribed and reproduced within the affordances of the conventions of print. Of networked writing generally, Crystal (2006) wrote, “Netspeak has far more properties linking it to writing than to speech. . . . Netspeak is better seen as written language which has been pulled some way in the direction of speech than as spoken language which has been written down” (p. 47). Our study supports this sense of written language that has been pulled some way in the direction of speech, rather than spoken language that has been written down. Additionally, while eye dialect might be understood as an attempt by writers to reproduce the rhythms of speech and pull writing in the direction of speech, it is important to note that eye dialect is not speech written down. As Walpole (1974) noted, eye dialect is quite distinct from linguistic transcription: “readers have been so well conditioned to the conventions of dialogue that accurate transcripts would dismay them” (p. 191). For example, in the case of wanna, the first and second vowels (both represented by “a”) are, in fact, different phonemes and the “n” would not be doubled (that is, not elongated). Eye dialect, that is, evokes informal speech, but it does not render oral features into written form, in the sense that 1) most oral features are not represented or are ignored in eye dialect, 2) features of oral production that are represented are represented inconsistently, and 3) some eye dialect features, like the double “n” in wanna or gonna, have no corollary in oral production. Eye dialect does not work, then, to bring together writing and speech; instead, it works as a cue to the paralinguistic cues that are absent in purely written contexts.

Slang

Linguists have defined slang in numerous ways, although nearly all understandings of it are grounded in oral language usage. Indeed, Labov (1992) pointed out that “a difficulty in studying slang terms stems from the fact that slang is generally spoken and rarely written down” (p. 347). Dumas and Lighter (1978) identified three characteristics of slang: it lowers the dignity of formal or serious speech or writing; its use implies the user’s familiarity with the class of people who use the term; and it is a taboo term in ordinary discourse with persons of higher social status or greater responsibility. Dumas and Lighter (1978) argued that slang is any expression which meets at least two of these criteria, the first of which is central—if not crucial—to slang.

Slang functions, then, as an expression of informality, of group membership, and of opposition to established authority. Slang typically arises from youth culture, particularly youth subcultures, and it functions within these groups as “simultaneously an act of featuring and obtruding the self within the subculture—by cleverness, by control, by up-to-dateness, by insolence, by virtuosities of audacious and
usually satirical wit, by aggression” (Chapman, 1986, pp. xii–xiii). In our study, we applied the term “slang” very broadly to include any expressions that fall under these descriptors. Our study identifies the presence of slang as a persistent feature in IM writing (as opposed to cataloging the vocabulary of slang as it appears in IM). Slang appeared 435 times across the corpus, comprising 18% of the IM features in our study’s corpus. Notably, slang appeared in every conversation (ranging from 1 occurrence per conversation to 66 per conversation) and thus, would seem to be a conventional feature of IM. The majority of slang (81% of slang occurrences) took one of three forms: seven specific commonly used expressions (ok, yeah, like, cool, wow, nope), profanities, and terms naming people. The remaining 19% of slang expressions appeared less than two times across the corpus and were coded as miscellaneous expressions.

Slang in the commonly used expression category account for 60% of the total slang expressions. They seem to function primarily to move the writing along rather than to impart any specific information, while profanities (such as shit, suck, ass, damn) were most often used as adjectives. Terms naming people took the form of direct address (goof, babe, hun, kiddo) as well as reference to self and others (jerk, freak, roomie, dork). The pervasiveness of slang in these transcripts of IM conversations (occurring in every conversation anywhere between once and 66 times) is suggestive of a common way writers draw from traditionally oral resources to pull the written communication in the direction of the oral.

**Emoticons**

Emoticons only comprised approximately 5% of IM features in the corpus, but given their prominence in characterizations of IM and other forms of digitally mediated writing, we turn to a closer examination of them as an instance of the paralinguistic in IM writing. Whereas emoticons might appear to be a visual feature

<table>
<thead>
<tr>
<th>Type of slang</th>
<th>Frequency</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commonly used expressions</td>
<td>262</td>
<td>ok &amp; its variants, yeah &amp; its variants, hey, cool, like, wow, nope/nah</td>
</tr>
<tr>
<td>Profanities</td>
<td>50</td>
<td>shit, suck, ass, damn, fuckin</td>
</tr>
<tr>
<td>Terms naming people</td>
<td>44</td>
<td>goof, whore, babe, hun, jerk, roomie</td>
</tr>
<tr>
<td>Other miscellaneous expressions</td>
<td>79</td>
<td>alrighty, awesome, freaked out, gee, iunno, peppy, problemo, totally, twisted</td>
</tr>
</tbody>
</table>


situated within a verbal context, they are actually a kind of visual representation that is composed of numerals and orthographic symbols from the computer keyboard. Emoticons use elements of the writing system (symbols, letters, and numbers) to create visual, representational characters meant to mimic the human face and capture what is communicated paralinguistically in face-to-face communication. Again, interlocutors are drawing on resources from one realm (face-to-face spoken communication) and literally writing them into the conversation. In addition to the familiar smiley, frown, and wink emoticons, our participants used the angel emoticon [0:], the frightened emoticon [=–0], the sticking-out-the-tongue emoticon [:–P], and the laugh emoticon [:–D], among others. The emoticon is a particularly intriguing feature in light of the qualities reminiscent of speech in this written form. The emoticon is designed to convey the paralinguistic and non-verbal cues present in speech—smiling [:–)], frowning [: ]-, indicated surprise [=–O]—to indicate how the written message is to be taken. In this way, they too suggest the immediacy of speech. Yet these paralinguistic indicators are carefully constructed using conventional signs of writing—punctuation marks and other textual indicators—as these are manifest on a QWERTY keyboard. As such, emoticons nicely gather up many of the complexities of the speech-writing dyad.

**Meta-markings**

While meta-markings were not a prevalent feature in our corpus, they are an interesting and unique instance of the impulse to inscribe the paralinguistic into written form. Metadiscursive markings are written features used to “comment on” the utterance as a whole. In our corpus, the most common meta-marking was the use of the asterisk to step outside of and comment upon a previous utterance.10 Sometimes asterisks point attention back to something that had been typed incorrectly, as in this example:

**BRAD:** its not your fault i let myself get to attached
**BRAD:** too*

Brad’s entry, too*, shows both his recognition of the spelling error and his correction of it. Additionally, meta-markings can also be innovative indicators of intent, as in:

**AARON:** u’re place or mine;-)
**AARON:** (hoping u’ll say yours
**AARON:** cuz mine’s disgusting)*hint

and

**CALISTA:** *warning Bette* I was acting totally. . .naughty lol. sorry.
Here, the meta-marking asterisk signals how the entry is to be “taken,” with its illocutionary force as a hint or as a warning. In this way, meta-markings further disambiguate meaning not only at the word level—as in spelling—but at the rhetorical level as well. Interestingly, in this way, writing becomes a self-referencing artifact: through the features of the writing, the text reflects back on its own meanings and intentions.

These features of IM suggest that rather than seeing IM as a hybrid of speech and writing, IM is writing: as Crystal (2006) said, writing pulled in the direction of speech, or, as we have argued, writing which attempts to inscribe the paralinguistic. Our taxonomy of language features as well as the frequency of those features suggests the many innovative, creative ways some writers are using written language to communicate and to reappropriate paralinguistic features of spoken communication into their written language. IM can be seen as an instance of people communicating meaningfully in writing while constrained by what have usually been demands of an oral situation. Gumperz and Berenz (1993) said that successful conversation rests on: “. . . speaker’s and listener’s interpretation of verbal and nonverbal signs or contextualization conventions, that is, systems of cues that guide conversational management [in] such matters as topic change or change in perspective, relative emphasis or salience of information, style shifting, degree of formality and the like, or to achieve the kind of rhythmic integration that often marks smooth interaction” (pp. 91–92). In order to achieve the rhythmic integration leading to smooth interaction that Gumperz and Berenz (1993) described, IM writers must anticipate and respond to actual or supposed listener reaction, include paralinguistic signals on how to “take” an utterance, and write in the social niceties that move a conversation along (and the development of an IM conversation, unlike other familiar written forms, is dependent on turn taking). These reappropriations and rewritings are playful but that does not lessen their importance to the success of the utterance; instead, their playfulness is in itself an important way the paralinguistic is written into the written text. IM writers use the conventional signs of writing to represent in accepted cultural ways (i.e., youth culture, Internet culture) some of the immediacy of speech.

Conclusion

Our research focus on the language features of IM (rather than the social uses of IM as a literate practice) is an important undertaking in a culture where popular conceptions of IM and other mediated writing tend to run to the damning. As Crispin Thurlow (2006) has so convincingly demonstrated, the dominant popular discourse has characterized IM and texting as ruining the English language. Reviewing 101 print-media accounts of new media (such as IM and texting) published between 2001 and 2005, Thurlow (2006) concluded that “one major narrative thread
in public discourse about emerging technologies involves concerns about the way language is affected . . . and any perceived threats to conventional or standard language practices.” Through his analysis, Thurlow (2006) suggested that “it appears that language and technology is (once again) not only being poorly represented, but also scapegoated for a range of adult anxieties about newness, change, and perceived threats to the status quo” (p. 689). These anxieties about newness and change fuel the attempt in many popular accounts to capture what these types of writing are; we fear that too often, these anxieties also fuel teachers’ perspectives on their students’ practices with instant messaging. We hope our study contributes to a more comprehensive and richer picture of what writers do at the level of the writing itself when they are faced with the challenge of communicating in digitally mediated, disembodied lifeworlds.

In the time since we first collected our data set, texting on mobile phones has quickly eclipsed IM as the most common computer-mediated communication among college students and mobile phones have evolved to support writing that more closely resembles standard written English (with full keypads as well as predictive text software that completes and automatically corrects spelling and grammar). Predictive text and SmartPhones with QWERTY keypads may indeed contribute to the eventual demise or significant revision of the language features we have identified here—but given the disembodiment of technologically mediated forms of written communication, inscribing the paralinguistic will likely remain a priority for writers. Indeed, in initial studies we have conducted on the language features of Facebook and texting, we have discovered many of these same features in writing in those contexts, as well. Our study then provides a jumping off point for further research in how technologies are not only mediating but potentially shaping written language features.

Teachers in a digital world need to pay attention to technologically mediated writing in which our students engage and build their lifeworlds, not only to understand our students but to understand what writing is now. What do our students do when they write? How are they building lifeworlds through writing? What can we do to prepare them for a world in which so much communication is literate? Our study suggests that we begin by understanding mediated forms of writing as writing: it is purposefully constructed of visible marks in graphic space (Harris, 1996; Bolter, 1991). We have suggested that the language features of IM are instances of written paralanguage—of attempts by writers to inscribe paralinguistic cues into the conventions of print visually, referentially, and phonetically. As the categories of our taxonomy—punctuation, letters, words, dialect, and metadiscursive markers—suggest, this inscription work is being done by writers in IM at the levels of the individual print symbols and letters as well as in commentary on the discourse as a whole. Beginning with writing allows us to contextualize our own work as writing teachers and researchers within an understanding of what writing
is now and how language changes. Crystal (2008) made the important point that healthy vital languages are always changing. Seen in this way, IM is hardly the ruin of the English language but rather a type of change that indicates the healthiness of written language (especially among the young people who are mostly doing the work of defining its parameters). This understanding of the inevitability and healthiness of language change is important, too, as it contextualizes the rapid changes some technologically mediated forms of language can take.

We began this article with a discussion of everyday literacies, the kinds of reading and writing practices that take place beyond the sanctioned walls of classrooms and halls of academe. Our project illustrates that a close, fine-grained look at language features provides one window into everyday literacy in its historical moment. The understanding that emerged for us was one of young people who have a true mastery of written language: the IM transcripts we studied were funny, clever, innovative, sometimes moving, and almost always delightful. We believe this picture emerged partly because of our inductive approach—to the best of our abilities, we let the language of participants speak to us and speak it did. Our study has revealed that literacy—at once innovative and playful, systematic and purposeful—is alive and well on the Internet and in the lives of the young people who use it. H. L. Mencken (1963) viewed slang as “a kind of linguistic exuberance, an excess of word-making energy” (p. 702). Mencken’s description of slang (a small part of IM) in some ways captures the spirit of IM writing as a whole. In a world where digital, multimodal communication is increasingly the standard practice, these users of IM are using the somewhat limited resources of the 68 typographic symbols on the computer keyboard to communicate visually, verbally, and bodily. The “linguistic exuberance” resulting from the creative adaptation of technological resources to an expansive understanding of textual features offers to say much to writing researchers and teachers interested in the evolution of written discourse.

NOTES
1. This research was funded by the NCTE Research Foundation, Kent State University Moulton Scholarships, the Kent State University Moulton Faculty Associates Program, and Kent State University Office of Research and Graduate Studies.
2. The contributions of the two lead authors were equal.
3. Our use of “word” is meant to capture not what a linguistic definition would identify as a word but a word-like unit of the language. For example, we counted 10 words in this IM turn: “nothing sittin around until i have 2 go to math” where a strictly linguistic definition would not see 2 as a word, and we counted 3 words in this IM turn: “yayyyyyyyy you ddi?!!!”
4. We focus in our analysis on the textual forms produced. Given our data set, we are not able to articulate the relationship between a writer’s background and the forms of her written language. However, given our sampling method and the larger population from which we drew (a predominantly white student body), it is very likely that the features of IM we describe in this article are
shaped in part by race, class, and purpose. Indeed, we agree with one reviewer of this article who suggested that “this does not mean that some of the features are not common across gender, or class, or race, but we cannot assume this given how languages and language varieties differ.” This reviewer suggests one direction for future researchers when she asks “Would IMing, for instance, at a historically Black college be the same?” Our current study cannot begin to answer that question, though we do suspect that the writer’s background and context influence the shape of the writing produced. Crystal (2006), for example, suggested that differences in a writer’s background and in the cultural uses of language might result in varying textual forms in digital settings when he described the more formal greetings in English language emails of non-native speakers as “suggestive of the existence of a pull toward traditional usage in non-native speaker settings” (p. 102). Future researchers might measure the influence of contextual factors, experiential knowledge, and cultural context of writers on IM.

5. The research team for this project has been a key feature of the methodology of the project and the analysis of the data. Two tenure-track professors (Haas and Takayoshi), three rhetoric and composition PhD students (Diana Awad, Emily Dillon, and Elizabeth Feltner), and four undergraduate students (Brandon Carr, Jessica Heffner, Kimberly Hudson, and Ross Pollock) have brought different institutional, intellectual, and experiential backgrounds at different times to the development of our taxonomy and understanding of the IM transcripts we analyzed. As active writing researchers, the tenure-track professors led the development of relevant research questions within the context of existing bodies of scholarship, guided students to readings on theoretical and methodological matters, led collaborative work on research design, data collection and analysis, and took the lead on the collaborative authoring of this journal article and a book chapter. Within that framework, undergraduates and graduate students worked as research team members, in particular bringing important insights to the research design, data collection, and analysis aspects of the research team. At the beginning of this study, IM was a widely used technology by college-aged students, although the faculty members of the research team had not used IM prior to the research. John Perry Barlow’s distinction between “native” and “immigrant” suggests in part the value of a cross-generational research team:

Barlow spells out a distinction between two mindsets that are brought to bear on cyberspace specifically and spaces of digitized practices more generally. Barlow refers to these two mindsets as ‘immigrant’ and ‘native’ mindsets respectively. We prefer to call them ‘outsider’ (or ‘newcomer’) and ‘insider’ mindsets respectively, since the terminology of ‘immigrants’ and ‘natives’ might reasonably be seen as offensive by members of some social groups. . . . Barlow’s distinction is between mindsets which related to how this space is constructed and controlled in terms of values, morals, knowledge, competence, and the like. Since ‘newcomers’ lack the experience, history, and resources available to them that ‘insiders’ have, they cannot—to that extent—understand the new space the way insiders do. On fundamental points and principles of cyber/information/virtual space, says Barlow, newcomers ‘just don’t get it’ (Lankshear & Knobel, 2003, p. 32).

The experiential knowledge native insiders can provide immigrant outsiders about the forms and functions of IM informs our analysis and interpretations, helping us understand the significance of this literate practice in young people’s lives. From initial involvement with undergraduates in our writing classes, we know that discussing emerging forms of communication technologies with younger users reveals facets and meanings of which we were not aware; as a profession of
writing researchers and teachers, we have much to learn about contemporary language practices through a close examination of new language forms and through the guidance and perspectives of the native users. Our methodology is described in more detail in both Haas, Takayoshi, and Carr (forthcoming) and Farkas and Haas (forthcoming).

6. Our taxonomy captures the presence, rather than the absence, of features in the corpus. Thus, missing end punctuation did not end up on the taxonomy, although we found it to be a pervasive feature of every conversation in our IM corpus. Similarly, initial capitalization was a pervasive feature across the conversations in our corpus. The only absence we coded was dropped letters, as that was an absence which left a trace that could be identified: for example, the dropped “g” in an “-ing” ending (droppin, sayin, talkin) was not present, but by virtue of it being the only letter in the word which was not present, its absence was coded as a deliberately dropped letter. The absence of ending punctuation, punctuation separating independent clauses, and apostrophes, on the other hand, left no partial trace of their potential presence.

7. While it is possible to IM with several participants, even in IMs among multiple participants, the discourse involves a limited and known number of participants. In our corpus, all conversations were one-to-one involving only two interlocutors.

8. Interestingly, the concept of eye dialect has also appeared in methodological discussions of the transcription of verbal data. Gumperz and Berenz (1993), for example, referred to eye dialect as one solution researchers draw upon when faced with “problems in capturing significant features of style and dialect that may need to be represented” (p. 96). They wrote, “some researchers use what is in effect a pseudophonetic form for certain words in the lexical stream while maintaining conventional orthography for others in what appears to be an insufficiently motivated way, so that in one transcript the word is might be spelled in the normal way, but the word because is spelled b’cuz. This approach has been called eye dialect due to its comic-strip-like effect” (p. 96). Our use of eye dialect is contextualized within literary studies in part because participants in our study were likely more influenced by conventions for oral speech presented in comic books, novels, and other fictional work than by conventions adopted by linguistic researchers. We are grateful to Paul Prior for this insight.

9. Crystal (2008) has described IM as “typographically idiosyncratic” but while we found typographical innovations especially in the eye dialect categories, these were not idiosyncratic to individuals, but rather conventionalized across the 32,000-word corpus. Indeed, most of the items in this category were of a conventional type, and we found few sound words that were new or unfamiliar.

10. The use of the asterisk may be a trace of synchronous chatroom discussion, particularly MOOs and MUDs. In those earlier online discourse forms, participants regularly entered extra-linguistic data about their actions using a software command, which would appear in the MOO/MUD alongside an asterisk as a narrative description of them. So if a writer named Becky were to enter:

```
/me walks into the room and smiles at everyone
```

It would appear on everyone else’s screens as:

```
* Becky walks into the room and smiles at everyone.
```

Although asterisks were the most common form for a meta-marking, asterisks were also used to indicate emphasis, as in:

```
Lu: hey…thanks again for staying up with me . . . get some rest!!!
Gus: no no, thank *you
```
11. Replaced letters worked in some instances to subtract the total number of keystrokes (as in 2morrow) but in others only maintained the same number of keystrokes (as in assumptions are a b**ch). We have included it as a subtracting feature of IM language because there were no instances of replaced letters adding to the production of the language.

REFERENCES


LEWIS, C., ENCISO, P., & MOJE, E. B. (Eds.) (2007). Reframing sociocultural research on


