

Analyzing Pauses in Computer-Assisted EFL Writing—A Computer-Keystroke-Log Perspective

Cuiqin Xu* and Yan Qi

Nanjing University, Nanjing, China // cqxu@nju.edu.cn // yzqiyan@sina.com

*Corresponding author

(Submitted April 10, 2016; Revised July 5, 2016; Accepted August 15, 2016)

ABSTRACT

Using computer keystroke logs, this study investigated how writing skill affected L2 writers' pausing patterns to gain insights into their management of the cognitive writing processes. The 59 participants, 29 in the more-skilled group and 30 in the less-skilled group, were recruited from a college English course at a key Chinese university. The two groups completed an argumentative essay in a computer classroom where Inputlog6.0 was installed to log their writing activities. Setting the pause threshold at 2 seconds, the study examined both the global pausing patterns and the interval pausing patterns by dividing each writing event into five equal intervals, and how the final text quality related to the pausing patterns. The results showed a significant effect of writing skill on the interval pausing patterns, but not on the global pausing patterns. Correlating significantly with the final text quality, the interval pausing patterns also revealed important differences in L2 writers' management of writing processes in terms of how one writing process dominates at specific intervals with interaction and shifts between other processes in a recursive fashion. Pedagogical implications are then discussed in light of these findings.

Keywords

L2 writing skill, Pauses, Writing process, Writing intervals

Introduction

Written text production involves complex cognitive processes that place heavy demands on the writers' working memory (Kellogg, 1996; Olive, Kellogg, & Piolat, 2008). When writers experience cognitive overload in their working memory, some processes would be suspended or even sacrificed to accommodate the immediate call for a specific process (DeKeyser, 2001; de Larios, Manchón, & Murphy, 2006) and pausing is a strategy that writers actively or passively adopt to free up attentional resources for processes of immediate priority. Therefore, pauses may serve as windows to writers' writing activities, such as "phrasing, memory such, decision, feedback, conceptual integration, and so forth" (de Beaugrande, 1984, p. 166). In addition, although writing consists of major processes as planning, translating (i.e., putting ideas into visible language) and revising (Hayes & Flower, 1980), these processes are not activated linearly but interact with one another in a recursive fashion, with one process dominating while other processes remaining dormant until being reactivated. While Olive and Kellogg (2002) suggest skilled writers can activate transcription concurrently with translating, planning and revising are mainly activated during pauses (Alves, Castro, & Olive, 2008; Olive, Alves, & Castro, 2009). Therefore, pausing may reveal possible problems and writing strategies behind such shifts of writing processes (Wengelin, 2006). As writing performance depends on how writing processes are organized during the composition process (Van den Bergh & Rijlaarsdam, 2007), pausing, being windows to writing activities and cognitive processes during text production, merits more research attention if the writing process is to be better understood.

Many pause studies were conducted in pen-and-paper settings and mostly adopted video-recording or think-aloud protocols to probe into the writing process. While both methods have produced enlightening results for the comprehension of writing and think-aloud protocols have, in particular, advanced the knowledge of how pauses are filled qualitatively, these methods are limited to case studies due to practical reasons. In addition, there have been wide concerns over the intimidating nature of video-recording and the reactivity issues of thinking-aloud in writing research (for a review, see Yang, Hu, & Zhang, 2014).

With computer arising to be the leading writing medium for academic purpose, writing has witnessed much difference from that in pen-and-paper settings (Alves, Castro, Sousa, & Strömqvist, 2007; Olive & Kellogg, 2002). Meanwhile, with the advent of computer science and technology, a number of keystroke logging tools have been developed and greatly promoted written production research (Leijten & Van Waes, 2006). These computer programs "log and time stamp keystroke activities to reconstruct and describe the text production processes" (Leijten & Van Waes, 2013, p. 359), thus enabling writing research to examine the online writing process both multi-dimensionally and at higher accuracy levels. More importantly, these programs, by running in the background, minimize writers' anxiety of being observed. Informed by recent development in writing

research and theory, these programs also provide rather sophisticated analyses of the writing process, thus more opportunities for refined writing research (Latif, 2009; Sullivan & Lindgren, 2006; Van Waes, Leijten, Wengelin, & Lindgren, 2012).

As different writing processes draw on the same working memory pool (Kellogg, 2001; McCutchen, 2000), the activation or deactivation of certain processes provides insights into the writers' management of their cognitive processes during the text production. With the rationale behind keystroke logging that "writing fluency and flow reveal traces of the underlying cognitive processes" (Leijten & Van Waes, 2013, p. 360), this study, using Inputlog6.0 to log L2 writing activities, endeavors to examine how pauses provide insights into L2 writers' management of their writing process in computer-assisted writing settings.

Literature review

Pauses in pen-and-paper writing settings

Early writing research identifies three major processes in writing, i.e., planning, translating and revising (Hayes & Flower, 1980) and suggests that the ability to consciously manage these writing processes is a fundamental component of writing skill, contributing much to producing good-quality texts (Levy & Ransdell, 1995). Many studies gained insights into these cognitive processes by probing into pauses during text production. Early pause studies in pen-and-paper writing settings mainly used video-taping or think-aloud protocols and had long pause thresholds for practical reasons. Schumacher, Klare, Cronin and Moses (1984), setting the pausing threshold at 10 seconds, videotaped and compared the pausing of 22 high school graduates and 20 college students through a 30-minute writing assignment. The study found no differences in pause frequency between the two groups, but the high school group paused averagely longer than the college group. Dividing the writing events into four intervals, the researchers reported that two groups did not differ significantly in pause duration at different writing intervals. However, more cognitive activities were documented for the college group while more grammatical activities were observed for the high school group during the pauses.

Many pausological studies have focused on the relationship between pause duration and the grammatical locations of such pauses, i.e., within-word, between words/clauses/sentences or at T-unit boundaries (Matsuhashi, 1987; Spelman Miller, 2000; Wengelin, 2006). These studies generally suggest that pause duration increases with the grammatical unit level, i.e., pauses within a word are shorter than those proceeding a word, and pauses between sentences are shorter than those between paragraphs. Other researchers reported similar findings that grammatical, discourse and morphological boundaries affect pause length (Nottbusch, Grimm, Weingarten, & Will, 2005; Spelman Miller, 2006). Some pausological studies tried to relate writing pauses with writing processes during the composition process. Using think-aloud protocols and reaction-time tasks, Beauvais, Olive and Passerault (2011) examined pauses at different writing stages and reported that text quality positively correlated with pause length at the prewriting stage. However, as there are wide concerns over to what extent think-aloud protocols would affect the writing process (see Yang, Hu, & Zhang, 2014) and whether the dual-task mode (reaction-time task and composition task) would genuinely reflect the actual writing process, further research into the writing processes through pauses are in need.

Pauses in computer-assisted writing settings

With the development of computer keystroke logging tools, pausological studies have been able to log and observe the writing process in a non-intruding manner and therefore have greatly pushed the research boundaries imposed by research instruments. Van Waes and Schellens (2003) compared the pausing behavior of experienced writers in computer keyboard settings against pen-and-paper settings. Setting the pausing threshold at 3 seconds and dividing writing into segments of 10 minutes, the study found that long pauses evenly distributed throughout the writing event in computer settings while half of the long pauses concentrated at the beginning stage of writing in paper settings. The researchers thus concluded that writers tend to begin writing sooner in keyboard settings than in paper settings and that the use of a word processor would result in more fragmented writing processes. As differences have been clearly documented in pausing patterns in pen-and-paper writings against computer keyboard writings, more studies are in need to better understand the nature of writing with computer keyboard.

Mixed research results have also been reported with regard to how writing skill affects pause frequency and pause duration in computer-assisted writings. Spelman Miller (2000) reported that L2 writers paused longer in

all grammatical locations compared with L1 writers, suggesting that pause duration is a function of both language proficiency and writing skill. Spelman Miller, Lindgren, and Sullivan (2008) further examined pause frequency and pause duration of L2 writers through a 3.5-year longitudinal study. The study found that as their writing abilities increased, L2 writers paused less frequently, but their pause duration did not change significantly over time, suggesting that writing ability affected pause frequency but not pause duration. Furthermore, Xu and Ding (2014) reported that while skilled and less-skilled L2 writers did not differ in their pausing patterns from a global perspective, skilled L2 writers paused less frequently and significantly longer at the prewriting stage than their less-skilled peers, showing that writing skill affected pause frequency and pause duration at different functional stages of writing. The study also found that text quality significantly correlated with the prewriting pause duration, rendering it important to examine pause in a more refined manner than from a global perspective.

To sum up, while previous pause studies adopt various and relatively long pause thresholds due to practical reasons, computer logging tools enable L2 writing researchers to examine pauses of shorter thresholds for better understandings of L2 writers' pausing behavior as well as their process management behind these pauses. In addition, although direct comparisons of pausing patterns at different stages or time segments are insightful, it is undeniable that a writer's pausing pattern carries much personal traits and the management of writing process is a function of personal knowledge system and writing expertise. Therefore, it would be insightful to examine how L2 writers' pausing patterns vary over the temporal development of writing, as indicated by the pausing patterns at different time intervals of the writing event. Such a probe could neutralize the effect of writer difference and gain more understanding of the L2 writing processes. However, precise estimate of pausing from an inter-writer perspective is still lacking.

This study attempts to investigate how pauses vary in terms of frequency and duration at different time intervals to gain insights into how writers temporally manage their writing processes. Specifically, this study attempts to investigate whether writing skill affects the pausing pattern globally and at different intervals of the writing event to gain insights into the cognitive processes underlying the written text production. For these ends, this study attempted to address the following research questions:

- Does writing skill affect EFL writers' global pausing patterns?
- Does writing skill affect EFL writers' pausing patterns at different intervals?
- How do EFL writers' global and interval pausing patterns relate to the quality of their final texts?

Method

Participants

The participants of the study ($N = 59$) were recruited from a pool of 118 sophomores who enrolled in a general college English course at a Chinese University. The students, between 18 to 20 years old, had been learning English for an average of 10 years, ranging from 8 to 11 years. Roughly, the students were of intermediate and upper-intermediate levels, or B1 and B2 as described in *Common European Framework of Reference for Languages: Learning, Teaching, Assessment* (Council of Europe, 2011). For the English course in which data were collected for this study, the 118 students were assigned into four classes according to the alphabetical order of their surnames. They regularly had four hours of English classes per week, two hours of listening and speaking plus two hours of reading and writing. As a routine, the participants had their writing sessions in a computer classroom, where each had a desktop computer at their disposal. The participants were fairly proficient with typewriting on computer as they had compulsory computer classes since primary schools with a focus on operation skills. In addition, college students in China, with few exceptions, have laptop computers and routinely hand in e-copies for their course work.

Upon entering college, all students took the English placement test, which included four multiple-choice sections of 85 points in total and a writing section of 15 points. Among the 118 students, those who ranked the top 35 in the writing section formed the more-skilled group and those who ranked the bottom 35 formed the less-skilled group. Eventually, 59 out of the 70 students who completed the study entered data analysis, with 29 in the more-skilled group and 30 in the less-skilled group. The 59 students participated in the study in their original classes during data collection and were only regrouped and combined for data analyses.

An independent-samples t -test was run on the 59 participants' writing scores in the placement test and the results suggested that the more-skilled group ($M = 11.26$, $SD = 1.04$) and the less-skilled group ($M = 6.32$, $SD = .90$) differed significantly in their writing scores ($t = 19.22$, $p = .00 < .05$, Cohen's $d = 5.08$).

Procedures

All data were collected in a computer classroom, where the students regularly had their reading and writing classes. Before the course began, a writing log program, Inputlog6.0, was installed on each computer and for each writing session, the students were asked to start the program and input the necessary personal data, i.e., name, age, gender, group, years of English learning. Before the data collection for this study, the students had completed two writing tasks using Inputlog and become familiar with writing in Inputlog.

At the beginning of the class when data were collected for this study, one drunk driving case reported by a local newspaper was briefly discussed in class and then the essential elements for argumentation learned earlier in their reading class were reviewed. Then, the students were asked to develop an argumentative essay on *whether drunk drivers should be imprisoned on their first offense*. They were told to develop a complete piece of argumentation on the given topic, with no requirement about the specific length of the final product. The students were asked to finish their writing in class, within a maximum length of 90 minutes.

With a click on the “Start Recording” button, a normal word processor popped up and the students began their writings, with Inputlog6.0 logging their writing events meanwhile. After finishing their writing, the students could simply click “Stop Recording” and the program would automatically exit and save the word documents in a pre-specified file folder. The instructor then collected the student writings through a management software pre-installed on all computers.

After the data collection, all students were informed of the research purpose of this study and signed the agreement of participation on a voluntary basis. As a return, one of the researchers offered the students individualized face-to-face discussion of their writings.

Data coding and analysis

Global analysis of pauses

Following previous pausological studies (Wengelin, 2006; Xu & Ding, 2014), this study also set the pause threshold at 2 seconds. Each writing event was analyzed globally in terms of event time, active writing time, pausing time, global pause frequency and pause duration. As the writing event time was different for each participant, raw global pause frequency was converted into standardized global pause frequency of 60 minutes through the following formula:

$$\text{Standardized global pause frequency} = (\text{Raw global pause frequency} / \text{event time (seconds)}) * 3600 \text{ (seconds)}$$

Interval analysis of pauses

Each writing event was divided into five temporal intervals of equal length by setting “Number of Intervals” at 5 in Inputlog6.0. Pause frequency and pause duration were compared respectively at each interval between the two groups and within each group across the five intervals. As the interval length varied from one writer to another, raw interval pause frequency was converted into standardized interval pause frequency of 10 minutes by following the formula below:

$$\text{Standardized interval pause frequency} = (\text{Raw interval pause frequency} / \text{Interval length (seconds)}) * 600 \text{ (seconds)}$$

Text scoring

Before scoring, all texts were coded with consecutive numbers to remove student identities. The writing rating rubric (see Appendix) was adapted from CET4 (College English Test, Band 4), a nation-wide English test in China. Two experienced college EFL teachers, who did not teach any of the students before or during data collection and had extensive experience in CET4 writing scoring, conducted the scoring separately. As the maximum score for a text is 15, the text would be returned to the two raters for rescoreing if the score discrepancy exceeded three points.

The inter-rater reliability was checked using Pearson correlation coefficients. The initial inter-rater reliability was $r = .83$ ($p = .00 < .05$) and it reached $r = .90$ ($p = .00 < .05$) after the rescoring for score discrepancy over three points. Then, the scores from two raters were averaged as the final text score for further analysis.

Statistical analysis

A series of independent-samples t -test were performed for possible between-group difference both globally and at different intervals, and repeated measure ANOVAs were conducted for the within-subject differences across the five intervals. In addition, Pearson correlation coefficients were computed to examine how writing pauses related to the final text quality. For all statistical tests, the significance level was set at .05.

Results

Global pausing patterns

Table 1 presents an overview of time allocation and pausing patterns for the two groups. The independent-samples t -test results suggest that the two groups did not differ significantly with regard to the total event time ($t = -1.64$, $p = .11 > .05$), active writing time ($t = -1.17$, $p = .25 > .05$), total pausing time ($t = -.68$, $p = .50 > .05$), global pause frequency ($t = -.22$, $p = .83 > .05$) or global pause duration ($t = -.06$, $p = .95 > .05$), suggesting that the two groups did not differ significantly in terms of time allocation or pausing patterns from a global perspective.

Table 1. Task overview

	Less-skilled ($n = 30$)		More-skilled ($n = 29$)		t	p
	M	SD	M	SD		
Total event time (s)	3808.10	610.54	4061.45	575.46	-1.64	.11
Active writing time (s)	2053.77	391.90	2195.62	528.25	-1.17	.25
Total pausing time (s)	1754.33	655.95	1865.83	593.18	-.68	.50
Global pause frequency	213.10	40.00	215.17	32.92	-0.22	.83
Global pause duration (s)	7.74	2.00	7.78	2.44	-0.06	.95

Interval pausing patterns

Interval pause frequency

Table 2 summarizes the statistics of standardized pause frequency at the five intervals respectively. The independent-samples t -test results suggest that the less-skilled group paused significantly less frequently than the more-skilled group at Interval 2 ($t = -2.34$, $p = .02 < .05$) and the two groups were not significantly different in pause frequency in any other intervals.

Table 2. Interval pause frequency

	Less-skilled ($n = 30$)		More-skilled ($n = 29$)		t	p	Cohen's d
	M	SD	M	SD			
Interval 1	34.69	10.07	30.61	6.78	1.82	.07	--
Interval 2	33.32	10.13	38.60	6.80	-2.34	.02*	.61
Interval 3	36.42	6.35	36.48	7.46	-0.03	.98	--
Interval 4	38.77	9.41	38.74	8.75	0.01	.99	--
Interval 5	34.43	9.49	34.91	7.89	0.21	.83	--

Note. * $p < .05$.

For how each group differed in pause frequency across the five intervals, repeated measures ANOVAs were conducted with the within-subjects factor being time intervals and the dependent variable being the standardized interval pause frequency. Figure 1 plots the pause frequency at each writing interval for the two groups.

For the less-skilled group, Mauchly's test indicated that the assumption of Sphericity was not violated, $X^2(9) = 10.23$, $p > .05$. Tests of within-subjects effects show that there was significant effect of time interval on pause frequency, $F(4, 116) = 2.71$, $p = .03 < .05$, $\eta_p^2 = .09$, suggesting that 9% of the variation in pause frequency of

the less-skilled group was accounted for by time interval. Pairwise comparisons suggest that the less-skilled group paused significantly more frequently at Interval 4 than at Interval 1, Interval 2 and Interval 5.

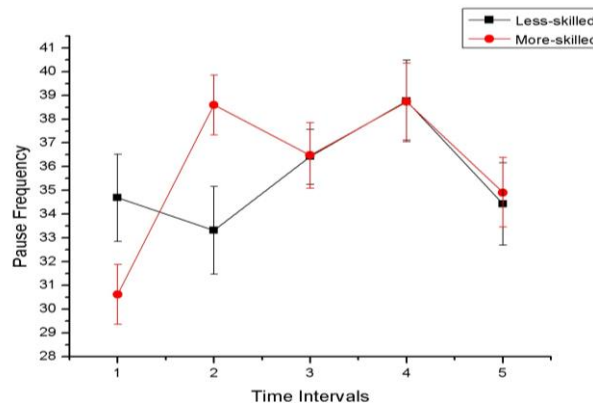


Figure 1. Pause frequency at the five intervals (Mean ± S.E.)

As for the more-skilled group, Mauchly's test indicated that the assumption of Sphericity was violated, $X^2(9) = 18.52, p < .05$; therefore, degrees of freedom were corrected using Huynh-Feldt estimates of Sphericity ($\epsilon = .87$). Tests of within-subjects effects show that there was significant effect of time interval on pause frequency, $F(3.47, 97.03) = 9.47, p = .00 < .05, \eta_p^2 = .25$, suggesting that 25% of the variation in pause frequency of the more-skilled group was accounted for by time interval. Pairwise comparisons suggest that the more-skilled group paused significantly less frequently at Interval 1 than in other four intervals and paused significantly more frequently at Interval 2 than at Interval 4.

Interval pause duration

Table 3 presents the statistics for pause duration at the five intervals. The independent-samples *t*-test show that the more-skilled group paused significantly longer than the less-skilled group at Interval 1 ($t = -2.52, p = .02 < .05$), but significantly shorter than the less-skilled group at Interval 2 ($t = 2.17, p = .04 < .05$).

Table 3. Interval pause duration

	Less-skilled ($n = 30$)		More-skilled ($n = 29$)		<i>t</i>	<i>p</i>	Cohen's <i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
Interval 1	7.34	2.68	10.11	5.38	-2.52	.02*	-.65
Interval 2	9.28	3.93	7.43	2.45	2.17	.04*	.56
Interval 3	7.70	2.26	7.60	2.83	0.14	.89	--
Interval 4	7.45	2.83	7.45	3.05	0.00	.99	--
Interval 5	8.16	4.19	8.04	5.27	0.10	.93	--

Note. * $p < .05$.

For how each group differed in pause duration across the five intervals, repeated measures ANOVAs were conducted with the within-subjects factor being time interval and the dependent variable being pause duration. Figure 2 plots the pause duration at each writing interval for the two groups.

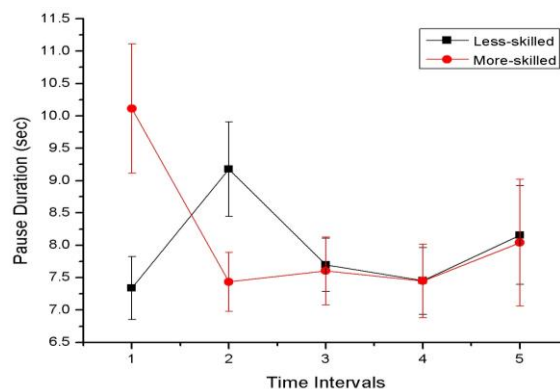


Figure 2. Pause duration at the five intervals (Mean ± S.E.)

For the less-skilled group, Mauchly's test indicated that the assumption of Sphericity was violated, $X^2(9) = 28.64, p < .05$; therefore, degrees of freedom were corrected using Greenhouse-Geisser estimates of Sphericity ($\epsilon = .61$). Tests of within-subjects effects show that there was no significant effect of interval on pause duration, $F(2.44, 70.66) = 2.69, p = .06, \eta_p^2 = .09$, suggesting that the effect of interval on pause duration approached the significance level. As Howell (2010) suggests that most of the multiple comparison procedures do not require an overall significant ANOVA group effect (pp. 372-373), pairwise comparisons were conducted. The results suggest that pause duration at Interval 2 were significantly longer than those at Interval 1, Interval 3 and Interval 4, with no significant difference in pause duration between any other two intervals.

For the more-skilled group, Mauchly's test indicated that the assumption of Sphericity was violated, $X^2(9) = 58.28, p < .05$; therefore, degrees of freedom were corrected using Greenhouse-Geisser estimates of Sphericity ($\epsilon = .59$). Tests of within-subjects effects show that there was significant effect of time interval on pause duration, $F(2.37, 66.44) = 3.76, p = .02 < .05, \eta_p^2 = .12$, suggesting that 12% of the variation in pause duration of the more-skilled group was accounted for by time interval. Pairwise comparisons suggest that pause duration at Interval 1 were significantly longer than those at Interval 2, Interval 3 and Interval 4, but not longer than that at Interval 5. This shows that Interval 1 is significantly different from the middle three intervals, displaying possible differences of the more-skilled group in managing their writing processes at different intervals.

In summary, the pausing patterns of more-skilled and less-skilled writers mainly differed during the first two intervals. The more-skilled group paused less frequently and longer at Interval 1, but paused shorter and more frequently as the writing went on; by contrast, the less-skilled group displayed more frequent yet shorter pauses at Interval 1, and less frequent yet longer pauses at Interval 2. In addition, both groups witnessed a slight increase in pause duration and a decrease in pause frequency at Interval 5.

The final text quality and pausing patterns

Table 4 presents the writing product data for the two groups in terms of process words (number of words produced during the process), product words (number of words in the final product) and the final text quality (text score). As shown in Table 4, compared with the less-skilled group, the more-skilled group wrote significantly more words during the process ($t = -2.84, p = .01 < .05$), produced significantly longer final texts ($t = -3.46, p = .00 < .05$) and their final texts were of significantly better quality ($t = -7.49, p = .00 < .05$).

Table 4. Writing product data

	Less-skilled ($n = 30$)		More-skilled ($n = 29$)		t	p	Cohen's d
	M	SD	M	SD			
Process words	315.87	74.22	371.48	76.47	-2.84	.01*	-0.74
Product words	241.80	42.96	284.34	51.19	-3.46	.00*	-.90
Text score	7.68	1.38	11.05	2.02	-7.49	.00*	-1.95

Note. * $p < .05$.

Two-tailed Pearson correlation coefficients were computed to see how the global and interval pausing patterns related to the final text quality. The results suggest that the final text quality did not correlate significantly with either global pause frequency ($r = -.08, p = .57 > .05$) or global pause duration ($r = .13, p = .33 > .05$). However, the final text quality correlated positively with pause duration at the first interval ($r = .38, p = .00 < .05$) but negatively with pause frequency at the first interval ($r = -.25, p = .05$). No significant correlation was observed between the final text quality and pause frequency or pause duration at other intervals.

To sum up, the more-skilled group were significantly more productive than the less-skilled group during the writing event and produced significantly better final texts. Significant correlations were observed between the final text quality and pauses at the first interval, showing that the final text benefited from long and infrequent pauses at the beginning of the writing event.

Discussion

This study set out to examine how writing skill affected writing pauses to gain insight into L2 writers' management of their writing processes. The answers to the research questions have specified important pausing patterns both globally and at different temporal intervals and how pausing patterns related to the final text quality.

First, more-skilled and less-skilled L2 writers did not differ significantly with regard to writing time allocation or global pausing patterns, showing that writing skill does not affect the intermediate L2 writers' allocation of writing time from a global perspective. This confirms Xu and Ding's (2014) finding that L2 writers of intermediate level did not distinguish themselves in the global pausing patterns, testifying to some commonality in their management of L2 writing processes. Together with the fact that the final text quality did not correlate with either global pause frequency or global pause duration, this study suggests the global pausing pattern is not a good indicator of either writing skill or final text quality. However, the fact that the two groups differed significantly in the final text length and quality suggests more productive use of the writing time for the more-skilled group, lending support to Ferrari, Bouffard, and Rainville's (1998) finding that poor college students wrote shorter texts than their peers. Pedagogically, this study shows writing productivity is not a simple function of writing time and L2 writers need instruction on more qualitative use of their writing time from a global perspective.

Second, more-skilled and less-skilled L2 writers differed in their interval pausing patterns and the final text quality significantly correlated with pauses during the first interval, showing that writing skill affects L2 writers' management of writing processes and consequently, their writing products. Spelman Miller (2000) reported that pause duration increased at larger grammatical unit boundaries and signaled significant planning (p. 142). In this study, the interval pausing patterns suggest the more-skilled group may engage in much global planning at Interval 1, displaying long and infrequent planning pauses; at Interval 2, this group engaged in more focused translating of ideas, displaying short and more frequent translating pauses. By contrast, the less-skilled group displayed frequent and short pauses in Interval 1, suggesting that this group started writing much sooner and Interval 1 related to both planning and translating. Moving to Interval 2, the less-skilled group displayed less frequent but longer pauses, showing that their focused translating process was frequently interrupted by other processes. While Spelman Miller et al.'s (2008) found that global pause frequency declined but global pause duration did not change as writing abilities developed, this study documented differences in the interval pausing patterns between more-skilled and less-skilled L2 writers and therefore suggests the importance of examining pauses at fine-grained temporal intervals.

The positive correlation between pause duration at Interval 1 and the final text quality supports Beauvais et al.'s (2011) finding on pen-and-paper writings that text quality benefits from long pauses at the prewriting stage. In addition, Ferrari et al. (1998) reported that long episodes of pausing during translating emerges as a result of insufficient planning at the prewriting stage. When writing breaks down, writers would have to strive for global conceptualization and planning in the middle of translating. As writing is constantly restructured and redefined by the produced text, insufficient global planning could impose extra cognitive load on working memory with assessing and integrating new content into the existing text. Since text quality depends on the ability to plan globally before actual drafting and use this plan to guide the writing process (Chai, 2006; De La Paz & Graham, 2002; MacArthur, Harris, & Graham, 1994; Xu & Ding, 2014), L2 writers should be pedagogically reassured about the importance of global planning at the beginning of writing and that a hasty start of drafting does not gain them advantage in their final products.

Thirdly, this study documents significant effects of time interval on pausing patterns, manifesting transitions of writing processes in a recursive fashion along the temporal development of writing. The pauses of the more-skilled group at Interval 1 were significantly different from other intervals except the last one. Informed by early writing models (Hayes & Flower, 1980; Kellogg, 1996), this pausing pattern clearly marks out the planning, translating and revising stages during the composition process. By contrast, the less-skilled writers paused significantly longer at Interval 2 than other intervals except the last one. This suggests that the less-skilled group had much shorter initial planning process and began their translating process much sooner (Ferrari et al., 1998). When their writing encountered breakdowns, the less-skilled group had to engage in further planning at Interval 2, followed by focused translating at Interval 3 and Interval 4 and revising at Interval 5. Therefore, the less-skilled group experienced planning, translating, planning, translating and revising in their composition process.

Although the writing process is highly interactive and recursive among various component processes (Hayes & Flower, 1980; Kellogg, 1996), this study suggests an effect of writing skill on how different writing processes dominate and shift at different intervals: the more-skilled group competently managed their writing processes with well-defined functional stages, while the less-skilled group exhibited less competent management of writing processes and juggled back and forth with poorly-defined functional stages. This confirms previous findings from think-aloud protocols that more-skilled writers concentrated more on formulation (translating in Hayes & Flower's model) during the central stages of writing (de Larios et al., 2006; de Larios, Manchón, Murphy, & Marín, 2008). As the transitions of different processes may signal writers' strategy use (Levy & Ransdell, 1995), this study shows L2 writing benefits from efficient global planning at the beginning and focused production

during the translating process. Pedagogically, instructors need to cultivate L2 writers' awareness of process management (see Mikulski & Elola, 2011) and strategy use (De Silva, 2015) to better scaffold their L2 writing development.

Conclusions and suggestions for future research

By analyzing how L2 writers of different writing skills paused in an argumentative writing task, this study has disclosed interesting pausing patterns that are insightful for comprehending the L2 writing process. First, writing skill does not affect time allocation or pausing patterns from a global perspective, testifying to the importance of qualitative use of task time. Second, writing skill influences interval pausing patterns, showing that writing products benefit from long planning pauses at the beginning of the writing. Finally, writing skill affects how writers manage their writing processes in terms of interaction and shifts between different component processes, with good writers displaying functionally well-defined writing stages while poor writers being forced to shift back and forth to sustain the writing. As effective writing demands competent process management and good strategy use (De Silva, 2015; Van den Bergh & Rijlaarsdam, 2001), L2 writers should strategically define clear goals for each writing stage to avoid draining their working memory capacity with multiple writing processes concurrently.

While good writing is an orchestration of good vocabulary, grammar, genre, topic and procedural knowledge, it is beneficial for L2 writers to consciously reflect on their writing processes and improve their management of such processes. Of course, it would oversimplify the issue by claiming that prolonged pauses contribute to text quality, but L2 writers need to be reassured that long pauses devoted to global planning lead to more fluent and productive writing processes. This positive understanding of pauses may help boost L2 writers' confidence in striving for sufficient global planning and lift their anxiety in long episodes of pausing during writing.

There are a number of limitations in the study that merit caution for interpreting the findings or designing future studies. First, this study investigated the temporal distribution of pauses, leaving the issue of how pauses were qualitatively filled unaddressed. Second, as short and frequent pauses were observed during the first temporal interval for the less-skilled group, this suggests the division of the writing event into five intervals was not sufficient to depict the transition from the planning process to the translating process for this group. Further research into pauses in more writing intervals would be necessary to fine grain the observations made in this study.

Acknowledgements

The research reported in this paper was funded by the Youth Project of Humanities and Social Sciences, Ministry of Education, China (No. 14YJC740095) and the National Social Science Foundation of China (No. 16BYY099). We would like to thank the reviewers and ETS editors for their insightful suggestions that help improve the paper.

References

- Alves, R., Castro, S., & Olive, T. (2008). Execution and pauses in writing narratives: Processing time, cognitive effort and typing skill. *International Journal of Psychology*, 43(6), 969-979.
- Alves, R., Castro, S., Sousa, L., & Strömquist, S. (2007). Influence in typing skill and pause-execution cycles in written composition. In M. Torrance, L. van Waes, & D. Galbraith (Eds.), *Studies in writing: Writing and cognition: Research and applications* (Vol. 20, pp. 55-66). Amsterdam, The Netherlands: Elsevier.
- Beauvais, C., Olive, T., & Passerault, J. (2011). Why are some texts good and others not? Relationship between text quality and management of the writing processes. *Journal of Educational Psychology*, 103(2), 415-428.
- Chai, C. (2006). Writing plan quality: Relevance to writing scores. *Assessing Writing*, 11(3), 198-223.
- Council of Europe (2011). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge, UK: Cambridge University Press.
- de Beaugrande, R. (1984). *Text production: Toward a science of composition*. Norwood, NJ: Ablex.

- De La Paz, S., & Graham, S. (2002). Explicitly teaching strategies, skills and knowledge: Writing instruction in middle school classrooms. *Journal of Educational Psychology, 94*(4), 687-698.
- de Larios, J., Manchón, R., & Murphy, L. (2006). Generating text in native and foreign language writing: A Temporal analysis of problem-solving formulation process. *The Modern Language Journal, 90*(1), 100-114.
- de Larios, J., Manchón, R., Murphy, L., & Marín, J. (2008). The foreign language writer's strategic behavior in the allocation of time to writing processes. *Journal of Second Language Writing, 17*(1), 30-47.
- De Silva, R. (2015). Writing strategy instruction: Its Impact on writing in a second language for academic purposes. *Language Teaching Research, 19*(3), 301-323.
- DeKeyser, R. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 125-151). Cambridge, UK: Cambridge University Press.
- Ferrari, M., Bouffard, T., & Rainville, L. (1998). What makes a good writer? Differences in good and poor writers' self-regulation of writing. *Instructional Science, 26*(6), 473-488.
- Hayes, J., & Flower, L. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3-30). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Howell, D. C. (2010). *Statistical methods for psychology* (8th ed.). Belmont, CA: Cengage Wadsworth.
- Kellogg, R. (1996). A Model of working memory in writing. In C. M. Levy & S. E. Ransdell (Eds.), *The Science of writing: Theories, methods, individual differences, and applications* (pp. 57-71). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Kellogg, R. T. (2001). Competition for working memory among writing processes. *American Journal of Psychology, 114*(2), 175-191.
- Latif, M. M. A. (2009). A State-of-the-art review of the real-time computer-aided study of the writing process. *International Journal of English Studies, 8*(1), 29-50.
- Leijten, M., & Van Waes, L. (2006). Inputlog: New perspectives on the logging of on-line writing. *Studies in writing, 18*, 73-94.
- Leijten, M., & Van Waes, L. (2013). Keystroke logging in writing research: Using Inputlog to analyze and visualize writing processes. *Written Communication, 30*(3), 358-392.
- Levy, C., & Ransdell, S. (1995). Is writing as difficult as it seems? *Memory and Cognition, 23*(6), 767-779.
- MacArthur, C. A., Harris, K. R., & Graham, S. (1994). Improving students' planning processes through cognitive strategy instruction. In E. C. Butterfield & J. C. Carlson (Eds.), *Advances in cognition and educational practice* (pp.173-198). Greenwich, CT: JAI Press.
- Matsuhashi, A. (1987). *Writing in real time: Modeling production processes*. Norwood, NJ: Ablex.
- McCutchen, D. (2000). Knowledge, processing, and working memory: Implications for a theory of writing. *The Educational Psychologist, 35*(1), 13-23.
- Mikulski, A., & Elola, I. (2011). Spanish heritage language learners' allocation of time to writing processes in English and Spanish. *Hispania, 94*(4), 715-733.
- Nottbusch, G., Grimm, A., Weingarten, R., & Will, U. (2005). Syllabic structures in typing: Evidence from hearing-impaired writers. *Reading & Writing, 18*(6), 497-526.
- Olive, T., & Kellogg, R. T. (2002). Concurrent activation of high- and low-level production processes in written composition. *Memory and Cognition, 30*(4), 594-600.
- Olive, T., Alves, R. A., & Castro, S. L. (2009). Cognitive processes in writing during pause and execution periods. *European Journal of Cognitive Psychology, 21*(5), 758-785.
- Olive, T., Kellogg, R. T., & Piolat, A. (2008). Verbal, visual and spatial demands during text composition. *Applied Psycholinguistics, 29*(4), 669-687.
- Schumacher, G. M., Klare, G. R., Cronin, F. C., & Moses, J. D. (1984). Cognitive activities of beginning and advanced college writers: A Pausal analysis. *Research in the Teaching of English, 18*(2), 169-187.
- Spelman Miller, K. (2000). Academic writers online: Investigating pausing in the production of text. *Language Teaching Research, 4*(2), 123-148.
- Spelman Miller, K. (2006). Pausing, productivity and the processing of topic in online writing. In E. L. Sullivan (Ed.), *Computer keystroke logging: Methods and applications* (Vol. 18, pp. 131-155). Oxford, UK: Elsevier.
- Spelman Miller, K. S., Lindgren, E. & Sullivan, K. P. (2008). The Psycholinguistic dimension in second language writing: Opportunities for research and pedagogy using computer keystroke logging. *TESOL Quarterly, 42*(3), 433-454.

- Sullivan, K., & Lindgren, E. (2006). *Computer key-stroke logging and writing*. Oxford, UK: Elsevier.
- Van den Bergh, H., & Rijlaarsdam, G. (2001). Changes in cognitive activities during the writing process, and relations with text quality. *Educational Psychology, 21*(4), 373-385.
- Van den Bergh, H., & Rijlaarsdam, G. (2007). The Dynamics of idea generation during writing: An On-line study. In M. Torrance, L. Van Waes, & D. Galbraith (Eds.), *Studies in writing: Writing and cognition: Research and applications* (Vol. 20, pp. 125-150). Amsterdam, The Netherlands: Elsevier.
- Van Waes, L., & Schellens, P. J. (2003). Writing profiles: The Effect of the writing mode on pausing and revision patterns of experienced writers. *Journal of Pragmatics, 35*(6), 829-853.
- Van Waes, L., Leijten, M., Wengelin, A., & Lindgren, E. (2012). Logging tools to study digital writing processes. In V. W. Berninger (Ed.), *Past, present, and future contributions of cognitive writing research to cognitive psychology* (pp. 507-533). New York, NY: Psychology Press.
- Wengelin, Å. (2006). Examining pauses in writing: Theories, methods and empirical data. In K. P. H. Sullivan & E. Lindgren (Eds.), *Computer key-stroke logging and writing: Methods and applications (Studies in Writing)* (Vol. 18, pp. 107-130). Oxford, UK: Elsevier.
- Xu, C., & Ding, Y. (2014). An Exploratory study of pauses in computer-assisted EFL writing. *Language Learning & Technology, 18*(3), 80-96.
- Yang, Y., Hu, G., & Zhang, L. (2014). Reactivity of concurrent verbal reporting in second language writing. *Journal of Second Language Writing, 24*(1), 51-70.

Appendix

Table 5. Rating Rubric for CET4 Writing

Points	Organization	Content	Language
0	No organization	Little idea	Incorrect spellings or a few isolated words
2	No focus; disorganized	Ideas random, inappropriate or illogical	Incomplete or incorrect sentences; severe errors that affect meaning
5	Attempts to focus; minimal organization	Ideas mixed; few transitions	Monotonous sentence structures; numerous errors that interfere with meaning
8	Single focus; some lapses or flaws in organization	Ideas not well supported or elaborated	Little variety in syntax; some evident errors
11	Single focus; logical organization	Ideas appropriate and varied	Varied sentence structure; few errors
14	Single, distinct focus; logical progression of ideas	Details effective, vivid, explicit and pertinent	Very few, if any, errors

Note. This rubric assigns each text into one of the levels above, and minor deviations from the description for each level except the “0” point level may result in the loss or gain of one point. Thus, a perfect score is 15 and the lowest score is zero.