Language Learning & Technology
http://llt.msu.edu/vol14num3/parkkinginger.pdf
October 2010, Volume 14, Number 3
pp. 31–50

WRITING/THINKING IN REAL TIME: DIGITAL VIDEO AND CORPUS QUERY ANALYSIS

Kwanghyun Park and Celeste Kinginger
Pennsylvania State University

The advance of digital video technology in the past two decades facilitates empirical investigation of learning in real time. The focus of this paper is the combined use of real-time digital video and a networked linguistic corpus for exploring the ways in which these technologies enhance our capability to investigate the cognitive process of learning. A perennial challenge to research using digital video (e.g., screen recordings) has been the method for interfacing the captured behavior with the learners' cognition. An exploratory proposal in this paper is that with an additional layer of data (i.e., corpus search queries), analyses of real-time data can be extended to provide an explicit representation of learner's cognitive processes. This paper describes the method and applies it to an area of SLA, specifically writing, and presents an in-depth, moment-by-moment analysis of an L2 writer's composing process. The findings show that the writer's composing process is fundamentally developmental, and that it is facilitated in her dialogue-like interaction with an artifact (i.e., the corpus). The analysis illustrates the effectiveness of the method for capturing learners' cognition, suggesting that L2 learning can be more fully explicated by interpreting real-time data in concert with investigation of corpus search queries.

INTRODUCTION

Technological innovations and increasingly ubiquitous access to networked computers via the Internet (Chun, 1994; Kern, 1995) are rapidly changing the scope of research in second language acquisition (SLA). The advance of digital video technology in the past two decades facilitates the collection of data on, and empirical investigation of, real time learning processes. Furthermore, technological revolutions in computer networking allow for innovative research methodologies that have been difficult, if not impossible, to conceptualize in the past. The focus of this paper is the combined use of digital video and a networked linguistic corpus for exploring the way in which these technologies enhance our capability to investigate the process of learning in SLA. The paper describes this exploratory method and applies it to the study of second language writing. This area is chosen for the known difficulty of tracking the learner's cognition during the composing process (Miller, 2005). This study will present an analysis of a second language (L2) writer's composing process, illustrating the effectiveness of the method for fine-grained portrayal of learners' cognition.

In the past fifty years or so, much research on the composing process has endeavored to understand the workings of the human mind as it constructs a text. A perennial challenge to research in this area has been to collect data that interfaces with the cognitive processes of writers (Humes, 1983; Stallard, 1974). Typically, this research has relied upon two major data sources: (a) retrospective accounts from participants (e.g., stimulated recall interviews) and (b) audio/video data collected in real time (e.g., keystroke logging and screen recording). Although these studies have significantly advanced the description of real-time, online writing processes, they have encountered difficulties in connecting the interpretation of data to internal cognitive processes.

Retrospective data is not so much a precise reproduction of the composing process as a reinterpretation of it, through which writers create coherence out of their past writing experiences. Reliance on the participant's account taken at face value, thus, can and does undermine the validity of research (Pavlenko, 2007). In order to provide more precise documentation of the composing process, process research has increasingly used computer-aided techniques for recording user activity, primarily keystroke analysis

(Miller, 2005; New, 1999; Sullivan & Lindgren, 2006) and screen recording studies (Geisler & Slattery, 2007).

Real-time methodology, however, has its own shortcomings: Lacking a means to connect the real-time video and keystroke transcripts to internal cognitive processes, analysis of these data relies on researchers' interpretations, a position exactly opposite to, and no better than, the retrospective method relying on the participant's perceptions. Thus, the contribution of the real-time data has been limited to descriptions of user behavior (e.g., the number and length of pauses) and has not been extended to enhance analytical validity in composing process research.

Due to this methodological gap, research on the composing process has been reduced to basing data interpretation either on the researcher's speculation (Miller, 2005, p. 311) or on the participant's perceptions. Although some studies have suggested supplementing real-time data with retrospective accounts (e.g., Geisler & Slattery, 2007; Lindgren & Sullivan, 2006), such suggestions only bring us back to the debate over whether retrospective narratives constitute valid research data. A new answer may reside in constructing an additional layer of data that reflects, and thus, allows the study of, the cognitive processes underlying the construction of text.

In this study, the analysis of real-time data is complemented with investigation of writers' search queries used to consult a corpus. The term "query" broadly refers to words and phrases that users enter into a database in order to retrieve relevant results. In this study, it refers to a particular kind of search that L2 writers use to consult a specialized academic corpus as a linguistic reference. The study's central hypothesis is that each query expresses an immediate need of the writer, and thus furnishes information about his or her thinking processes.

An exploratory proposal in this paper is that, using the corpus queries as research data, we can begin to reconstruct internal cognitive processes without undermining analytical validity. This paper describes a methodological framework that triangulates three types of data (i.e., screen recordings, retrospective reflections in oral and written forms, and corpus search queries). Based on these data, the study documents the L2 composing process through a moment-by-moment microanalysis of text construction. Consistent with the post-cognitive perspective in writing research (Atkinson, 2003) and the growing interest in artifact-mediated language learning and development (Chapelle, 2000; Chun, 1994; Kern, 1995; Kinginger & Belz, 2005; Kramsch & Thorne, 2002; Salaberry, 1999; Thorne, 2003; Warschauer, 1997, 1998; Warschauer & Kern, 2000), the study takes into account the role of a particular artifact (i.e., an online corpus) as an integral part of the focal rhetorical situation.

After briefly reviewing studies on the composing process and current issues in real-time methodology, the paper develops a corpus-based methodology responding to these issues. This methodology is then applied to an analysis of the composing process and the findings are discussed in regard to the role of the methodology in enhancing our understanding of the composing process. The paper concludes with a discussion regarding the implications of the methodology and potential directions for corpus-enhanced research as it relates to the investigation of learners' cognitive processes.

A Brief Overview of Composing Process Research

Two theoretical orientations have made major contributions to composing process research: the psychologically focused, cognitive-processing model (Flower & Hayes, 1981; Hayes & Flower, 1980) and the more socially oriented, post-cognitive approach (Kent, 1999; Trimbur, 1994). In the cognitive-processing model, the human mind is typically conceptualized as a processor whose internal mechanisms are the focus of research (Bereiter & Scardamalia, 1987; Hayes, 1996). Drawing on information-processing theory (Newell & Simon, 1972), the goal of this approach is to create a cognitive model for describing internal brain operations during the production of a text. Conceptually, the model represents machinery that takes input (information and knowledge); repeats planning, translating, and reviewing; and

generates a text as output. Previously influential in L1 and L2 writing research (Krapels, 1990; Raimes, 1983, Zamel, 1982, 1983), the model has lost much of its allure because it sidesteps consideration of sociocultural contexts and writers' development in academic settings (Silva, 1990); yet, its epistemological stance remains pervasive in writing research.

The post-cognitive perspective emphasizes sociocultural aspects of the acquisition of knowledge, positing that cognition takes place in a "network." Unlike the processor metaphor, wherein the mind is an inviolate entity operating in isolation, the network metaphor suggests that the human mind is inseparably embedded within a social-material situation (Kinginger, 2004; Lantolf, 2006; Lantolf & Pavlenko, 2001). Accordingly, research on composing from this perspective has expanded to include cultural, historical, and socio-political contexts in understanding the construction of texts (Atkinson, 2003; Matsuda, 2003).

With the shift toward the post-cognitive era, we are in a better position than that afforded by the processing model to address the changing ecology of academic writing. In this period of digital communication environments, the contemporary writing context now offers an extensive array of networked artifacts, including Web search engines, academic databases, and linguistic corpora. In such an environment, the post-cognitive perspective enhances our understanding of the composing process by bringing to the fore the nature of writing as a multi-resourced and multi-party activity.

Real-Time Data Collection Methodologies in Composing Process Research

Methods for real-time data collection and analysis first became visible in the literature in the mid-1990s. Two state-of-the-art real-time techniques, keystroke logging and screen recording, are reviewed in Table 1.

Table 1. *Real-time Computer Techniques in Composing-Process Research*

Technique and Area	Study
Keystroke logging	
Behavioral features (pause, deletion, and repetition)	Eklundh (1994); Lansman, Smith, & Weber, (1993); Levy & Ransdell (1994); Miller (2000, 2006); Miller, Lindgren, & Sullivan (2008); Ransdell (1995); Thorson (2000);
Quantity and types of revision	Bonk & Reynolds (1992); Flinn (1987); New (1999); Ransdell & Levy (1994); Reynolds & Bonk (1996)
Task type and text structure	Eklundh & Kollberg (2003)
Topic development	Miller (2005)
Screen recordings	
Screen videotaping	Collier (1983)
L1 revision in word-processing	Owston, Murphy, & Wideman (1992)
Personal writing	Geisler (2001)
EFL collaborative writing	Glendinning & Howard (2001, 2003)
Learner perception	Luoma & Tarnanen (2003)
Technical writing	Slattery (2005, 2007)
Communication in the workplace	Swarts (2004, 2010); Van Ittersum (2009)

Keystroke logging refers to computer software and techniques to record keyboard presses and mouse movements (Lindgren & Sullivan, 2006). Studies of this type have mainly focused on describing formal

behavioral features including pauses, deletions, and repetitions of key presses. In L2 composing process studies, the use of keystroke logging in writing pedagogy has been discussed with regard to charting out the revision types in French (New, 1999; Scott & New, 1994) and the comparison of L1 and L2 writing in German (Thorson, 2000) and Swedish (Lindgren & Sullivan, 2006). When compared with keystroke logging, screen recording offers a better visualization than a keystroke replay. These programs record the computer screen actions and create digital video clips (for review, see Latif, 2008).

Issues in Real-Time Method in Research on Composing Processes

Despite the growing popularity of these techniques, the value of real-time methods is not well defined or widely recognized in pedagogical research. Researchers in this area have not articulated what the quantity and type of keystrokes and behavior captured in screen recordings reveal about internal cognitive processes of writers in the construction of texts. Geisler and Slattery (2007) suggested that the key to this problem might be found in taking account of "the writers' consciousness" (p. 199); therefore, "suitable prodding" in recall interviews would reveal "*why* a writer has employed a given sequence of actions" (p. 197). However, this suggestion immediately reminds us of the issue of internal validity raised in narrative research: Participants create coherence out of past events through story-telling (i.e., narratives) and their accounts are not exact reproductions of the events (Ochs & Capps, 2001; Pavlenko, 2007). Regarding the difficulty of connecting real-time data to cognitive processes, Miller (2005) observes:

> From the outset, it has been clear that keystroke logging yields only partial and indirect data about a highly complex human activity. . . . Although triangulation of data from concurrent or retrospective interviews can support the teacher/researcher's process of interpretation, the observations remain speculative. We can at best infer what underlies the overt behaviour we are able to capture. (p. 311)

Although these points concerning the speculative nature of data interpretation are well taken, the limitation does not necessarily imply that real-time technique cannot be used for studying a highly complex human activity such as (L2) writing; nor does it mean that researchers have no better means of inquiry than retrospective speculation. A high degree of precision (and less speculation) in data interpretation can be obtained by introducing appropriately innovative methodology. For example, Lindgren and Sullivan (2006) used keystroke data to perform a step-by-step analysis of the linguistic hypotheses that 12^{th}–15^{th}-grade Swedish writers made about L2 English. Instead of asking the participants to produce a holistic reflection on their writing, the researchers (p. 162) described in detail the instances of revision and inferred the cognitive process that the writers were experiencing. For example, one student wrote:

> I want you to come to sweden therefor Sweden <4.8> 6 ⌫ → the summer holidays
> In Sweden <3.0> are really great.

The keystroke replay shows that the student originally wrote, "I want you to come to sweden therefor Sweden," then paused (4.8 seconds) and deleted "Sweden," and added, "the summer holidays in Sweden are really great." In regard to the revision, Lindgren and Sullivan (2006) discussed two potential hypotheses, which can be summed up as follows:

> **Hypothesis 1**: The revision is *formal*. "therefore" is a mistranslation for a Swedish word with a similar morphology, "därfür," for English "because." The writer was working on the form (i.e., sentence structure) intending to create a subordinate sentence that would provide the reason for the invitation.
>
> **Hypothesis 2**: The revision is *conceptual*. In the midst of a global revision, the writer might have an (unknown) intention to "adjust" the sentence according to "the overall plan" (p. 162). Therefore, the significance of the deletion of "Sweden" is minimal.

Although this analysis marks a step forward in keystroke analysis by breaking down the cognitive activity into smaller units to trace learner hypotheses, the analysis still relies on the researchers' interpretation; therefore, it is more or less speculative. The two hypotheses are based upon inferences about the writer's cognitive process that the key presses may or may not represent.

How then do we remove this speculative reasoning from an analysis of real-time data? This paper proposes that one answer lies in the provision of a further layer of data—one capable of connecting real-time data to cognitive processes. Although such a move cannot save us from speculation altogether, it can more securely anchor interpretation to results. Suppose we have a hypothetical piece of data showing that the writer in the above example eventually replaced "therefore" with "because." With this pair of words, it becomes much clearer that the writer was focusing on these particular items, rather than on a global-scale adjustment, and thus, the pair allows us to accept Hypothesis 1 and reject Hypothesis 2 with greater confidence. Such pairs, in other words, serve as an additional layer of data revealing writers' thinking processes during composing. This paper explores a methodology for augmenting analysis of real-time data with additional data in the form of corpus queries.

THE STUDY

Objective

This study is a part of a larger project that longitudinally investigated the composing processes of L2 writers in the context of corpus-assisted writing instruction at the tertiary level (Park, 2010). The present study seeks to achieve a high level of precision in real-time data interpretation and composing process documentation through a methodological innovation. Specifically, this study evaluates one special kind of data—corpus search queries—for its potential to enhance the analysis of real-time data. The study focuses on two research questions:

1. How can a methodology be developed to collect and analyze corpus search queries?
2. What is the role of corpus search queries in connecting the interpretation of real-time data and retrospective data to the needs and the internal cognitive processes of an L2 writer?

Context

The research site was an intermediate/advanced ESL writing course at a large American university. The class met twice a week in a classroom. Additionally, students attended three writing conferences in the teacher's office during the semester. This paper tracks one female student, Yilin (a pseudonym), and documents her composing process. Yilin was a first-year student in business with an L1 background in Mandarin Chinese. Yilin had one year's experience of taking general English courses in an international school in China and had a high level of proficiency both in written English as attested in the scores she received for her written assignments.

As a resource for the course and a research tool as well, the first author created a corpus-based system with a database of academic texts and a custom search engine (Google CSE) as a companion searching device to the database. An electronic collection of academic texts within a relevant topic area, the corpus had about 350,000 words and consisted of online academic journals in the broad topic areas of communication and technology, language use, and language learning.

While the custom search engine is similar to a concordancer, a popular corpus search tool, within pedagogical applications the search engine and the concordancer function in fundamentally different ways (for concordancing in pedagogy, see Davies, 2008). Concordancers produce lines in a format called keyword-in-context (KWIC). Advocates of pedagogical concordancing argued that the KWIC representation helps learners to notice linguistic patterns (e.g., collocations) through focus-on-form activities, and thus facilitates language learning (Cobb, 1997; Flowerdew, 1996; Gaskell & Cobb, 2004;

Johns, 1986, 1991; Tribble, 1990). In addition to the KWIC lines, most concordancers provide analytical information such as word frequency, collocation, keywords, and n-grams, which learners may find useful. Custom search engines differ from concordance software in that they do not offer analytical information or present search results in KWIC format. The search engine, however, comes with other benefits: First, it is a prefabricated service that allows users to create a corpus accessible via the Internet, which is a capability that most concordancers lack (but see Lu, 2009). Instead of collecting texts, users can simply provide hyperlinks of the online texts to the engine and create a searchable database. Secondly, the search engine provides a relatively richer context than does the KWIC format, showing sentences that co-occur with the search terms and highlighting all search terms instead of one single keyword. Third, multiple-word search is much easier with a search engine than with concordancers. This functionality is important for the writing course in this study, as the corpus system was introduced to allow students to access the corpus for writing references rather than for form-focused activities. A useful feature of the engine was its flexibility, which allowed the course designers to easily compile and modify the contents of the corpus by simply creating hyperlinks to the online texts. In addition to search capability, it is important that the system have a logging device to record the searches that users enter into the system. A computer program was written for this purpose and attached to the system. The program received and saved the queries in a separate computer database, from which a log of search queries could be retrieved. The log contained three kinds of key information among others: timestamps, search queries, and sign-in pseudonyms. Figure 1 shows an example of the retrieved query log.

No.	User	Queries	IP Address	Date
1307	kup133	but it will also	99.71.00.000	2008-11-16 21:56
1726	kup133	benifit from	99.71.00.000	2008-11-22 21:59
1728	kup133	benifit	99.71.00.000	2008-11-22 21:59
1729	kup133	benifit them to	99.71.00.000	2008-11-22 22:01
1731	kup133	benifit to	99.71.00.000	2008-11-22 22:01
1735	kup133	beneficial	99.71.00.000	2008-11-22 22:09
1736	kup133	beneficial to	99.71.00.000	2008-11-22 22:09
2028	kup133	but it will also be	99.71.00.000	2008-12-04 16:04

Figure 1. Screenshot of a sample query log. (*Note*. IP addresses are not the actual addresses due to privacy concerns.)

Data and Collection Procedures

This study analyzes three types of data: (1) the query log, (2) screen recordings, and (3) oral and written reflections. Screen recordings and corpus queries were collected during writing tasks in the classroom. The classroom was equipped with networked computers offering students access to the corpus system. A query is a search that writers enter in order to consult the course corpus. Thus, the query log in this study simply refers to a list of searches that writers have entered while consulting the corpus. The log contains three kinds of information: timestamps, search queries, and sign-in pseudonyms. The computers had screen-recording software (iShowU) installed as well, which the students used to record their screens. The data were collected while the students were engaged in their in-class writing tasks. As for the oral reflection data, writing conferences took place in the teacher's office and each session lasted approximately twenty to thirty minutes. In the sessions, students watched their screen-recording clips and commented on their corpus consultations. The written reflections were collected from student papers and electronic bulletin board postings.

Among the three kinds of data, the query log is of special importance to this study, as it constitutes the "additional layer of data" to address the validity issue in real-time writing research. In contrast to research

emphasizing the writer's reflections (e.g., Geisler & Slattery, 2007), this paper suggests that we may improve analytical validity through a kind of data that mediates between real-time data and cognition. Specifically, this paper hypothesizes that search queries closely interface with writers' thinking processes, as queries are, by definition, created to address (and thus reflect) their cognitive needs. In that sense, no query is random and all queries have a specific purpose. Analysis of these search queries may lend stronger support to our inferences regarding the cognitive processes of writers given that we will base our data interpretation on logical reasoning, rather than on our intuition and/or the writers' perceptions.

Data Analysis

Data analysis was performed using a two-step procedure. In order to develop basic research tools (i.e., to define *what* to analyze and *how*) a pilot analysis was conducted. The objective of the pilot analysis was two-fold: to identify the unit of analysis and to establish a data-coding scheme. Three primary data—screen video clips, oral/written reflections, and the query log—were reviewed multiple times to identify consistent patterns across these data. After the pilot analysis had identified the analytical unit and the coding scheme, a main analysis was performed on the focal student's data to document her composing process.

The pilot analysis of the data suggested that the focal student's composing processes were organized in chunks forming discernable patterns. These patterns seemed to reflect her writing needs followed by an effort to address those needs. The present paper refers to such a sequence as "transaction." Conceptually, a transaction in a composing process is identified based on three characteristics: (a) it responds to a particular problem or challenge, (b) it looks for a single target discourse item, and (c) it normally begins and ends with a visual signal in the screen recordings (e.g., beginning with mouse cursor highlighting and ending with scrolling away). The next step was to formally define a transaction. The following example obtained from a student's screen recording illustrates a formal transaction boundary:

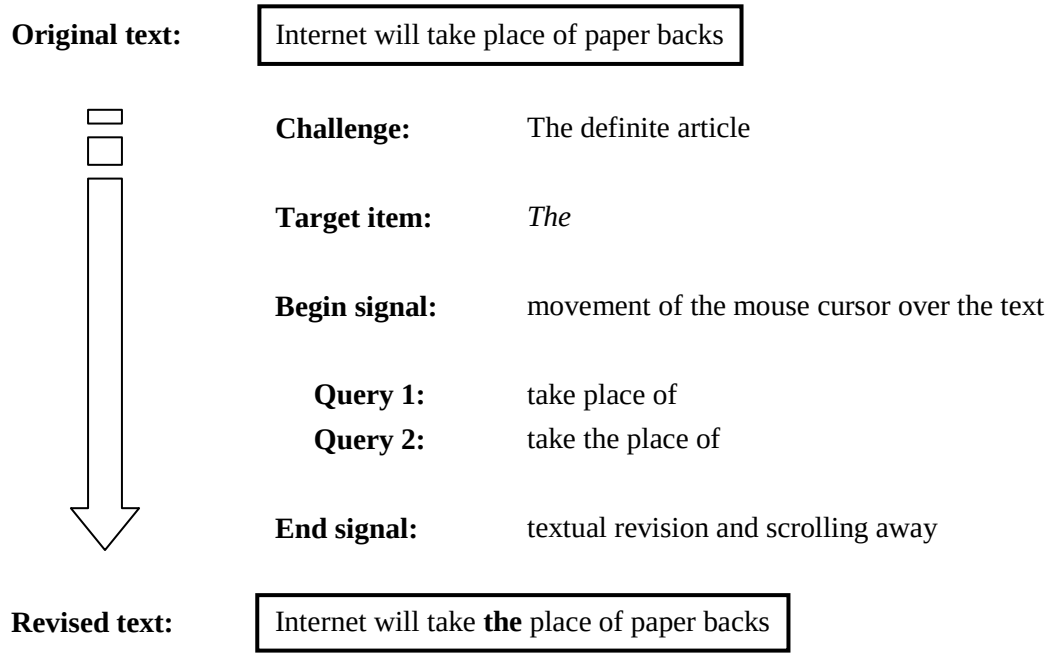


Figure 2. Illustration of a formal transaction boundary sequence.

Figure 2 shows that a transaction is a formally identifiable phenomenon across three types of data (i.e., corpus query, screen recordings, and reflections). In the query log, the student's searches (Query 1 and 2) have formal similarities and are considered a potential transaction. These queries can be connected to the

corresponding visual signals in the screen recordings. Typically, the writer begins a reflection by commenting on the begin signal (moving mouse cursor) and concludes the narrative with a comment that corresponds with the end signal (revision and scrolling away) in the screen clip. Furthermore, description in stimulated recalls shows that the boundary of the transaction corresponds to the reflection data as well: The writer describes the writing process in a series of narrative episodes, each of which has an identifiable beginning and end. This sequence shows that the writer perceives a challenge and an ensuing effort as an identifiable event, which gives further support to the proposal that the transaction is a valid unit of analysis. Traditional units such as words, sentences, and paragraphs do not capture the writer's conceptualization. In contrast, transactions allow us to reorganize composing processes within meaningful boundaries of writers' needs and responsive effort in a way that is consistent with the writer's perceptions.

Analyzing Transactions

This paper offers a formal and functional analysis of the composing process. In formal analysis, the objective is to identify transactions and describe the needs based only on the form. In the above example, two searches, "take place of" and "take **the** place of," constitute a transaction, which allows us to locate the learner's need and make a guess about it. By comparing the two related queries, we may conclude that the transaction revolves around a syntactic issue (i.e., the article system and a lexical issue, namely a formulaic expression). At this point, of course, the conclusion is based only on the form reflected in the query log.

Functional analysis connects formally identified transactions to a writer's hypotheses and strategies by considering the content of the data. If formal analysis seeks to answer what the challenge was, functional analysis focuses on how the challenge was met. In addition to the formal analysis of the query log, the screen video clips further show that the writer was comparing two example sentences in the search results (Figure 3):

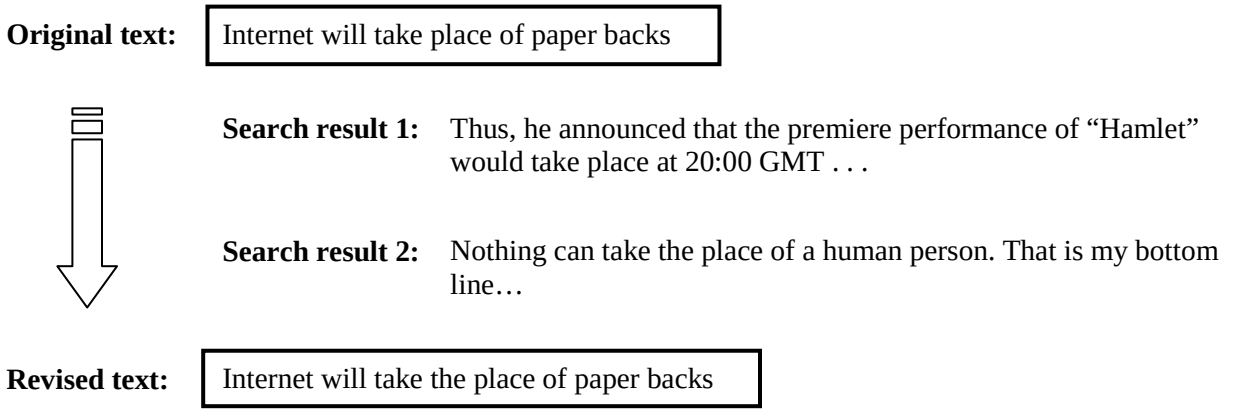


Figure 3. Sample sequence of queries, search results, and revision.

These two example sentences from the corpus lead to the subsequent revision in which "the" is added to the writer's sentence, resulting in "Internet will take **the** place of paper backs." From the first search result, the writer noticed the discrepancy between the intended meaning in the original text ("replace") and the meaning in the example sentence ("occur"). Then the writer formed a hypothesis about a correct form, tested it out in the next query, and accepted the hypothesis by confirming it in the second search result.

The next step in the analysis is to connect our interpretation of the transaction to the writer's perception of the process as it is given in the reflective narrative. This step looks for consistency of content between the data sources:

> **I was not sure** about whether "take place of" or "take the place of" is correct. **At the beginning**, it [the corpus] didn't solve my problem, because I didn't pay attention to the actual results. Only by simply comparing the number of results from "take place" and "take the place." I chose to use take place here. [However], **I noticed that** although "take place" provides more results, none of them are in the form of "take place of." **Then** I tried "take the place" to see it's correct.
>
> (*Note*. Excerpt from the writer's self-reflection)

Despite its brevity, the writer's reflection has all the essential components and maintains the canonical sequence of a narrative (Labov & Waletsky, 1967). The writer began the reflection with an orientation, "I was not sure," and presented the content of the story. Then there came the complication, when the writer described how the initial assumption was challenged by a later discovery ("I chose to use take place here. [However], I noticed that…"). Finally, resolution was reached when the writer tried a new query ("Then I tried…"). Thus, reflective accounts are organized as narrative "episodes," wherein problems are noticed (analyzing the search results, noticing the discrepancy), then solved (setting up and testing a hypothesis).

To conclude, the pilot analysis shows that the transaction is a valid unit for analysis. Although the example in the sample analysis is a very simple one, it still shows that even this simple transaction involves a complicated procedure of identifying the challenge, evaluating the resources, and testing a hypothesis.

Data Coding

In coding the query data, two formal categories of transaction, simple and complex, emerged and were used to distinguish between transactions with only one query (simple) and those with multiple queries (complex). Figure 4 exemplifies these two kinds of transactions.

Transaction ID	tran-type	Query ID	query words	timestamp
yvw-08-0047	simple	1307	but it will also	2008-11-16 20:01
yvw-08-0048	simple	1308	good writing skil	2008-11-16 21:56
yvw-08-0049	simple	1310	job field	2008-11-16 22:06
yvw-08-0050	simple	1311	difference with	2008-11-16 22:10
yvw-08-0051	simple	1312	to have almost	2008-11-16 22:30
yvw-08-0052	simple	1314	try to anyway	2008-11-16 22:32
yvw-08-0053	simple	1332	fluent essay	2008-11-16 22:57
yvw-08-0054	simple	1333	sentence	2008-11-16 23:53
yvw-08-0055	simple	1334	plarism	2008-11-16 23:54
yvw-08-0056	simple	1344	confused	2008-11-17 0:26
yvw-08-0057	complex	1345	task	2008-11-17 18:20
yvw-08-0057	complex	1509	toughtask	2008-11-17 18:29
yvw-08-0057	complex	1510	tough task	2008-11-20 15:03
yvw-08-0057	complex	1512	tough task	2008-11-20 15:03
yvw-08-0057	complex	1513	tough task	2008-11-20 15:03
yvw-08-0057	complex	1515	task	2008-11-20 15:03
yvw-08-0057	complex	1518	task	2008-11-20 15:03
yvw-08-0057	complex	1520	hard task	2008-11-20 15:04
yvw-08-0057	complex	1521	hard task	2008-11-20 15:04
yvw-08-0057	complex	1524	difficult task"	2008-11-20 15:04
yvw-08-0057	complex	1526	difficult task	2008-11-20 15:05

Figure 4. Simple and complex transactions.

There were two reasons to code transactions as either simple or complex. First, there were noticeable differences between the distributions of simple and complex transactions in the query log: Complex

transactions tend to cluster together forming a visible group in the query log, while simple queries are scattered between the complex query clusters. This suggests that the analysis might benefit from isolating and examining them separately. Second, complex transactions seem to reflect the writer's needs and her strategy to address the needs. In other words, complex transactions make visible the writer's cognitive processes in a moment-by-moment manner. Thus, complex transactions offer richer insight into the writing processes than do simple transactions.

Once these kinds of transactions had been identified, it was clear that interpreting simple transactions is much more difficult than interpreting complex transactions. For example, in Figure 4, a pair of queries "hard task" (#1521) and "difficult task" (#1524) form a complex transaction, as they differ minimally in regard to the adjectives, "hard" and "difficult", indicating that the writer's needs revolved around a lexical choice (i.e., an adjective for "task"). With a simple transaction, however, (e.g., "confused," #1344), such reasoning would be extremely difficult. Therefore, the analytic focus of this study is the complex transaction.

RESULTS

For the present analysis, we collected a total of 109 minutes of screen recordings for the focal student. We also retrieved 194 corpus searches from the focal writer's query log and identified 118 transactions. Among them, simple transactions (N = 92) outnumbered complex transactions (N = 26) by approximately three times.

Documenting Transactions and Multi-Dimensional Needs

Transactions originate from writers' needs and concluded in a follow-up effort for addressing these needs. An analysis of the query log suggests that needs arise from syntactic, lexical, and morphological issues:

Table 2. *Breakdown of Learner Needs in Transactions*

Motivation	Query ID	Query
Syntactic	751	no excuse from
	752	no excuse for
Lexical	1582	change essay
	1625	revise essay
Morphological	1743	miss used word usage
	1744	misused word usage

Table 2 exemplifies the rather simple, two-query transactions. By comparing two queries that minimally differ from each other, we can zero in on the particular needs of the writer. We identify a syntactic need (i.e., a grammar issue based on comparison of "for" and "from"), a lexical need (i.e., a word choice, by comparing "change" "and "revise"), and a morphological concern in the third pair (i.e., "miss used" and "misused")

These needs are further complicated, if multiple linguistic foci come into play in a single transaction. Then, a mix of needs may be identified in a single transaction. Table 3 illustrates three mixed needs (lexico-morphological, lexico-syntactic, and syntactic-morphological) in the query log:

Table 3. *Mixed Needs in Transactions*

Mixed motivation	Query ID	Query example	Motivation
Lexico-morphological	647	relat**ive** to	lexical/morphological
	648	relat**ed** to	lexical/morphological
Lexico-syntactic	1585	**on** a website	syntactic
	1587	**in a website**	syntactic
	1589	in **the web**	syntactic/lexical
Syntactic-morphological	2461	**aquire** information	syntactic
	2462	aquire **of** information	syntactic
	2463	aquire of information	
	2464	**acquire** of information	morphological
	2465	**acquiring** of information	syntactic
	2466	**acquirance** of information	morphological
	2467	**acquisition of** information	morphological

The first transaction consists of one minimal query pair that contrasts two words ("relative" and "related"), which are etymologically connected, but differ in meaning. Apparently, the morphological similarity between the two lexically separate items was a source of confusion for the writer. The query log has other similar examples including "respectably" and "respectively," and "beneficial" and "beneficiary." The second transaction shows a syntax-focused contrast between two prepositions ("on"/ "in") in the first minimal query pair, while it extends to syntactic ("a"/"the") and lexical ("website"/"web") concerns in the second pair. The final example has multiple queries that are not only plural in number, but recurrent throughout the transaction. We see a series of revisions with a syntactic focus ("acquire of information" and "acquiring of information") as well as a morphological focus ("aquire" and "acquirance"). These needs alternate in a recursive manner between concerns about syntax and morphology following the path of the writer's problem-solving.

For the purpose of demonstrating a method to analyze transactions, one transaction from the focal student's query was retrieved and coded:

Table 4. *Coding a Transaction*

Query ID	Query example	Needs
2435	**obtain** skill	Lexical
2437	**have** skill	Lexical
2440	**writing** skill	Lexical
2441	**develop** writing skill	Lexical
2442	develop writing **ability**	Lexical
2444	develop writing **abilities**	Syntactic
2446	develop writing **skill**	Lexical

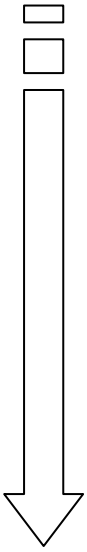
Table 4 shows that the focal writer's transaction was motivated by lexico-syntactic needs. The coded transaction serves as ground data for the next step in the transaction analysis.

Tracking Hypotheses

In the next step of analysis, the query log is compared with screen recordings and reflections to confirm (or revise) the transaction analysis so far. Based on the comparison, the composing process will be described. The description, however, does not directly rely on either kind of data—screen recordings or reflections. Instead, the present analysis will reconstruct the focal writer's hypotheses by triangulating from multiple data sources. The reconstructed description of hypotheses, then, will serve as an intermediary dataset to support inferences about the composing process.

The first step to describe the hypothesis development is to list the queries in a complex transaction in parallel with the writer's reflections. A complex transaction in the query log suggests that the writer is evaluating the corpus search results; a subsequent refinement of a query indicates that the writer is formulating a hypothesis based on the evaluation and is testing it against the corpus. Therefore, by examining the query refinement, we can possibly track the writer's hypotheses as they have developed. Figure 5 exemplifies this step by tabulating the focal writer's queries and her reflections:

Original text: **Obtaining** a decent writing skill will not only help college students



Query	**Reflection**
obtain skill	it sounds like anyone could go out and buy their skillfulness in writing. So I searched "obtain skill"
have skill	(no comment)
writing skill	**most of the sentence** use "obtain + property", but skill is ability, and I **did not find any reference** contains "obtain skill." Then I typed in "writing skill"
develop writing skill	I saw **sometimes they use** "development of writing skill," then I search "develop writing skill."
develop writing ability	**I found people do use it**. But I also saw they use "develop writing ability" I was considering whether I should use "ability" instead of "skill," then I searched it, **but they do not use it a lot**. So I keep using develop writing skill

Revised text: **Developing** a decent writing skill will not only help college students

Figure 5. Tabulating the learner's queries and reflection.

The side-by-side tabulation of the queries and the accompanying comments allows us to track the development of Yilin's thinking. Just by examining the queries (left column), we can draw a tentative conclusion that these queries focus on lexical issues (i.e., word choice) rather than on syntactic or morphological issues. Then, by taking account of the writer's reflection (right column), we can confirm our conclusion with greater confidence.

The next step in tracking the development of the learner's hypotheses is to actually reconstruct them within the available datasets. Based on triangulation of these data, the present analysis provides a moment-by-moment representation of the development of her hypotheses, again, in a tabulated presentation:

Original text: **Obtaining** a decent writing skill will not only help college students

Query	Learner's hypothesis	Evaluation
obtain skill	**H1**: "skill" may not be something that can be obtained in a short time	H1: Accepted
	H2: "obtain" does not highlight the fact that the object ("skill") takes time and effort to achieve	H2: Accepted
have skill		
writing skill	**H3**: "writing skill" may elicit the collocating verbs	H3: Partially accepted
develop writing skill	**H4**: "develop skill" is an appropriate collocation	H4: Partially accepted
develop writing ability	**H5**: "develop ability" is more appropriate than "develop skill"	H5: Rejected

Revised text: **Developing** a decent writing skill will not only help college students

Figure 6. Visualizing the learner's hypothesis development.

The tabulated representation (Figure 6) shows that the writer evaluated corpus results and tested her hypotheses based on the evaluation of the results. On the surface, the only textual change is the substitution of "developing" for "obtaining." However, the analysis shows that the revision is a result of a complex process involving the writers' linguistic analysis and hypothesis.

Yilin's first and second queries originated from her sensitivity to language use and an effort to articulate her thoughts about learning to write as an L2 writer. In her first query, "obtain skill," Yilin hypothesized that there was a better verb than "obtain" for emphasizing the work it takes for an international student to become a good writer in English. For her, achieving a high level of skill as an L2 writer was something that would require a longitudinal effort; it was not a skill that one could simply come by in a short period of time. H1, her first hypothesis, accordingly states that the semantics of the noun, "skill," entails a longitudinal effort. She felt, somehow, that the verb "obtain" did not connote this painstaking effort. Yilin commented that "obtain" sounded as if the writing expertise were something that "anyone could go out and buy," such that its use would downplay the challenges involved in the longitudinal learning process. H2, the second, subsequent hypothesis, states that the verb, "obtain," does not correspond with the semantics of "skill."

The corresponding query, "obtain skill," was her strategy to confirm her hypothesis about the non-existence (and therefore inappropriateness) of this verb–noun collocation. The screen capture data shows that the collocations for "obtain" retrieved from the corpus in response to her query included "information," "outcome," "meaningful data," "goals," and "many skills." Perusing these results, Yilin concluded that the verb "obtain" did not encode the aspect of temporal longitude, as it depicted rather a finite status of achievement. Thus, she accepted two hypotheses and concluded that "obtain" did not collocate well with "skill." Then, in her second query, "have skill," seems to reflect her effort to

determine the semantic property of "skill" by contrasting "obtain" with "have," which represents an aspect of a finite state.

Yilin's next step was, naturally, to look for a verb that would express the longitudinal acquisition of writing "skill." Her third query ("writing skill") is an effort to find such a verb, from which the third hypothesis can be inferred. However, the following hypothesis, H3, is quite different in its purpose from the two preceding hypotheses. In the two earlier hypotheses (H1 and H2), Yilin already had in mind the target form ("obtain skill") and intended to confirm the (non) use of this particular form. In her third hypothesis, however, she did not have the target form and, therefore, had to elicit one from the corpus through her query. Although she felt that "obtain" did not work for "skill," she did not know what verb would substitute for "obtain." In response to the need, H3 states that the query, "writing skill" will elicit some verbs to replace "obtain." As a result of the search, Yilin did find a candidate verb, but in a nominal form ("development"). Now her query only partially served the purpose (of testing the hypothesis). Therefore, we may argue that she partially confirmed H3 by finding a semantic replacement and yet partially rejected it as well by retrieving a nominal form.

Then her logical next step was to confirm whether or not the collocation "develop skill" is an appropriate one, as reflected in H4. Although Yilin eventually accepted the hypothesis, her decision process was a convoluted one because she noticed that the verb "develop" co-occurs not only with "skill" but also with "ability." Now she had to choose one noun over the other. Yilin's final query, "develop writing ability," was intended to help make a decision between two collocations in her final hypothesis (H5). As the corpus results suggested that "skill" was used more often with "develop" than was "ability," Yilin rejected H5 and retained her original choice, "develop skill."

In sum, the analysis shows that the focal student engaged in a complex process of hypothesis development, shuttling between finding lexico-grammatical resources from the corpus and testing her hypotheses. Yilin's composing process can be understood as a multi-stage procedure consisting of three recursive steps: hypothesis testing through a query, analysis and evaluation of the search results, and revision (or optional query refinement).

DISCUSSION

Composing as a Developmental Process

The present analysis shows that Yilin's composing process was organized around dialogue-like interactions between her and an artifact. In her transaction, problem-solving began with a corpus query for hypothesis testing, followed by an evaluation of the search results, and finally an optional query refinement for further hypothesis testing. In its essence, it was a process in which the focal writer asked a question and received candidate answers from the corpus.

The dialogic pattern in the writer's corpus consultation bears a striking similarity to the way in which an experienced tutor offers help to a novice writer (described in Aljaafreh & Lantolf, 1994). With regard to help in a tutoring session, Aljaafreh and Lantolf defined the qualities of graduation and contingency: Tutors offer graduated assistance by aligning with the students' current stage of development; they offer contingent help by maintaining a high level of sensitivity to the student's performance and withdrawing that level of assistance once the student has reached the next developmental level.

Corpus consultation emulates such tutoring in two important respects. First, the corpus system dynamically adjusts exemplar materials to the increasingly sophisticated level of the writer's skill as reflected in a series of queries (graduation). Based on this graduated help, our focal writer achieved micro-level development, moving on from one dimension of challenge to the next, solving one little problem after another. Once the writer had solved the problem, the corpus consultation immediately came to an end and she started transferring the enhanced awareness to performance (contingency).

This analysis suggests that the composing process is fundamentally a developmental process. This study, thus, challenges the cognitive-processing model of Flower and Hayes (1981) as well as all studies neglecting the role of development in the construction of text. The data in this study show that each step in text production is both a process and a result of development, and L2 writing, as a whole, is a site where the struggle to achieve that development takes place.

Composing as Problem-Solving

While it has been conventional in L2 writing research to describe the composing process as consisting of temporal stages (i.e., planning, writing, and revising) the data in this study suggests otherwise by showing that planning, writing, and revising occur at the same time and across multiple linguistic dimensions (e.g., syntactic-lexical needs). The focal writer's composing process simply cannot be described as a sequence of discrete stages. Rather, planning, writing, and revising are difficult to separate from each other and do not necessarily occur in a temporal sequence. Specifically, the findings in this study suggest that the writer's composing process is organized around problem-solving work. And, inhering in this work is a dense locus of cognitive activity focused on hypothesis testing aimed at addressing the immediate challenges of producing text.

Composing as a Distributed Activity

The complex hypothesis testing of the focal writer suggests that she constructed her text in cognitive collaboration with the corpus system, rather than through an isolated processing activity. In other words, the focal writer and the corpus system engaged in a cognitive division of labor: The search engine performed all repetitive and time-consuming tasks such as collecting register-specific, grammatically acceptable texts and retrieving relevant sentences in response to the learner's needs. As the system had done this basic groundwork on the lexicogrammatical level, the cognitive burden on the writer became lighter, thus allowing the writer to operate on the more creative rhetorical level, concentrating on testing her linguistic hypotheses and articulating her thoughts.

CONCLUSION

This study sought to answer a question that is of significance to technology-enhanced research on learners' cognitive processes: How can we base our interpretation of real-time data on empirical evidence so as to enhance our understanding of the learner's cognition during the composing process? The challenge is to develop a method to index real-time data to the learners' cognition. The study has shown that a combination of screen recordings and corpus search queries, as an additional layer of research data, can play a significant role in providing a basis for data interpretation—one that is founded on logical reasoning rather than on our intuition or the participant's perception. When applied to the area of writing, the method allowed for a moment-by-moment analysis of the query data and traced a complicated process of hypothesis testing. The analysis therefore showed that this exploratory method can enhance our understanding of the L2 composing processes beyond what current research methods have offered. It is unclear that real-time techniques alone, with or without the support of reflective narrative, are sufficient to provide data from which to make inferences about the writer's internal cognitive process.

One limitation of the suggested method is that it examines only the portion of composing processes that are reflected in corpus queries. No doubt there were episodes when Yilin did not appeal to the corpus query in her composing process. These other moments are not captured in the query log and thus cannot be examined. The challenge of adequately describing these episodes remains unresolved. A related issue is learners' level of proficiency. It can be hypothesized that a novice learner will not produce as many complex queries as an advanced learner such as Yilin does. Complex queries suggest a process of hypothesis testing, one that requires both analytical skill and relatively high proficiency. Then, analysis of the novice learners' queries may not be as revealing as scrutiny of advanced learners' queries. Another limitation is that the method still depends on the learner's reflections for interpreting the query log. While

the triangulation technique is effective in identifying consistency between the real-time data and the learner's reflections, the method suffers when the learner's reflections are not available. Although a technique for representing the learner's cognitive process without such reflections is beyond the scope of this study, this is clearly a key issue for future inquiry.

The combination of screen recording video and corpus query analysis is an unexplored area of potential benefit for methodological innovations in learning process research. An important next step in this area is to conduct a larger-scale study to examine the development of L2 learners through query analysis and the real-time data collected in pedagogical as well as professional situations. Future research also needs to extend real-time research techniques to issues including learner behavior analysis, assessment of learners' performance and development, and evaluation of technology/software for language learning and teaching. Such efforts may yield ecologically valid applications for technology-enhanced pedagogy in contemporary contexts of language learning and teaching—contexts that are increasingly mediated, distributed, and collaborative.

ACKNOWLEDGEMENTS

We are very grateful to the anonymous reviewers for valuable suggestions leading to significant improvements in the paper. We are also grateful to our colleagues in the College of Liberal Arts at the Pennsylvania State University, particularly Karen Johnson, Xiaofei Lu, and Xiaoye You.

ABOUT THE AUTHORS

Kwanghyun Park recently received his Ph.D. in Applied Linguistics at the Pennsylvania State University. His research interests include applied corpus linguistics and learner language analysis with a focus on developing methods to investigate learner development using a computational system and real-time video data.

E-mail: kpark90@gmail.com

Celeste Kinginger is Associate Professor of Applied Linguistics at the Pennsylvania State University where she is affiliated with the Center for Advanced Language Proficiency Education and Research (CALPER).

E-mail: cxk37@psu.edu

REFERENCES

Aljaafreh, A., & Lantolf, J. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *The Modern Language Journal, 78*, 465–487.

Atkinson, D. (2003). L2 writing in the post-process era: Introduction. *Journal of Second Language Writing, 12*, 3–15.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Erlbaum.

Bonk, C., & Reynolds, T. (1992). Early adolescent composing within a generative-evaluative computerized prompting framework. *Computers and Human Behavior, 8*(1), 39–62.

Chapelle, C. (2000). Is network-based learning CALL? In M. Warschauer & R. Kern (Eds.), *Network-*

based language teaching: Concepts and practice (pp. 204–228). New York: Cambridge University Press.

Chun, D. (1994). Using computer networking to facilitate the acquisition of interactive competence. *System, 22*, 17–31.

Cobb, T. (1997). Is there any measurable learning from hands-on concordancing? *System 25*(3), 301–315.

Collier, R. (1983). The word processor and revision strategies. *College Composition and Communication, 34*(2), 149–155.

Davies, G. (Ed.). (2008). *Information and communications technology for language teachers (ICT4LT)*. Slough, Thames Valley University. Retrieved from http://www.ict4lt.org

Eklundh, K. (1994). Linear and nonlinear strategies in computer-based writing. *Computers and Composition, 11*, 203–216.

Eklundh, K., & Kollberg, P. (2003). Emerging discourse structure: computer-assisted episode analysis as a window to global revision in university students' writing. *Journal of Pragmatics, 35*(6), 869–891.

Flinn, J. (1987). Case studies of revision aided by keystroke recording and replaying software. *Computers and Composition, 5*, 31–43.

Flower, L. S., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication, 32*, 365–387.

Flowerdew, J. (1996). Concordancing in language learning. In M. Pennington (Ed.), *The power of CALL* (pp. 97–113). Houston, TX: Athelstan.

Gaskell, D., & Cobb, T. (2004). Can learners use concordance feedback for writing errors? *System, 32*(3), 301–319.

Geisler, C. (2001). Textual objects: Accounting for the role of texts in the everyday life of complex organizations. *Written Communication, 18*(3), 296–325.

Geisler, C., & Slattery, S. (2007). Capturing the activity of digital writing: Using, analyzing, and supplementing video screen capture. In H. A. McKee & D. N. DeVoss (Eds.), *Digital writing research: Technologies, methodologies, and ethical issues* (pp. 185–200). Cresskill, NJ: Hampton Press.

Glendinning, E., & Howard, R. (2001). Examining the intangible process: Lotus ScreenCam as an aid to investigating student writing. *Edinburgh Working Papers in Applied Linguistics, 11*, 42–58.

Glendinning, E., & Howard, R. (2003). Lotus ScreenCam as an aid to investigating student writing. *Computer Assisted Language Learning, 16*, 31–46.

Hayes, J. R. (1996). A new model of cognition and affect in writing. In M. Levy & S. Ransdell (Eds.), *The science of writing* (pp. 1–30). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Hayes, J. R., & Flower, L. S. (1980). Identifying the organisation of writing processes. In L. W. Gregg & E. R Steinberg (Eds.), *Cognitive processes in writing* (pp. 3–30). Hillsdale, NJ: Lawrence Erlbaum Associates.

Humes, A. (1983). Research on the composing process. *Review of Educational Research, 53*(2), 201–216.

Johns, T. (1986). Micro-concord: A language learner's research tool. *System, 14*(2), 151–162.

Johns, T. (1991). Should you be persuaded: Two examples of data driven learning. *ELR Journal, 4*, 1–16.

Kent, T. (Ed.). (1999). *Post-process theory: Beyond the writing process paradigm*. Carbondale: Southern Illinois University Press.

Kern, R. G. (1995). Restructuring classroom interaction with networked computers: Effects on quantity

and quality of language production. *The Modern Language Journal, 79*, 457–476.

Kinginger, C. (2004). Bilingualism and emotion in the autobiographical works of Nancy Huston. *Journal of Multilingual and Multicultural Development, 25*, (2&3), 159–178.

Kinginger, C., & Belz, J. (2005). Sociocultural perspectives on pragmatic development in foreign language learning: Case studies from telecollaboration and study abroad. *Intercultural Pragmatics, 2*(4), 369–421.

Kramsch, C., & Thorne, S. (2002). Foreign language learning as global communicative practice. In D. Block & D. Cameron (Eds.), *Globalization and language teaching* (pp. 83–100). New York: Routledge.

Krapels, A. R. (1990). An overview of second language process research. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 37–56). Cambridge: Cambridge University Press.

Labov, W., & Waletzky, J. (1967/1997). Narrative analysis: Oral versions of personal experience. In J. Helms (Ed.), *Essays on the verbal and visual arts* (pp. 12–44). Seattle: University of Washington Press. (Reprinted in *Journal of Narrative and Life History, 7*, 3–38).

Lansman, M., Smith, J., & Weber, I. (1993). Using the writing environment to study writers' strategies. *Computers and Composition, 10*, 71–92.

Lantolf, J. P. (2006). Sociocultural theory and L2: State of the art. *Studies in Second Language Acquisition, 28*, 67–109.

Lantolf, J. P., & Pavlenko, A. (2001). (S)econd (L)anguage (A)ctivity theory: Understanding second language learners as people. In M. Breen (Ed.), *Learner contributions to language learning: New directions in research* (pp. 141–158). New York: Longman.

Latif, M. (2008). A state-of-the-art review of the real-time computer-aided study of the writing process. *International Journal of English Studies, 8*, 29–50.

Levy, C., & Ransdell, S. (1994). Computer-aided protocol analysis of writing processes. *Behavior Research Methods, Instruments, & Computers, 26*, 219–223.

Lindgren, E., & Sullivan K. (2006). Analyzing on-line revision. In G. Rijlaarsdam (Series Ed.) and K. P. H. Sullivan & E. Lindgren (Vol. Eds.), *Studies in writing, computer keystroke logging: Methods and applications* (pp. 157–188). Oxford: Elsevier.

Lu, X. (2009). *A corpus-based approach to tracking L2 development*. Paper presented at the 2009 Annual Convention of the American Council on the Teaching of Foreign Languages (ACTFL-09). San Diego, CA.

Luoma, S., & Tarnanen, M. (2003). Creating a self-rating instrument for second language writing: From idea to implementation. *Language Testing, 20*, 440–465.

Matsuda, P. K. (2003). Process and post-process: A discursive history. *Journal of Second Language Writing, 12*, 65–83.

Miller, K. (2000). Academic writers online: Investigating pausing in the production of text. *Language Teaching Research, 4*, 123–148.

Miller, K. (2005). Second language writing research and pedagogy: A role for computer logging? *Computers and Composition, 22*, 297–317.

Miller, K. (2006). Pausing, productivity and the processing of topic. In K. Sullivan & E. Lindgren (Eds.), *Computer key-stroke logging: Methods and applications* (pp. 131–156). Oxford: Elsevier.

Miller, K., Lindgren, E., & Sullivan, K. (2008). The psycholinguistic dimension in second language writing: Opportunities for research and pedagogy using computer keystroke logging. *TESOL Quarterly, 42*(3), 433–454.

New, E. (1999). Computer–aided writing in French as a foreign language: A qualitative and quantitative look at the process of revision. *The Modern Language Journal, 83*(1), 80–97.

Newell, A., & Simon, H. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Ochs, E., & Capps, L. (2001). *Living narrative*. Cambridge, MA: Harvard University Press.

Owston, R., Murphy, S., & Wideman, H. (1992). The effects of word processing on students' writing ability. *Research in the Teaching of English, 26*, 249–276.

Park, K. (2010). *Enhancing lexicogrammatical performance through corpus-based mediation in L2 academic writing instruction*. Unpublished doctoral dissertation, The Pennsylvania State University, USA.

Pavlenko, A. (2007). Autobiographic narratives as data in applied linguistics. *Applied Linguistics, 28*(2), 163–188.

Raimes, A. (1983). Tradition and revolution in ESL teaching. *TESOL Quarterly*, *19*, 535–552.

Ransdell, S. (1995). Generating thinking-aloud protocols: Impact on the narrative writing of college students. *American Journal of Psychology, 108*(1), pp. 89–98.

Ransdell, S. & Levy, C. (1994). Writing as process and product: the impact of tool, genre, audience, knowledge, and writer expertise. *Computers in Human Behavior, 10*, 511–527.

Reynolds, T., & Bonk, C. (1996). Facilitating college writers' revisions within a generative-evaluative computerized prompting framework. *Computers and Composition, 13*, 93–108.

Salaberry, R. (1999). CALL in the year 2000: Still developing the research agenda. A commentary on Carol Chapelle's CALL in the year 2000: Still in search of research paradigms. *Language Learning & Technology, 3*(1), 104–107. Retrieved from http://llt.msu.edu/vol3num1/comment/index.html

Scott, V. M., & New, E. (1994). Computer-aided analysis of foreign language writing strategies. *CALICO Journal, 11*, 5–18.

Silva, T. (1990). Second language composition instruction: Developments, issues and directions in ESL. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 11–23). New York: Cambridge University Press.

Slattery, S. (2005). Technical writing as textual coordination: An argument for the value of writers' skill with information technology. *Technical Communication, 52*(3), 353–360.

Slattery, S. (2007). Un-distributing work through writing: How technical writers manage texts in complex information environments. *Technical Communication Quarterly, 16*(3), 311–325.

Stallard, C. K. (1974). An analysis of the writing behavior of student writers. *Research in the Teaching of English, 8*, 206–218.

Sullivan, K., & Lindgren, E. (Eds.). (2006). *Computer key-stroke logging and writing: Methods and applications*. Oxford: Elsevier.

Swarts, J. (2004). Technological mediation of document review: The use of textual replay in two organizations. *Journal of Business and Technical Communication, 18*(3), 328–360.

Swarts, J. (forthcoming 2010). Recycled writing: Assembling actor-networks from reusable content. *Journal of Business and Technical Communication*.

Thorne, S. L. (2003). Artifacts and cultures-of-use in intercultural communication. *Language Learning & Technology, 7*(2), 38–67. Retrieved from http://llt.msu.edu/vol7num2/thorne/default.html

Thorson, H. (2000). Using the computer to compare foreign- and native-language writing processes: A statistical and case study approach. *The Modern Language Journal, 84*, 55–70.

Tribble, C. (1990). Concordancing in an EAP writing program. *CAELL Journal, 1*(2), 10–15.

Trimbur, J. (1994). Taking the social turn: Teaching writing post-process. *College Composition and Communication, 45*, 108–118.

Van Ittersum, D. (2009). Distributing memory: rhetorical work in digital environments. *Technical Communication Quarterly, 18*(3), 259–280.

Warschauer, M. (1997). Computer-mediated collaborative learning: Theory and practice. *The Modern Language Journal, 81*, 470–481.

Warschauer, M. (1998). Researching technology in TESOL: Determinist, instrumental, and critical approaches. *TESOL Quarterly*, *32*(4), 757–761.

Warschauer, M., & Kern, R. (Eds.). (2000). *Network-based language teaching: Concepts and practice*. New York: Cambridge University Press.

Zamel, V. (1982). The process of discovering meaning. *TESOL Quarterly*, *16*(2), 195–209.

Zamel, V. (1983). The composing processes of advanced ESL students: Six case studies. *TESOL Quarterly, 17*, 165–187.