Studies in Second Language Acquisition **41** (2019), 529–554
doi:10.1017/S027226311900010X

Research Article

WHAT CAN L2 WRITERS' PAUSING BEHAVIOR TELL US ABOUT THEIR L2 WRITING PROCESSES?

Khaled Barkaoui*

York University, Toronto

Abstract

When responding to a writing task, writers spend a significant amount of their time not writing. These periods of physical inactivity, or pauses, during writing provide observable and measurable cues as to when, where, and how long writers halt to plan and/or revise their texts. Consequently, examining writers' pausing patterns can provide important insights into the cognitive processes that writers employ when composing and the impact of various individual, task, and contextual factors on those processes. This article discusses theory and research on writers' pausing behavior; how pause analysis can be used to investigate second language (L2) learners' writing processes; challenges in researching writers' pausing behavior (e.g., defining pauses); and some strategies to address these challenges. Next, the article illustrates how L2 writers' pause data can be collected, analyzed, and interpreted, using keystroke logging data from a research project that aimed to examine the effects of task type, L2 proficiency, and keyboarding skills on L2 learners' writing processes when writing on the computer. The article concludes with a call for more research on L2 writers' pausing behavior, particularly how L2 writers' pausing behavior relates to L2 writing outcomes and development across learners, contexts, and time.

One striking characteristic of the writing process is that a large amount of time is spent not writing (Alves, Castro, de Sousa, & Stromqvist, 2007). Alamargot et al. (2007) estimated that "pausing can occupy up to 60 or 70% of total composition time" (p. 13). These "moments of physical inactivity during writing" (Matsuhashi, 1981, p. 114) provide observable and measurable cues as to when, where, and how long writers halt to plan and/or revise their texts (Schilperoord, 1996). Consequently, examining writers' pausing patterns can provide important insights into the cognitive processes that writers

The keystroke data used in this article is from a study funded by Educational Testing Service (ETS) under a Committee of Examiners and TOEFL research grant. ETS does not discount or endorse the methodology, results, implications, or opinions presented by the researcher. The opinions expressed in the report are those of the author.

*Correspondence concerning this article should be addressed to Khaled Barkaoui, Faculty of Education, York University, 253Winters College, 4700 Keele Street, Toronto, Ontario, M3J 1P3. E-mail: barkaoui@yorku.ca

Copyright © Cambridge University Press 2019

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

employ when composing and the impact of various individual and contextual factors on those processes. This article discusses theory and research on writers' pausing behavior; how pause analysis can be used to investigate second language (L2) learners' writing processes; challenges in researching writers' pausing behavior; and some strategies to address these challenges. The article then illustrates how L2 writers' pause data can be collected, analyzed, and interpreted, using keystroke logging data from a research project that aimed to examine the effects of task type, L2 proficiency, and keyboarding skills on L2 learners' writing processes when writing on the computer.

EXAMINING WRITERS' PAUSING BEHAVIOR

As Spelman Miller (2000, 2006b) explained, the interest in the pausological features of written text production was inspired by research into the temporal aspects of speech production such as the duration and distribution of pauses and utterances (e.g., Boomer, 1965; Butterworth, 1980). In this line of research, performance phenomena such as silent pauses, hesitations, repetitions, and false starts are seen as indicative of the mental processes underlying speech production, such as the local and global planning of the content and form of utterances. Pauses may also be indicative of problems in the planning or execution of utterances, such as difficulties in conceptualization and word finding (Wengelin, 2006, 2007). Similarly, in research on the temporal aspects of writing, pauses are seen as visible traces of such covert cognitive processes as goal setting, activating conceptual and linguistic knowledge, planning, memory and lexical searches, and internal revisions that contribute to written discourse production and quality (Chanquoy, Foulin, Fayol, 1996; Chenu, Pellegrino, Jisa, Fayol, 2014; Matsuhashi, 1981; Schilperoord, 1996; Schilperoord & Sanders, 1999; Spelman Miller, 2000, 2006a, 2006b; Wengelin, 2006, 2007).

Cognitive models of writing (e.g., Fayol, 1999; Field, 2004; Hayes, 1996, 2006; Kellogg, 1996; Torrance & Galbraith, 2006) offer an explanation of how and why pausing behavior can provide insights into writers' cognitive processes while writing (Alves et al., 2007; Chenoweth & Hayes, 2001; Fayol, 1999; McCutchen, 1996, 2000; Olive & Kellogg, 2002). According to these models, writing is a complex activity that requires the coordination of a variety of cognitive processes (e.g., planning, translation, execution, monitoring, revising) that place heavy demands on the writer's working memory (Fayol, 1999; Field, 2004; Kellogg, 1996; Kormos, 2012; Torrance & Galbraith, 2006). As Xu and Qi (2017) explained, with increasing demand by some processes writers may experience cognitive overload. As a result, other processes "would be suspended or even sacrificed to accommodate the immediate call for a specific process" (p. 24). In addition, writers may be able to activate some processes such as transcribing and translating concurrently, but may need to pause to activate other processes such as planning and revising (Xu & Qi, 2017). As Alamargot et al. (2007) argued, pauses can be intentional, based on a "strategic choice" by the writer to pause to engage in other activities (e.g., planning, revising), or imposed by a "limited cognitive capacity" that does not allow the writer to pursue multiple processes simultaneously (p. 27). From this perspective, pauses can serve as a window into writers' cognitive processes such as processing difficulties, planning, and evaluation that interrupt fluent text production (Baaijen, Galbraith, and de Glopper, 2012; Chenu et al., 2014; Fayol, 1999; Kellogg,

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

1996; Medimorec & Risko, 2017; Schilperoord, 1996; Xu & Qi, 2017; Zhang & Deane, 2015).

Research on pausing behavior in relation to planning illustrates how pauses during writing are interpreted (e.g., Schilperoord, 1996; Spelman Miller, 2000, 2006a, 2006b). This research assumes that pauses reflect planning and that longer pauses reflect more planning (Hayes & Nash, 1996). As Schilperoord (1996) explained, observed in real time, text production comes to us as patterns of pausing and writing, which reflect the planning-execution cycles that writers go through when writing. According to Schilperoord (1996), "[P]auses provide observable and measureable cues as to *where* and *how long* writers halt to plan and decide upon what to write next, whereas the linguistic output [be it producing new text or revising existing text] provides cues as to the nature of the hidden decision and planning processes" (p. 21; emphasis added). For example, long pauses are likely to be associated with macroplanning processes such as paragraph organization, whereas shorter pauses are likely to reflect microplanning processes such as grammatical and lexical choices (cf. Spelman Miller, 2006b).

It is important to emphasize here that pauses provide only *indirect* evidence of writers' underlying cognitive processes in writing (Chanquoy et al., 1996; Matsuhashi, 1981; Spelman Miller, 2006b; Wengelin, 2006). As Matsuhashi (1981) explained, "[T]he length of pauses … and their location in the text … provide a temporal taxonomy or description of real time aspects of written language production from which *inferences* about planning and decision making can be made" (p. 114; emphasis added). Alamargot et al. (2007) noted that, to be able to draw inferences about the distribution and sequencing of writers' underlying cognitive processes, studies of writers' pausing behavior make four key assumptions:

1. Pause duration varies as a function of the complexity of the processes engaged in;
2. Pause position within the hierarchical structure of the text [e.g., paragraph, sentence, word] indicates the nature of this processing;
3. The processes that occur during a pause concern the part of the text that will be written [or revised] immediately afterwards (processing adjacency principle); and
4. As the more demanding controlled processes cannot be engaged in parallel with graphomotor execution, they impose a writing pause (processing sequentiality principle). (p. 14)

These assumptions have been tested and confirmed in several studies on L1 (e.g., Alves et al., 2007; Chanquoy et al., 1996; Chenu et al., 2014; Matsuhashi, 1981; Medimorec & Risko, 2017; Schilperoord, 1996; Van Waes & Leijten, 2015; Wengelin, 2006, 2007) and L2 writing (e.g., Deane, 2014; Deane & Zhang, 2015; Phinney & Khouri, 1993; Révész, Kourtali, & Mazgutova, 2017; Sasaki, 2000; Spelman Miller, Lindgren, & Sullivan, 2008; Xu & Ding, 2014; Xu & Qi, 2017; Zhang & Deane, 2015). These studies show that pause locations and durations in writing are not arbitrary. For example, writers tend to pause more frequently and longer before paragraphs and before sentences than they do within and between words and phrases, suggesting different levels of processing at these text boundaries (e.g., Chanquoy et al., 1996; Matsuhashi, 1981; Medimorec & Risko, 2017; Phinney & Khouri, 1993; Schilperoord, 1996; Spelman Miller, 2000; Wengelin et al., 2009).

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

Other studies have found that pause frequency and duration vary significantly across the writing session (e.g., Xu & Ding, 2014; Xu & Qi, 2017), a finding that is consistent with research on writing processes, which suggests that the distribution of writing activities varies over the writing process (e.g., Breetvelt, Van den Bergh, & Rijlaarsdam, 1994; Roca de Larios, Manchón, Murphy, & Marín, 2008; Van der Hoeven, 1999; Van Weijen, Van den Bergh, Rijlaarsdam, & Sanders, 2008). Roca de Larios et al. (2008), for example, found that while low-proficiency writers maintained the same pattern of time allocation throughout the writing process, more proficient writers showed a more diversified time allocation to different writing activities during the writing process (e.g., more formulation in early stages and more revision in later stages). One implication of these findings is that it is important to examine not only *what* writers do (e.g., pause or write) but also the temporal aspects of writing (i.e., *when* they pause) (Breetvelt et al., 1994; Roca de Larios et al., 2008; Van der Hoeven, 1999; Van Weijen et al., 2008). Time during the writing process is a proxy variable for the changing "task situation" or writing context (Roca de Larios et al., 2008; Van Weijen et al., 2008). As the text evolves (e.g., through the addition of new content and/or the revision of already written text), the task situation changes and writers would adapt their writing activities to those changes.

Previous research also shows that pause frequency, duration, and location vary depending on various individual and contextual factors (e.g., Deane & Zhang, 2015; Medimorec & Risko, 2017; Révész et al., 2017; Xu & Ding, 2014; Xu & Qi, 2017). This is consistent with a cognitive view of writing; if task requirements or contextual factors affect L2 learners' writing processes, it is likely that such effects will manifest themselves in the form of different patterns of pausing behavior across writers and contexts. Deane and Zhang (2015), for example, noted that for proficient writers, "text tends to be produced efficiently in longer bursts [and] pauses are more likely to happen at natural loci for planning such as clause and sentence boundaries," while for less proficient writers, "text tends to be produced less efficiently, and pauses appear in locations that suggest difficulties in typing, spelling, word-finding, and other transcription processes" (p. 1).

Few studies have shown that the pausing patterns of skilled and less skilled writers differ significantly (e.g., Sasaki, 2000; Xu & Ding, 2014; Xu & Qi, 2017). Sasaki (2000), in a study comparing the pausing behaviors of expert and novice L2 writers, found that the experts paused longer before starting to write than did their novice counterparts. Additionally, the novices tended to pause at each thematic episode boundary more often than did the experts, which suggests that the novices were employing a "what next strategy," whereby they have to stop and plan what they are going to write next each time they finish writing one semantically coherent chunk. By contrast, once they made their global plan, the expert writers did not stop and think while writing as frequently as the novices did (cf. Van der Hoeven, 1999). Xu and Ding (2014) found that skilled L2 writers tended to pause less frequently and significantly longer at the prewriting stage than did less skilled writers. Similarly, Xu and Qi (2017) reported that the skilled and less skilled L2 writers in their study differed significantly in terms of their pausing patterns at different intervals of the writing process. For example, skilled L2 writers tended to pause longer but less frequently than did their less skilled peers in the first interval, perhaps because skilled writers engaged in more global planning at the beginning of the writing session. In contrast, less skilled writers tended to pause more frequently and for shorter durations in the first interval likely because they started writing much sooner. In the

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

second interval, the skilled writers engaged in more focused translating of ideas, displaying short and more frequent translating pauses, while the less skilled writers displayed less frequent but longer pauses, suggesting that their translating process was frequently interrupted by other processes such as lexical retrieval (p. 32). Finally, Spelman Miller et al. (2008) examined changes in the online writing processes of adult L2 learners. They found that, over a period of three years of L2 study, the participants increased both their fluency and the amount of text produced between interruptions (pauses and revisions); required fewer, but equally long, pauses; and decreased their total pause time.

Task characteristics can also affect writers' pausing patterns in L1 (e.g., Matsuhashi, 1981; Medimorec & Risko, 2017; Schilperoord, 1996) and L2 (Révész et al., 2017; Spelman Miller, 2000). In L1 writing, Matsuhashi (1981) found that the mean pause time was longest in the production of more cognitively demanding tasks (persuading and generalizing) than a less cognitively demanding task (reporting). The longer pause duration for the persuading and generalizing tasks was associated with increased attention to revisions reflecting conceptual, rather than formal, planning concerns. Schilperoord (1996) found that different writing tasks were more likely to influence the length of pauses between paragraphs, and to a lesser extent pauses between sentences and clauses, but not pauses between words and phrases. Schilperoord concluded that task differences are more likely to be manifested by differences in length of pauses at paragraph transitions indicating that task differences influence higher levels of planning (often conceptual planning). Word- and phrase-level processes, however, seem to be unaffected by task characteristics. Medimorec and Risko (2017) compared the pausing behaviors of university students when composing narrative and argumentative essays in L1. They found that pause rates were higher at word and sentence boundaries in argumentative essays.

Révész et al. (2017) examined the effects of task complexity (i.e., absence vs. presence of content support) on L2 writers' fluency, pausing, and revision behaviors. They found that providing content support led to significantly fewer pauses between sentences and more revisions below the clause level. In addition, more frequent pauses when content support was present were associated with the production of more complex language. Révész et al. concluded that content support can reduce processing burden on planning processes, allowing writers to devote more attention to linguistic encoding processes. In contrast, Spelman Miller (2000) found no significant differences in pause duration and location across descriptive and evaluative writing tasks, perhaps because the L2 writers in her study either lacked sensitivity to the rhetorical demands of the two types of tasks or were unwilling or unable to react to the specific directives given in the tasks.

Writers' pausing patterns can also vary as a function of other individual and contextual factors such as writing mode (computer or paper) and typing skills. First, several studies have indicated that writing on the computer calls for a different distribution of pausing patterns compared to writing on paper in L1 (e.g., Haas, 1989; Van Waes & Schellens, 2003) and L2 (e.g., Lee, 2004). Lee (2004), for instance, found that L2 learners paused longer and planned less initially when writing on the computer than they did when writing on paper, suggesting that planning and text production are more interwoven on the computer than they are on paper. Second, some studies have shown that typing skills can significantly affect writers' pausing patterns when writing on the computer in L1

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

(e.g., Alves et al., 2007) and L2 (e.g., Phinney & Khouri, 1993). Phinney and Khouri (1993), for example, found that both pause frequency and duration appeared to be influenced by L2 writers' level of word-processing experience.

Alves et al. (2007) compared the pausing patterns of slow and fast typists when writing in L1 on the computer. They found that the writing processes of slow typists were characterized by a significantly higher number of pauses and significantly shorter execution periods than those of faster typists. Additionally, slow typists took longer to compose their texts and tended to produce shorter texts. The author concluded that high typing skills allow the concurrent activation of writing processes. Because they could not think and type at the same time, slow typists seemed to use a serial way of composing, whereby they devoted pauses to high-level writing processes such as planning and monitoring and execution periods to typing. Slow typists also could not sustain execution periods for as long as fast typists because more resources were directed toward motor execution proper and thus fewer resources were available to other processes involved in online writing (p. 58). As Alves et al. put it, "[S]low typists need to expend so much cognitive effort on low-level processes [i.e., typing] during writing, that they often lose track of the content structure they had planned to realize, and therefore need to pause more often than fast typists, in order to retrieve the thread" (p. 64). As typing skill increases, the strategy of activating writing processes might shift from serial to parallel.

RESEARCHING WRITERS' PAUSING BEHAVIOR

The typical approach to investigating the writing process is the concurrent think-aloud protocol (e.g., Cumming, 1989). However, this approach has its limitations; in particular it is time-consuming and labor intensive and can alter the writing process and/or its outcomes (Barkaoui, 2011; Smagorinsky, 1994; Stratman & Hamp-Lyons, 1994). Alternative approaches to think-aloud protocols include retrospective stimulated recalls (e.g., Barkaoui, 2015; Bosher, 1998; Sasaki, 2000), analysis of video recordings of writing sessions (e.g., Matsuhashi, 1981, 1987; Van Waes & Schellens, 2003), and keystroke logging (e.g., Leijten & Van Waes, 2013; Spelman Miller et al., 2008). Keystroke logging, in particular, provides an unobtrusive and efficient approach to the study of writing processes (Abdel Latif, 2001; Leijten & Van Waes, 2006, 2013; Spelman Miller & Sullivan, 2006; Spelman Miller et al., 2008).

Keystroke logging programs are computer programs that log or record all the writing activities (e.g., keystrokes, pauses, mouse actions, scrolling, use of editing functions, deletion, insertion) that take place when a writer is writing on the computer together with the exact time of each action. During the writing process, this information is stored in a log file for later processing. This continuous data storage does not interfere with the writing process, creating an ecologically valid research context (Leijten & Van Waes, 2013). The output of the log file is a highly detailed record of the temporal features of the writing activity. For example, pauses are indicated in milliseconds within brackets, while revisions are indicated either as deletions or insertions of text. Furthermore, keystroke logging programs often have a replay facility that allows the replay of text exactly as it was inscribed (Abdel Latif, 2001; Spelman Miller et al., 2008). This replay function allows the observation and online analysis of the writing process "in real time as it unfolds" (Lindgren & Sullivan, 2006a, p. 157).

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

Keystroke logging data can be analyzed automatically and/or manually to reconstruct the on-screen writing process and examine the dynamics of text production in real time, particularly when and how long writers pause and when and what they revise (Abdel Latif, 2001; Baaijen et al., 2012; Barkaoui, 2016b; Spelman Miller et al., 2008; Sullivan & Lindgren, 2006a). Keystroke logging thus allows the close analysis of the linguistic and textual output to draw inferences about the cognitive processes from which it is generated (Spelman Miller, 2006; Spelman Miller & Sullivan, 2006). Keystroke logging can also offer insights into how texts develop under certain conditions (e.g., different task types) and problems encountered during the writing process (Leijten & Van Waes, 2013; Leijten, Van Waes, Schrijver, Bernolet, & Vangehuchten, 2019; Lindgren & Sullivan, 2006a). As Deane (2014) argued, "[T]he analysis of keystroke logging features can provide information about the kinds of behaviors in which writers engage during the writing process" that can be used to draw *inferences* about "the prevalence of planning and editing behaviors and about how these behaviors are interleaved with text production" (p. 2).

Figure 1 shows a portion of the keystroke logging data for one writing session together with the corresponding section from the final text. Comparing the final text and the logging data shows how often, when, how long, and where the writer paused while writing this section of her text. The log file also shows what revisions the writer made to her text during the writing process and when she made each of them (Barkaoui, 2016b). As Figure 1 illustrates, keystroke logging offers an effective tool for collecting very

Keystroke Logging Data:

[ENTER2]Second,·{8844}if·someone·{2890}choose·a·subject·mea[BS]rely·for·the·future·job·or
·career,·he·may·{7484}suffer[BS6]{9406}not·{4016}[BS3]ve[BS3]ve·[BS3]be·not·{2531}very
·glad·about·his·{2984}life.·{32000}A[BS]T[BS]For·exampke,[BS3]lw[BS]e,·the·well·known
·{4344}subjects·that·is·[BS3]are·{2063}easy·{3546}for·su[BS]tudents·to·find·a·job·is·[BS3]are
·marketing·and·accounting.·However,·{2156}some·students·may·be·[BS3]get·bored·on·the
·accounting·issue{2062}s{2032}·and·{4782}[LEFT][RIGHT5][BS]be·not·used·to[END]get
·{2062}bored·about·the·{3156}nun[BS]mbers·and·forms.·As·a·resl[BS]ult,·they·mo[BS]ay
·[BS9]even·if·they·finally·find·a·good·job·with·high·pa{2140}[BS2]salary,·tehy[BS3]hey·may·not
·be·{4047}satisfied·about·their·life·because·of·the·boring·daily·work·that·they·are·not
·interest.[BS]ed.[ENTER2]

Note: {4344}= pause in milliseconds; BS2= backspace (number of keystrokes); LEFT/RIGHT= Arrow keys (number of characters left/right); UP/DOWN= Arrow key (number of lines up/down); Movement= mouse movement; END= End key; · = space bar; LeftButton= mouse click left; ENTER2= enter + number.

Final Text:

Second, if someone choose a subject merely for the future job or career, he may be not satisfied about his life. For example, the well known subjects that are easy for students to find a job are marketing and accounting. However, some students may be not used to the accounting issues and get bored about the numbers and forms. As a result, even if they finally find a good job with high salary, they may not be satisfied about their life because of the boring daily work that they are not interested.

FIGURE 1. Example of keystroke logging data (from *Inputlog*) for a portion of a writing session and the corresponding final text.

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

detailed data about the writing process that cannot be obtained by other methods (Abdel Latif, 2001).

CHALLENGES IN RESEARCHING WRITERS' PAUSING BEHAVIOR

The preceding discussion highlights the potential contributions of pause analysis to research on writers' cognitive process and the effects of various factors, such as task type, writing mode, L2 proficiency, and typing skills, on those processes. However, pause analysis is not a straightforward process. Defining and interpreting pauses, in particular, can be challenging. First, as Medimorec and Risko (2017) noted, while pauses usually refer to physical inactivity during writing, there is no objectively defined pause threshold in the literature (cf. Chenu et al., 2014; Van Waes & Leijten, 2015; Wengelin, 2006, 2007). Instead, different studies define and measure pauses differently. Chenu et al. (2014) described two approaches to defining a threshold for pauses in writing research: a temporally driven approach and a linguistically driven approach. The temporally driven approach consists in defining a threshold (e.g., 2 seconds) and observing when and where such pauses occur during the writing process. The linguistically driven approach consists in postulating the existence of linguistic units (e.g., word, sentence) and then examining pause frequency and duration between or within these units. In this approach all pauses that the program identifies are counted, regardless of their length.

In a review of studies of writers' pausing behavior that use a temporally driven approach, Medimorec and Risko (2017) found that while the most commonly used pause thresholds in adult writing (both handwriting and typing) are 1 and 2 seconds, some studies have used much lower pause thresholds (e.g., 250ms), while others did not use any thresholds. For example, several studies have used the criterion of 2 seconds (e.g., Alves et al., 2007; Spelman Miller et al., 2008; Wengelin, 2006, 2007). Wengelin (2006) defined a pause as "a transition time between two keystrokes, which is longer than what can be expected to be necessary for the time needed to merely find the next key" (p. 111). She operationalized this definition as a transition between two keystrokes that is longer than 2 seconds, reasoning that this pause criterion is at least twice as long as a "normal transition" even for the slowest typist (cf. Alves et al., 2007). As Chenu et al. (2014) explained, the use of 1 or 2 seconds as a threshold allows the exclusion of those pauses that are associated with motor activities, which tend to be brief and are not relevant to cognitive processes.

The pause criterion chosen can significantly affect the findings of pause studies. For example, the lower the pause criterion the more frequent pauses will appear (Medimorec & Risko, 2017; Spelman Miller et al., 2008; Wengelin, 2006). In addition, inconsistency in pause definition can limit the extent to which the results of different studies can be compared (Medimorec & Risko, 2017). One advantage of the linguistically driven approach to the study of writers' pausing behavior is that it avoids "the thorny question of a threshold" (Chenu et al., 2014, p. 2). Consequently, Chenu et al. (2014) have recommended the use of both approaches in the study of writing dynamics. Another solution is to define pauses in terms of multiple thresholds and examine how variation in pause threshold affects the results as Medimorec and Risko (2017) did.

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

A second problem in pausing research is that it is often difficult to interpret *why* pauses occur (Alves et al., 2007; Spelman Miller, 2000, 2006a, 2006b; Wengelin, 2006). As Wengelin (2006) noted, "[A]lthough we can observe exactly when a pause occurs, it is difficult to infer the cognitive activity it signals" (p. 130). Writers can pause for "physical (e.g., fatigue, motor execution of keyboarding), socio-psychological (e.g., writer's block, daydreaming), and/or cognitive (e.g., cognitive overload) reasons" (Alves et al., 2007, p. 56). Consequently, it is often difficult to determine whether a writer is pausing to plan what to write next, monitor the text produced so far, daydream, rest, or look for a key on the keyboard (Wengelin, 2006). Even if a pause can be interpreted as relating to a particular process, such as planning or revising, it is often difficult to determine the level or levels at which planning or revising may be occurring (Chenu et al., 2014; Spelman Miller, 2006b). Furthermore, as Chenu et al. (2014) cautioned, a particular cognitive process may manifest itself in different pausing patterns and a given pausing pattern may be related to different cognitive processes at different stages of the writing process. Additionally, a pause that occurs during a certain activity is not necessarily related to that activity but could be associated with a previous or a following activity (Chenu et al., 2014).

Three strategies have been suggested in the literature to help address problems concerning the interpretation of writers' pausing data (e.g., Baaijen et al., 2012). First, pauses should be interpreted on the basis of the text generation and revision activities surrounding the pause (Spelman Miller, 2006b). Second, it is important to look at periods between consecutive pauses (i.e., p-bursts) and at variation in writing speed as well (Alamargot et al., 2007). In particular, variation in writing or typing speed is closely associated with variation in writing processing demands (e.g., Alamargot et al., 2007; Hayes & Chenoweth, 2006). As Alamargot et al. (2007) explained, the increase or decrease of writing flow reflects the presence or absence of processes with various cognitive loads operating in parallel with writing (or typing) (p. 27). Consequently, fluency or production rate (e.g., number of characters written per minute) can provide additional clues to fluent and hesitant phases during writing (Alamargot et al., 2007; Hayes & Chenoweth, 2006; Spelman Miller, 2000, 2006a; Spelman Miller et al., 2008). Finally, the use of other data collection tools, such as writer stimulated recalls, think-aloud protocols, and/or eye-tracking, can help clarify what writers are doing when they pause during writing (Leijten & Van Waes, 2013; Wengelin, 2006).

AN EMPIRICAL ILLUSTRATION

To illustrate how pausing data can be collected, analyzed, and interpreted in research on L2 writing, in this section I describe the methods and some results of a study examining the effects of task type, L2 proficiency, and keyboarding skills on the pausing behavior of a group of L2 learners when responding to the TOEFL iBT independent and integrated writing tasks on the computer (Barkaoui, 2014, 2015). Because the purpose of this section is to illustrate how pause data can be analyzed and interpreted to address specific research questions, only those pause features that are typically examined and discussed in the literature on pausing behavior were examined.

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

THE DATA

Keystroke logging data were obtained from 68 students who each responded to two TOEFL iBT writing tasks on the computer. The participants belonged to four groups (i.e., 17 students per group): two English language proficiency (ELP) levels (high and low) by two keyboarding skill levels (high and low). The high ELP groups included post-admissions students in their first or second year of university study, while the low ELP groups included preadmission students who were enrolled in low- to high-intermediate ESL classes.[1] Keyboarding skill level was determined based on the results of two typing tests administered at the beginning of the study. The low keyboarding skill groups included students with average net typing speed (i.e., typing speed adjusted for typing accuracy) of 30 words per minute (WPM) or less, while the high keyboarding skill groups included students with average net typing speed of 40 WPM or more (see Barkaoui, 2015 for more details). About half the participants ($n = 36$) were males. Their ages ranged between 18 and 46 years ($M = 23$, $SD = 4$) and they spoke 21 different first languages (L1), with the majority being L1 speakers of Chinese, Spanish, and Korean.

Each participant responded to a research version of the TOEFL iBT integrated and independent tasks on a local PC. The independent task consisted in writing an argumentative essay of at least 300 words about a general topic in 30 minutes, while the integrated task consisted in listening to a lecture and reading a text about the same topic (5 minutes) and then writing a 300-word summary of both the lecture and the reading in 20 minutes.[2] Participants had access to the reading text, but not the lecture, during the writing interval of the integrated task. With both writing modes, participants were given scratch paper at the beginning of each writing session and informed that they could take notes and/or draft their responses on the scratch paper if they wanted. With both tasks the participants had access only to three editing functions: cut, paste, and undo. All writing sessions were recorded using the keystroke logging program Inputlog 5 (Leijten & Van Waes, 2013).

PAUSE ANALYSIS

Various measures of pause patterns were examined. Following previous studies (e.g., Alves et al., 2007; Spelman Miller et al., 2008; Wengelin, 2006, 2007), a pause was defined as an interruption to keyboard and mouse activity of more than 2 seconds. Inputlog identifies each pause that is longer than the cutoff specified by the program user, in this case 2 seconds, and then automatically generates data concerning the frequency, duration, and linguistic and temporal location of pauses for each writing session. Using output from Inputlog, the following measures of pausing behavior were computed:

1. *Writing and Pausing Time*: Inputlog reports the amount of time spent typing (i.e., writing time) and a *mean pause duration*, that is, the average duration of all pauses longer than 2 seconds, for each writing session. Using Inputlog data, a pause ratio, that is, ratio of time spent pausing to time spent actively writing, was computed for each writing session.
2. *Pause Frequency*: Inputlog reports the total number of pauses that are at least 2 seconds long for each writing session. Using information from Inputlog, the following two indices were computed for each writing session: (a) pauses per keystroke: the total number of pauses divided

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

by the total number of keystrokes (Wengelin, 2006)[3] and (b) pauses per minute: the total number of pauses divided by the length of the writing session (in minutes).

3. *Linguistic Location of Pause*: Inputlog reports the *number* and *mean duration* of pauses at the following linguistic locations: within word, between words, between sentences, between paragraphs, and other. Other pauses (called *unknown pauses* in Inputlog) are pauses that cannot automatically be identified as either within or between words or sentences such as pauses that occur between noncharacter actions (e.g., pauses between two delete actions, pauses between two cursor movements).
4. *Temporal Location of Pause*: To examine the temporal location of pauses, each writing session was divided into three equal intervals. Inputlog then computed the *number* and *mean duration* of pauses for each interval for each writing session.
5. *Pause Type*: Each pause was classified as being either a formulation or a revision pause based on the micro-context in which it occurred. Based on guidelines from Wengelin (2006) and Alves et al. (2007) and a preliminary examination of the keystroke logging data in this study, more than 25 pause patterns (i.e., a pause and the actions immediately preceding and following it) were identified. A concordance program specifically developed for this study was then used to count the frequency of each of the 25 pause patterns in each log file. Twelve of these patterns accounted for 98% of the frequencies of all the instances of the 25 pause patterns in all Inputlog data. These 12 patterns were classified as either formulation pauses ($n = 9$) or revision pauses ($n = 3$) as follows (see Table 1).
   a. *Formulation pauses*: Pauses immediately followed by the keyboarding of additional text at the point of inscription (i.e., immediately after the last-typed character).
   b. *Revision pauses*: Pauses followed by the revision of existing text. These included pauses followed by the deletion of text at the point of inscription or followed by typing or deletion after cursor movement away from the point of inscription (using the mouse and/or the key arrows).
6. *Fluency*: Using data from Inputlog, one index of fluency was examined: number of characters typed per minute (Spelman Miller et al., 2008).

Next, measures of pause patterns were analyzed statistically to (a) describe the characteristics of the participants' pausing patterns and (b) compare these patterns across tasks and groups (defined in terms of L2 proficiency and keyboarding skills). In addition to descriptive statistics, several mixed repeated-measures univariate analysis of variance (ANOVA) were conducted to compare pausing patterns across tasks and groups. Only main and interaction effects that are significant at $p < .05$ are reported in the following text.[4]

RESULTS

The purpose of presenting the following results is to *illustrate* how to analyze and interpret pausing data, but it is important to acknowledge some limitations of the dataset and analyses. Specifically, small samples of participants and tasks were included; participants had to respond to the writing tasks within a specific time frame and in the same order (i.e., integrated task first); only one pause threshold, 2 seconds, was used; and only the duration, location, and frequency of pauses were examined. Participants were not asked to explain why they paused. These limitations need to be taken into account when interpreting the results that follow.

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

TABLE 1. Pause patterns identified in keystroke logging data (adapted from Alves et al., 2007 and Wengelin, 2006)

Pause Context	Pattern	Definition
Formulation Pauses		
Opening a sentence	._^a	The writer paused after finishing a sentence.
After closing a sentence	^._a	The writer paused before starting the first word of a new sentence.
Closing a sentence	a^.	The writer paused after the last word of a sentence and before keyboarding a full stop, perhaps to think about whether to continue the sentence or to start a new one.
After minor delimiter	,^_a	The writer paused after finishing a phrase.
New word after minor delimiter	,_^a	The writer paused before starting the first word of a new phrase.
Before a minor delimiter	a^,	The writer paused after the last word of a phrase and before keyboarding a comma, perhaps to think about whether to continue the phrase or to start a new one.
Closing a word	a^_a	The writer paused after finishing a word within a phrase.
Opening a word	a_^a	The writer paused before starting a new word within a phrase.
Within a word	a^a	The writer paused within a word.
Revision Pauses		
Before deletion	aD	The writer paused before revising their text (by deleting text).
After deletion	D^a	The writer deleted some text, paused, and then continued to write.
Between deletions	DD	The writer deleted some text, paused, and then deleted more text.

Note: The following symbols are used in Table 1 (from Wengelin, 2006): "^" denotes a pause (of 2 seconds or more); "_" denotes space; "a" stands for any letter (lower or upper case); "D" denotes deletion (or backspace); "." denotes a major delimiter (i.e., a full stop, question mark, or exclamation mark); and "," denotes a minor delimiter (i.e., comma, colon, or semicolon). The order in which these symbols are presented indicates the sequence of actions. For example, the pattern ",_^a" indicates that the write first typed a minor delimiter (e.g., a comma), hit the space bar, then paused for at least 2 seconds, before typing the first letter of a word.

WRITING AND PAUSING TIME

On average, the participants completed the independent task in 27 minutes ($SD = 4.6$) and the integrated task in 17.6 minutes ($SD = 4.9$). Pausing time constituted 47% of the total time of the writing session, on average ($SD = 12\%$). However, the ratio of pausing to writing time varied significantly across tasks: $F(1, 64) = 90.34$, $p = .0001$, $\dot{\eta}^2 = .59$. Specifically, the participants paused, on average, 42% of the time when responding to the independent task ($SD = 11\%$) and 53% of the time when responding to the integrated task ($SD = 10\%$). Overall, the ratio of pausing to writing time did not differ significantly across L2 proficiency and keyboarding skill groups at $p < .05$, although, generally, participants with higher L2 proficiency and higher keyboarding skills needed slightly less time to complete the writing tasks.

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

TABLE 2. Descriptive statistics for the number and ratio of pauses per minute and per keystroke

	Number of Pauses		Pauses per 100 Keystrokes		Pauses per Minute	
	M	*SD*	*M*	*SD*	*M*	*SD*
Independent Task	107.01	30.03	5.00	2.00	7.06	2.18
Integrated Task	75.53	22.29	5.00	2.00	7.51	2.36
Low L2 Proficiency	99.18	28.37	6.00	2.00	7.92	2.36
High L2 Proficiency	84.65	31.23	4.00	2.00	6.74	2.07
Low Keyboarding	98.01	31.91	6.00	2.00	7.77	2.41
High Keyboarding	84.12	27.89	4.00	1.00	6.77	2.02
Total	*91.27*	*30.72*	*5.00*	*2.00*	*7.28*	*2.28*

PAUSE FREQUENCY

The participants paused on average 91 times ($SD = 31$), with the number of pauses ranging between 21 and 171 pauses. The participants paused on average seven times per minute ($M = 7.28$) or five pauses per every 100 keystrokes. As Table 2 shows, the participants paused significantly less frequently per minute with the independent task ($M = 7.06$) than they did with the integrated task ($M = 7.51$): $F(1, 64) = 5.73$, $p = .02$, $\dot{\eta}^2 = .08$. The average ratio of pauses per 100 keystrokes was identical for the two tasks (i.e., 5 pauses per 100 keystrokes).

Table 2 shows that the low-proficiency participants paused more frequently ($M = 99.18$) than did the high-proficiency group ($M = 85$ pauses). While the low-proficiency group paused 7.92 times per minute (6 pauses per 100 keystrokes), the high-proficiency group paused 6.74 times per minute (4 pauses per 100 keystrokes). Differences between L2 proficiency groups were statistically significant for number of pauses per minute, $F(1, 64) = 5.03$, $p = .028$, $\dot{\eta}^2 = .07$, and for number of pauses per 100 keystrokes, $F(1, 64) = 15.84$, $p = .0001$, $\dot{\eta}^2 = .20$.

Differences between keyboarding skill groups were also statistically significant both for number of pauses per minute, $F(1, 64) = 3.87$, $p = .05$, $\dot{\eta}^2 = .06$, and for number of pauses per 100 keystrokes, $F(1, 64) = 32.17$, $p = .0001$, $\dot{\eta}^2 = .33$. While the low keyboarding skill group paused 7.77 times per minute (6 pauses per 100 keystrokes), the high keyboarding skill group paused 6.77 times per minute (4 pauses per 100 keystrokes). The roughly similar ratio of pauses per minute and higher ratio of pauses per keystroke for the low keyboarding group suggest that this group typed fewer characters per minute because of their low typing speed. Overall, participants with lower L2 proficiency or lower keyboarding skills paused significantly more frequently than did participants with higher L2 proficiency or higher keyboarding skills.

PAUSE LINGUISTIC LOCATION

Table 3 shows that the participants paused most frequently between words ($M = 40$ pauses), followed by other pauses ($M = 32$), and within-word pauses ($M = 10$). Other

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

TABLE 3. Descriptive statistics for pause frequency, mean pause duration, and ratio of pause location

	Pause Frequency		Ratio of Pause Location		Pause Duration (in seconds)	
Pause Location	*M*	*SD*	*M*	*SD*	*M*	*SD*
Within word	9.82	7.47	.11	.07	4.15	2.84
Between words	40.09	17.12	.44	.10	4.85	1.16
Between sentences	5.12	3.60	.05	.03	5.73	4.84
Between paragraphs	2.43	2.11	.04	.03	5.96	8.34
Other	32.06	14.98	.35	.12	7.21	3.42

pauses are pauses that are preceded and/or followed by deletion or cursor movement. Differences in the frequencies of pauses across linguistic locations were statistically significant: $F(2.44, 156.32) = 342.04$, $p = .0001$, $\dot{\eta}^2 = .84$.[5] Pauses between words accounted for 44% of the pauses, while other pauses accounted for 35% of the pauses, followed by within-word pauses (11%).

In terms of mean pause duration, Table 3 shows that the longest pauses, on average, were other pauses ($M = 7.21$ seconds), followed by pauses between paragraphs ($M = 5.96$ seconds), and pauses between sentences ($M = 5.73$ seconds). Differences in mean pause duration across linguistic locations were significant: $F(1, 64.55) = 149.73$, $p = .0001$, $\dot{\eta}^2 = .70$. The differences in the frequency, mean duration, and ratio of different pause locations relative to each other were consistent across tasks and L2 proficiency and keyboarding skill groups. However, as the large standard deviations in Table 3 show, there was large variability in terms of the number and duration of pauses across individuals. For example, the large standard deviation for the frequency of pauses between words (SD = 17.12) indicates that participants varied substantially in terms of the number of times they paused between words (range: 7 to 105 pauses). The duration of pauses between paragraphs (SD = 8.34) also varied substantially across participants (range: 0 to 68 seconds).

There was a significant main task effect, $F(1, 64) = 17.28$, $p = .0001$, $\dot{\eta}^2 = .21$, and a significant task-by-location interaction effect, $F(1.96, 125.59) = 9.36$, $p = .0001$, $\dot{\eta}^2 = .13$, on the ratios of pause locations. As Table 4 shows, participants tended to have a larger proportion of other pauses with the independent task ($M = 38\%$) than they did with the integrated task ($M = 32\%$). They also tended to pause less frequently within words with the independent task ($M = 9\%$) than they did with the integrated task ($M = 12\%$).

There was a significant main task effect, $F(1, 64) = 45.94$, $p = .0001$, $\dot{\eta}^2 = .42$, and a significant task-by-location interaction effect, $F(1.01, 64.92) = 46.37$, $p = .0001$, $\dot{\eta}^2 = .42$, on mean pause duration too. Table 4 shows that the participants paused for a longer period, on average, between paragraphs with the independent task ($M = 8.05$ seconds) than they did with the integrated task ($M = 3.87$ seconds). This may be because they had to plan what to write next and/or generate ideas for a longer period with the independent task than they did with the integrated task. There was no significant difference between

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

TABLE 4. Descriptive statistics for mean pause duration (in seconds) and pause location ratio by task and group

Task	Independent				Integrated			
	Mean Duration		*Location Ratio*		*Mean Duration*		*Location Ratio*	
Pause Location	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Within word	3.92	2.17	.09	.06	4.37	3.39	.12	.08
Between words	4.57	.94	.43	.10	5.13	1.29	.44	.09
Between sentences	5.46	3.35	.05	.03	6.01	5.99	.06	.04
Between paragraphs	8.05	10.23	.03	.03	3.87	5.15	.03	.03
Other	6.74	2.56	.38	.12	7.69	4.06	.32	.10
L2 Proficiency	*Low*				*High*			
	Mean Duration		*Location Ratio*		*Mean Duration*		*Location Ratio*	
Pause Location	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Within word	3.97	1.86	.11	.07	4.29	3.47	.10	.07
Between words	4.93	1.19	.44	.10	4.79	1.14	.43	.10
Between sentences	6.69	6.40	.06	.04	4.93	2.79	.05	.03
Between paragraphs	5.57	9.78	.02	.02	6.28	6.95	.03	.03
Other	7.36	3.33	.33	.12	7.09	3.51	.36	.11
Keyboarding Skill	*Low*				*High*			
	Mean Duration		*Location Ratio*		*Mean Duration*		*Location Ratio*	
Pause Location	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Within word	3.73	1.18	.13	.07	4.58	3.87	.08	.06
Between words	4.85	1.15	.46	.09	4.85	1.18	.41	.10
Between sentences	5.99	5.60	.05	.03	5.45	3.91	.06	.04
Between paragraphs	4.78	4.37	.03	.02	7.20	11.00	.03	.03
Other	7.21	3.32	.30	.10	7.22	3.54	.40	.11

L2 proficiency groups in terms of the ratio of different pause locations or mean pause duration across linguistic locations. Note also the large standard deviations for the mean duration and ratio of pauses between sentences and between paragraphs in Table 4, which indicate that students varied substantially in terms of how often and how long they paused at these locations.

There were no significant main effects for keyboarding skill group on the ratio or mean duration of pauses at different linguistic locations. However, there was a significant keyboarding skill by pause location interaction effect on pause ratios: $F(1.91, 121.94) = 13.19$, $p = .0001$, $\dot{\eta}^2 = .17$. As Table 4 shows, participants with low keyboarding skills tended to pause more frequently within and between words ($M =$ 13% and 46%, respectively), than did participants with high keyboarding skills ($M =$ 8% and 41%). Within-word pauses may relate to low-level difficulties such as slow typing speed, correcting typos, and/or uncertainty about spelling. In contrast, participants with high keyboarding skills had a higher proportion of other pauses ($M =$ 40%) than did the low keyboarding skill group ($M = 30\%$). There were no significant

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

main or interaction effects of keyboarding skills on mean pause duration, but, overall, participants with high keyboarding skills paused longer between paragraphs (M = 7.20 seconds) than did participants with low keyboarding skills (M = 4.78 seconds). Pauses between paragraphs are likely related to planning paragraph content and structure.

PAUSE TEMPORAL LOCATION

Each writing session for each participant and each task was divided into three equal intervals, and the number and mean duration of pauses in each interval were computed and compared across intervals. Given that writing intervals (for each participant) are of equal length, the number of pauses per interval could be compared across intervals without any transformation. The average length of the intervals was 8.3 minutes. Intervals for the independent task were longer (M = 9 minutes) than those for the integrated task (M = 5.8 minutes). As Table 5 shows, the participants tended to pause

TABLE 5. Descriptive statistics for pause frequency and mean pause duration (in seconds) by writing interval

	Pause Frequency		Mean Duration					
All Participants	*M*	*SD*	*M*	*SD*				
Interval 1	21.58	14.17	52.73	121.69				
Interval 2	35.02	11.76	5.30	1.57				
Interval 3	34.67	11.88	5.91	2.67				
Task	*Independent*				*Integrated*			
	Pause Frequency		*Mean Duration*		*Pause Frequency*		*Mean Duration*	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Interval 1	31.22	11.89	25.36	124.34	11.94	8.64	80.11	113.39
Interval 2	37.62	11.38	5.24	1.37	32.43	11.64	5.36	1.76
Interval 3	38.18	12.39	5.73	3.11	31.16	10.29	6.09	2.14
L2 Proficiency	*Low*				*High*			
	Pause Frequency		*Mean Duration*		*Pause Frequency*		*Mean Duration*	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Interval 1	24.76	14.15	41.87	137.22	18.92	13.72	61.84	107.09
Interval 2	37.58	10.84	5.48	1.86	32.88	12.14	5.16	1.29
Interval 3	36.84	11.36	5.78	2.19	32.85	12.08	6.02	3.02
Keyboarding	*Low*				*High*			
	Pause Frequency		*Mean Duration*		*Pause Frequency*		*Mean Duration*	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Interval 1	24.34	15.61	59.05	149.21	18.65	11.89	46.04	83.87
Interval 2	37.19	12.29	5.30	1.56	32.73	10.79	5.31	1.60
Interval 3	36.49	11.77	5.60	2.02	32.74	11.78	6.24	3.20

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

significantly less frequently in the first interval of the writing session ($M = 21.58$ pauses), than they did in the second ($M = 35.02$) and third ($M = 34.67$) intervals: $F(1.55, 99.09) = 95.99$, $p < .05$, $\dot{\eta}^2 = .60$. However, pauses in the first interval tended to be significantly longer ($M = 52.73$ seconds) than those in the second ($M = 5.30$ seconds) and third ($M = 5.91$ seconds) intervals: $F(1, 64) = 20.04$, $p < .05$, $\dot{\eta}^2 = .24$. In other words, the participants tended to pause less frequently, but for a longer period, at the beginning of the writing session, most likely for reading the task and/or global planning of the content and structure of their responses. More frequent, but much shorter, pauses in the second and third intervals are likely associated with local planning and/or text revision. The frequency and mean duration of pauses did not differ significantly ($p > .05$) between intervals 2 and 3.

Table 5 shows that this pattern of fewer but longer pauses in the first interval and more frequent but shorter pauses in the second and third intervals was consistent across tasks and learner groups. However, there were some differences across tasks and learner groups in terms of the frequency and mean duration of pauses *within* intervals. First, there was a significant task-by-interval interaction effect on pause frequency, $F(1.87, 119.71) = 32.37$, $p = .0001$, $\dot{\eta}^2 = .34$, and mean pause duration, $F(1, 64.06) = 6.23$, $p = .015$, $\dot{\eta}^2 = .09$. Participants paused significantly more frequently in the first interval with the independent task ($M = 31.22$ pauses) than they did with the integrated task ($M = 11.94$). However, pauses in the first interval were much longer on average for the integrated task ($M = 80.11$ seconds) than were those for the independent task ($M = 25.36$ seconds). This is likely because the participants spent time rereading the reading passage at the beginning of the integrated task. Additionally, Table 5 shows that the number of pauses is distributed almost equally across the three intervals with the independent task, unlike the integrated task, where the participants paused three times more frequently in the second and third intervals than they did in the first interval. However, with both tasks, participants tended to pause much longer in the first interval than they did in the second and third intervals.

Second, there was a significant L2 proficiency-by-interval interaction effect on mean pause duration: $F(1, 64) = 5.95$, $p = .017$, $\dot{\eta}^2 = .09$. As Table 5 shows, the low-proficiency group paused more frequently in the first interval ($M = 24.76$ pauses) than did the high-proficiency group ($M = 18.92$ pauses). But the latter group paused longer, on average, in the first interval ($M = 61.84$ seconds) than did the low-proficiency group ($M = 41.87$ seconds), which suggests that the high-proficiency group spent more time planning their responses at the beginning of the writing session than did the low-proficiency group. The numbers and mean duration of pauses in the second and third interval are similar for the two L2 proficiency groups.

Third, as Table 5 shows, the low keyboarding skill group paused significantly more frequently in the first interval ($M = 24.34$ pauses) than did the high keyboarding skill group ($M = 18.65$ pauses). In addition, the low keyboarding skill group paused longer, on average, in the first interval ($M = 59.05$ seconds) than did the high keyboarding skill group ($M = 46.04$ seconds). The number and duration of pauses in the second and third intervals are similar for the two keyboarding skill groups. Finally, correlational analyses indicated that, generally, participants who tended to pause frequently in any interval paused frequently in the other intervals too (correlations of pause frequencies across intervals varied between .39 and .72). However, participants who paused longer in the

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

TABLE 6. Descriptive statistics for pause types by task and group

	Frequency				Percentage			
	Formulation		Revision		Formulation		Revision	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Independent Task	48.46	19.64	14.25	7.22	76.73	10.33	23.27	10.33
Integrated Task	35.32	13.00	8.68	5.67	80.58	9.21	19.42	9.21
Low L2 Proficiency	47.18	17.74	13.08	7.04	78.01	10.32	21.99	10.32
High L2 Proficiency	37.46	16.81	10.11	6.80	79.21	9.64	20.79	9.64
Low Keyboarding	47.71	17.58	11.69	7.22	80.81	9.27	19.19	9.27
High Keyboarding	35.71	16.08	11.23	6.90	76.37	10.17	23.63	10.17
Total	41.89	17.85	11.46	7.05	78.66	9.93	21.34	9.93

first interval had shorter pauses in the second and third intervals, and vice versa (correlations of pause duration across intervals were between –.10 and –.01).

PAUSE TYPE

Pauses were classified based on their context as being either formulation or revision pauses. Table 6 shows that the participants made almost four times as many formulation pauses ($M = 41.89$ pauses) as revision pauses ($M = 11.46$). Of all the pauses the participants made, 79% were formulation pauses. This pattern of more formulation than revision pauses, which was consistent across tasks and learner groups, suggests that the participants were more concerned with generating/formulating new content (what and how to write next) than revising what they have written.

Comparing the percentages of pause types in Table 6, it is clear that there were no large differences across L2 proficiency groups, but there were significant differences across keyboarding skill groups and tasks. First, the participants paused significantly more frequently to revise ($M = 23.27\%$) and significantly less frequently to formulate ($M = 76.73\%$) with the independent task than they did with the integrated task ($M = 19.42\%$ and 80.58%, respectively): $F(1, 64) = 11.39$, $p = .001$, $\dot{\eta}^2 = .15$. The participants were, thus, more likely to pause to revise their texts with the independent task than they were with the integrated task.

Second, participants with high keyboarding skills paused significantly more frequently to revise ($M = 23.63\%$) and significantly less frequently to formulate ($M = 76.37\%$), than did the low keyboarding skill group ($M = 19.19\%$ and 80.81%, respectively): $F(1, 64) = 4.90$, $p = .03$, $\dot{\eta}^2 = .07$. While this effect is weak, it seems that participants with low keyboarding skills were less likely to pause to revise their texts than were those with high keyboarding skills.

FLUENCY

As Table 7 shows, the average number of characters per minute was 161.53 (range 71 to 252 characters per minute, or 14 to 50 words per minute).[6] There was no significant

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

TABLE 7. Descriptive statistics for fluency by task and group

	Characters Typed per Minute	
	M	*SD*
Independent Task	162.89	35.60
Integrated Task	160.17	41.69
Low L2 Proficiency	146.49	36.63
High L2 Proficiency	174.13	35.88
Low Keyboarding	136.99	27.67
High Keyboarding	187.55	30.87
Total	*161.53*	*38.65*

difference in terms of mean number of characters typed per minute across tasks. Table 7 shows that participants with higher L2 proficiency typed significantly more characters per minute, $F(1, 64) = 16.09$, $p < .05$, $\dot{\eta}^2 = .20$, than did the low-proficiency group. This indicates a higher level of fluency for the first group. Keyboarding skill had a significant main effect on the number of characters typed per minute, $F(1, 64) = 66.28$, $p = .0001$, $\dot{\eta}^2 = .51$. Table 7 shows that participants with higher keyboarding skills typed more characters per minute than did the low keyboarding skill group.

DISCUSSION

How do we interpret the results of pause analysis? This section discusses the key results of the preceding analyses and how they compare to previous studies. First, overall, higher order transition points (e.g., between paragraphs) were associated with longer pauses than were lower order transition points (e.g., between words). This finding is consistent with previous L1 and L2 studies (e.g., Schilperoord, 1996; Spelman Miller, 2006b; Wengelin et al., 2009) and indicates that the participants engaged in major planning at higher level boundaries and in minor planning and/or revision at lower order boundaries.

Second, the participants paused most frequently between and within words, which is consistent with previous studies (e.g., Chanquoy et al., 1996; Schilperoord, 1996; Spelman Miller, 2006b; Wengelin et al., 2009), but also may indicate a concern among participants to write fast and pause only to check spelling and typos or to search for and retrieve vocabulary, instead of pausing to plan and revise at more global levels. Time pressure might have affected how often and how long the participants paused at different linguistic locations. Hall (1991) also found that students tended to pause longer and inscribe at a considerably faster rate during a timed writing task than they did during an untimed task.

Third, the participants made fewer, but much longer, pauses in the first interval of the writing session than they did in the second and third intervals. They paused significantly more frequently, but for shorter periods, in intervals 2 and 3. These patterns suggest that they paused to plan the overall content and structure of their responses in the first interval (cf. Wengelin, 2006), and paused to read sections of their texts, plan,

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

retrieve, and/or revise lower-level units (e.g., words, phrases) in later intervals. It is also possible that participants had more to say at the beginning and/or that their motivation declined later. Additionally, those participants who paused for a longer period in interval 1 generally paused for shorter periods in intervals 2 and 3. This suggests that those participants who planned more extensively at the beginning of the writing session did not feel the need to pause as often later and/or felt that they cannot pause as frequently later because of time pressure. Finally, the participants paused more frequently to formulate what to write next than to revise what they had written. These patterns indicate a concern with generating ideas, more than with revising and improving one's text, perhaps because of the requirement to write at least 300 words in 30 minutes.

TASK EFFECTS

First, the participants paused longer, on average, with the integrated task, particularly in the first interval, than they did with the independent task, perhaps because they needed to reread (segments of) the reading text with the integrated task. Second, participants paused for a longer period between paragraphs with the independent task than they did with the integrated task. This suggests that participants needed to think about and plan what to write next and/or to generate ideas for longer periods with the independent task than they did with the integrated task (cf. Schilperoord, 1996). Third, the independent task resulted in a higher proportion of revision pauses than did the integrated task, perhaps because the participants felt that they need to revise and improve their texts when generating their own content (i.e., for the independent task), while when drawing content from the reading and listening (i.e., with the integrated task), they did not feel a need to revise their responses as frequently.

L2 PROFICIENCY EFFECTS

Pausing patterns varied significantly across L2 proficiency groups. First, low-proficiency participants paused more frequently on average than did high-proficiency participants. Second, the high-proficiency participants paused longer at the beginning of the writing session than did low-proficiency participants. These findings suggest that while more proficient participants tended to pause less frequently overall, they tended to pause longer at the beginning of the writing session, perhaps to read the task and/or plan their responses, than did less proficient participants. These findings are consistent with previous studies that show that proficient writers tend to pause longer to plan their texts in more detail and at a global level at the beginning of the writing process, while less proficient writers tend to generate throughout the writing process, adopting a knowledge-telling approach (e.g., Sasaki, 2000; Van der Hoeven, 1999; Xu & Ding, 2014; Xu & Qi, 2017). Finally, participants with high L2 proficiency had higher fluency rates than did those with low proficiency. The finding that participants with lower L2 proficiency had slower production rate suggests that they might have experienced cognitive overload imposed by higher order processes (e.g., planning) that slowed their typing speed (cf. Alamargot et al., 2007; Fayol, 1999).

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

KEYBOARDING SKILL EFFECTS

Keyboarding skill was significantly related to various aspects of the participants' pausing behavior. First, participants with low keyboarding skills paused more frequently, on average, than did those with high keyboarding skills (cf. Alves et al., 2007). Second, participants with low keyboarding skills paused more frequently within and between words, while participants with high keyboarding skills had a higher proportion of other pauses and tended to pause longer between paragraphs than did those with low keyboarding skills. This suggests that participants with low keyboarding skills had to pause more frequently at low-level units (words), while participants with high keyboarding skills engaged in more revisions (i.e., other pauses) and more higher level planning (i.e., pauses before paragraphs) (cf. Alves et al., 2007).

Third, participants with low keyboarding skills paused longer than did those with high keyboarding skills in the first interval. These participants tended to plan and draft their responses on paper before typing them on the computer (Barkaoui, 2015). Fourth, participants with high keyboarding skills had higher fluency rates and a higher proportion of revision pauses than did those with low keyboarding skills. A lower percentage of revision pauses for participants with low keyboarding skills suggest that they did not make many changes to their texts perhaps because (a) they extensively planned and drafted their responses on paper at the beginning of the writing session (as suggested by longer pauses in interval 1) and/or (b) they did not take advantage of the computer potential for revising their texts once they had typed them (cf. Phinney & Khouri, 1993).

To summarize, participants with low L2 proficiency and low keyboarding skills paused more frequently than did participants with high L2 proficiency and high keyboarding skills. These patterns indicate that the writing of slow typists and less proficient writers was characterized by shorter execution periods and a higher number of pauses than that of faster typists and more proficient writers (cf. Alves et al., 2007). It seems that slow typists and less proficient writers suffered "from a trade-off between the execution and formulation systems" (Alves et al., 2007, p. 63). Because of their low abilities, these writers might have devoted pauses to high-level writing processes (e.g., planning, revising) and execution periods to low-level processes (i.e., typing). These findings also highlight the possible overlap between L2 writing ability and typing speed, which likely develop together.

FUTURE DIRECTIONS

As noted previously, the dataset and analyses have their limitations. Despite these limitations, the analyses demonstrate how pause analysis can be conducted and interpreted. The findings highlight the advantages of pause analysis as a tool for examining variability in L2 learners' writing processes. As the preceding findings illustrate, pause analysis can also provide significant insights about the effects of various individual and contextual factors on L2 writing processes. Future studies could use pause analysis to examine the effects of other factors such as writing mode, task complexity, learner L1, and time constraints on L2 learners' writing processes. For instance, examining writers' pausing behavior when responding to timed and untimed

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

tasks could help identify whether and how L2 learners' pause patterns, particularly the frequency and duration of pauses at different linguistic locations, under time pressure compare to those that they manifest in untimed tasks (e.g., when writing a take-home paper for a course) (cf. Hall, 1991). To compare pausing behavior across writing modes, smartpens can be used to examine pausing behavior when writing on paper (Barkaoui, 2016a). Task complexity can be examined using a repeated-measures design similar to the one employed in this study, that is, comparing the pausing behavior of the same learners when responding to tasks that vary in terms of their complexity. Such studies could include L2 learners from different backgrounds (in terms of L1, L2 proficiency, etc.) to examine the interaction effects of task complexity and learner variables on pausing behavior.

Furthermore, future studies could examine changes in L2 learners' pausing behavior over time and/or in relation to L2 writing instruction. There is little longitudinal research on changes in L2 learners' writing processes or the effects of L2 instruction on L2 learners' writing processes. Pause analysis, particularly when combined with other methods such as eye-tracking, observation, interviews, and stimulated recalls (Alamargot et al., 2007; Chukharev-Hudilainen, Saricaoglu, Torrance, and Feng, 2019; Leijten & Van Waes, 2013; Révész, Michel, & Lee, 2019; Wengelin, 2006; Wengelin et al., 2009), can provide significant insights into whether and how L2 writing ability develops over time. Such research can also shed light on the aspects of writing that L2 learners with different backgrounds and L2 proficiency levels attend to, as well as when and why, when responding to different writing tasks under different conditions in various contexts. Such a program of research can contribute significantly to our understanding of L2 writing proficiency and the effects of individual and contextual factors on L2 writing processes and development.

Another area worth exploring concerns the relationships between pausing behavior and text quality. Few studies have examined such relationships in L1 writing (e.g., Medimorec & Risko, 2017; Van den Bergh et al., 1994), and even fewer such studies were conducted in relation to L2 writing (e.g., Spelman Miller et al, 2008; Xu & Ding, 2014; Xu & Qi, 2017). Xu and Ding (2014) and Xu and Qi (2017), for example, found that prewriting pause duration was significantly associated with text quality. Spelman Miller et al. (2008), in contrast, found that text quality scores were strongly associated with fluency, but weakly associated with pausing behavior. These mixed findings led Deane (2014) to conclude that the relationship between keystroke logging process data, including pausing patterns, and writing quality is not clear (cf. Révész et al., 2017). One way to address this question is to examine the relationships between the writing processes that L2 writers employ, the characteristics of the texts they produce, and human ratings of text quality.

However, to further advance research on pausing behavior, we need more theorizing and research on what keystroke logging data and pausing behavior can tell us about L2 learners' cognitive processes when writing in L2. As a result, future studies need to examine what criteria to use to define and identify pauses; which of the pause indices discussed in the literature provide better indicators of L2 writers' cognitive processes; which specific processes each index measures; which indices are sensitive to variation in writing processes across individuals, tasks, and contexts; and which indices can detect changes in L2 learners' writing processes over time. These studies need to combine

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

keystroke logging with other data collection and analysis strategies (e.g., stimulated recall, think aloud protocols). Such a program of research has the potential to further our understanding of the cognitive processes underlying L2 learners' pausing behavior; shed more light on the role of pausing behavior in L2 writing performance and development; and inform L2 writing instruction and assessment by, for example, helping L2 learners' reflect on and improve their L2 writing (cf. Lindgren & Sullivan, 2003) and guiding the design of diagnostic and formative assessments of L2 writing processes (cf. Deane, 2014).

NOTES

[1]The cut score for admission to the university where the study was conducted is 83 or higher on TOEFL iBT and Band 6.5 or higher on IELTS. Postadmission students had scores at or higher than the admission cut scores, while the preadmission students had scores lower than the cut score.

[2]For more details about, and examples of, the writing tasks used in this study, please see the TOEFL iBT website: https://www.ets.org/toefl/ibt/about/content/.

[3]Some studies use pause per word, but as Wengelin (2006) argued, "[P]ause frequency is a process measure rather than a product measure and 'number of words' is a product measure—counted in the final edited texts. The writer could easily—especially in computer keyboarding—have changed his or her text many times and deleted many of the words in which the pauses occurred. Pause/keystroke, therefore, more adequately measures how often pauses occur in the writing process" (p. 117).

[4]For significant effects, effect size was estimated using *partial eta-squared* (*partial* $\dot{\eta}2$). *Partial* $\dot{\eta}2 \geq .01$ indicates a small effect size; *partial* $\dot{\eta}2 \geq .09$ indicates a medium effect; and *partial* $\dot{\eta}2 \geq .25$ indicates a large effect (Field, 2009).

[5]For repeated-measures tests that involve comparing more than two levels (e.g., linguistic location, writing interval), the spherecity assumption should be examined. Mauchly's tests indicated that the assumption of spherecity was violated for several of these tests (e.g., linguistic location). Consequently, degrees of freedom for all these tests were corrected using Greenhouse-Geisser estimates of spherecity (Field, 2009).

[6]The commonly accepted standard word length in English typing tests (e.g., Standards Australia, 2001) is 5 keystrokes = 1 word.

REFERENCES

Abdel Latif, M. (2001). 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.

Alamargot, D., Dansac, C., Chesnet, D., & Fayol, M. (2007). Parallel processing before and after pauses: A combined analysis of graphomotor and eye movements during procedural text production. In M. Torrance, L. van Waes, & D. Galbraith (Eds.), *Writing and cognition: Research and applications* (pp. 13–29). Amsterdam, The Netherlands: Elsevier.

Alves, R. A., Castro, S. L., de Sousa, L., & Stromqvist, S. (2007). Influence of keyboarding skill on pause-execution cycles in written composition. In M. Torrance, L. van Waes, & D. Galbraith (Eds.), *Writing and cognition: Research and applications* (pp. 55–65). Amsterdam, The Netherlands: Elsevier.

Baaijen, V. M., Galbraith, D., & de Glopper, K. (2012). Keystroke analysis: Reflections on procedures and measures. *Written Communication*, *29*, 246–277.

Barkaoui, K. (2011). Think-aloud protocols in research on essay rating: An empirical study of their veridicality and reactivity. *Language Testing*, *28*, 51–75.

Barkaoui, K. (2014). Examining the impact of L2 proficiency and keyboarding skills on scores on TOEFL iBT writing tasks. *Language Testing*, *31*, 241–259.

Barkaoui, K. (2015). *Test-takers' writing activities during TOEFL-iBT writing tasks: A stimulated recall study* (TOEFL iBT Research Report No. 25; ETS Research Report No. RR-15-04). Princeton, NJ: Educational Testing Service.

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

Barkaoui, K. (2016a). Examining the cognitive processes engaged by Aptis Writing Task 4 on paper and on the computer (Research Report AR-A/2016/1). Retrieved from https://www.britishcouncil.org/examining-cognitive-processes-engaged-aptis-writing.

Barkaoui, K. (2016b). What and when second-language learners revise when responding to timed writing tasks on the computer: The roles of task type, second language proficiency, and keyboarding skills. *The Modern Language Review*, *100*, 320–340.

Boomer, D. S. (1965). Hesitation and grammatical encoding. *Language and Speech*, *8*, 148–158.

Bosher, S. (1998). The composing processes of three Southeast Asian writers at the post-secondary level: An exploratory study. *Journal of Second Language Writing*, *7*, 205–241.

Breetvelt, I., Van den Bergh, H., & Rijlaarsdam, G. (1994). Relations between writing processes and text quality: When and how? *Cognition and Instruction*, *12*, 103–123.

Butterworth, B. (1980). Evidence from pauses in speech. In B. Butterworth (Ed.), *Language production: Speech and talk* (pp. 155–176). New York, NY: Academic Press.

Chanquoy, L., Foulin, J. N., & Fayol, M. (1996). Writing in adults: A real time approach. In G. Rijlaarsdam, H. van den Bergh, & M. Couzijn (Eds.), *Theories, models and methodology in writing research* (pp. 36–43). Amsterdam, The Netherlands: Amsterdam University Press.

Chenoweth, N., & Hayes, J. R. (2001). Fluency in writing: Generating text in L1 and L2. *Written Communication*, *18*, 80–98.

Chenu, F., Pellegrino, F., Jisa, H., & Fayol, M. (2014). Interword and intraword pause threshold in the writing of texts by children and adolescents: A methodological approach. *Frontiers in Psychology*, *5*, 182.

Chukharev-Hudilainen, E., Saricaoglu, A., Torrance, M., & Feng, H. (2019). Combined deployable keystroke logging and eyetracking for investigating L2 writing fluency. *Studies in Second Language Acquisition*, *41*, 583–604.

Cumming, A. (1989). Writing expertise and second language proficiency. *Language Learning*, *39*, 81–141.

Deane, P. (2014). *Using writing process and product features to assess writing quality and explore how those features relate to other literacy tasks* (Research Report No. RR-14-03). Princeton, NJ: Educational Testing Service.

Deane, P., & Zhang, M. (2015). *Exploring the feasibility of using writing process features to assess text production skills* (Research Report No. RR-15-02). Princeton, NJ: Educational Testing Service.

Fayol, M. (1999). From online management problems to strategies in written composition. In M. Torrance & G. Jeffery (Eds.), *The cognitive demands of writing: Processing capacity and working memory effects in text production* (pp. 13–23). Amsterdam, The Netherlands: Amsterdam University Press.

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Thousand Oaks, CA: Sage.

Field, J. (2004). *Psycholinguistics: The key concepts*. London, UK: Routledge.

Haas, C. (1989). How the writing medium shapes the writing process: Effects of word processing on planning. *Research in the Teaching of English*, *23*, 181–207.

Hall, E. (1991). Variations in composing behaviors of academic ESL writers in test and non-test situations. *TESL Canada Journal*, *8*, 9–33.

Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications* (pp. 1–27). Mahwah, NJ: Lawrence Erlbaum.

Hayes, J. R. (2006). New directions in writing theory. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 28–40). New York, NY: Guilford Press.

Hayes, J. R., & Chenoweth, N. A. (2006). Is working memory involved in the transcribing and editing of texts? *Written Communication*, *23*, 135–149.

Hayes, J. R., & Nash, J. G. (1996). On the nature of planning in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications* (pp. 29–55). Mahwah, NJ: Lawrence Erlbaum.

Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications* (pp. 57–71). Mahwah, NJ: Lawrence Erlbaum.

Kormos, J. (2012). The role of individual differences in L2 writing. *Journal of Second Language Writing*, *21*, 390–403.

Lee, H. K. (2004). A comparative study of ESL writers' performance in a paper-based and a computer-delivered writing test. *Assessing Writing*, *9*, 4-26.

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

Leijten, M., & Van Waes, L. (2006). Inputlog: New perspectives on the logging of on-line writing processes in a Windows environment. In K. P. H. Sullivan & E. Lindgren (Eds.), *Computer key-stroke logging: Methods and applications* (pp. 73–93). Oxford: Elsevier.

Leijten, M., & Van Waes, L. (2013). Keystroke logging in writing research: Using Inputlog to analyze and visualize writing processes. *Written Communication*, *30*, 358–392.

Leijten, M., Van Waes, L., Schrijver, I., Bernolet, S., & Vangehuchten, L. (2019). Mapping master's students' use of external sources in source-based writing in L1 and L2. *Studies in Second Language Acquisition*, *41*, 555–582.

Lindgren, E., & Sullivan, K. P. (2003). Stimulated recall as a trigger for increasing noticing and language awareness in the L2 writing classroom: A case study of two young female writers. *Language Awareness*, *12*, 172–186.

Lindgren, E., & Sullivan, K. P. H. (2006a). Analysing online revision. In K. P. H. Sullivan & E. Lindgren (Eds.), *Computer key-stroke logging: Methods and applications* (pp. 157–188). Oxford: Elsevier.

Matsuhashi, A. (1981). Pausing and planning: The tempo of written discourse production. *Research in the Teaching of English*, *15*, 113–134.

Matsuhashi, A. (1987). Revising the plan and altering the text. In A. Matsuhashi (Ed.), *Writing in real time: Modeling production processes* (pp. 197–223). Norwood, NJ: Ablex.

McCutchen, D. (1996). A capacity theory of writing: Working memory in composition. *Educational Psychology Review*, *8*, 299–325.

McCutchen, D. (2000). Knowledge, processing, and working memory: Implications for a theory of writing. *Educational Psychologist*, *35*, 13–23.

Medimorec, S., & Risko, E. F. (2017). Pauses in written composition: On the importance of where writers pause. *Reading and Writing*, *30*, 1267–1285.

Olive, T., & Kellogg, R. T. (2002). Concurrent activation of high- and low-level production processes in written composition. *Memory and Cognition*, *30*, 594–600.

Phinney, M., & Khouri, S. (1993). Computers, revision and ESL writers: The role of experience. *Journal of Second Language Writing*, *2*, 257–277.

Révész, A., Kourtali, N. E., & Mazgutova, D. (2017). Effects of task complexity on L2 writing behaviors and linguistic complexity. *Language Learning*, *67*, 208–241.

Révész, A., Michel, M., & Lee, M. (2019). Exploring second language writers' pausing and revision behaviors: A mixed-methods study. *Studies in Seond Language Acquisition*, *41*, 605–631.

Roca de Larios, J., Manchón, R., Murphy, L., & Marín, J. (2008). The foreign language writers' strategic behaviour in the allocation of time to writing processes. *Journal of Second Language Writing*, *17*, 30–47.

Sasaki, M. (2000). Toward an empirical model of EFL writing processes: An exploratory study. *Journal of Second Language Writing*, *9*, 259–291.

Schilperoord, J. (1996). The distribution of pause time in written text production. In G. Rijlaarsdam, H. van den Bergh, & M. Couzijn (Eds.), *Theories, models and methodology in writing research* (pp. 21–35). Amsterdam, The Netherlands: Amsterdam University Press.

Schilperoord, J., & Sanders, T. (1999). How hierarchical text structure affects retrieval processes: Implications of pause and text analysis. In M. Torrance & D. Galbraith (Eds.), *Knowing what to write: Conceptual processes in text production* (pp. 12–30). Amsterdam, The Netherlands: Amsterdam University Press.

Smagorinsky, P. (1994). Think-aloud protocol analysis: Beyond the black box. In P. Smagorinsky (Ed.), *Speaking about writing: Reflections on research methodology* (pp. 3-19). Thousand Oaks, CA: Sage.

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

Spelman Miller, K. (2006a). Pausing, productivity and the processing of topic in online writing. In K. P. H. Sullivan & E. Lindgren (Eds.), *Computer key-stroke logging: Methods and applications* (pp. 131–155). Oxford, UK: Elsevier.

Spelman Miller, K. (2006b). The pausological study of written language production. In K. P. H. Sullivan & E. Lindgren (Eds.), *Computer key-stroke logging: Methods and applications* (pp. 11–30). Oxford, UK: Elsevier.

Spelman Miller, K., & Sullivan, K. P. H. (2006). Keystroke logging: An introduction. In K. P. H. Sullivan & E. Lingden (Eds.), *Computer keystroke logging: Methods and applications* (pp. 1–9). Oxford, UK: Elsevier.

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press

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

Standards Australia. (2001). *Keyboarding speed tests* (3rd ed.). Sydney: Standards Australia.

Stratman, J., & Hamp-Lyons, L. (1994). Reactivity in concurrent think-aloud editing protocols: Issues for research. In: P. Smagorinsky (Ed.), *Speaking about writing: Reflections on research methodology* (pp. 89–112). Thousand Oaks, CA: Sage Publications.

Torrance, M., & Galbraith, D. (2006). The processing demands of writing. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 67–80). New York, NY: Guilford Press.

Van der Hoeven, J. (1999) Differences in writing performance: Generating as an indicator. In M. Torrance & D. Galbraith (Eds.), *Knowing what to write: Conceptual processes in text production* (pp. 65–77). Amsterdam, The Netherlands: Amsterdam University Press.

Van Waes, L., & Leijten, M. (2015). Fluency in writing: A multidimensional perspective on writing fluency applied to L1 and L2. *Computers and Composition*, *38*, 79–95.

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*, 829–853.

Van Weijen, D., Van den Bergh, H., Rijlaarsdam, G., & Sanders, T. (2008). Differences in process and process-product relations in L2 writing. *ITL - International Journal of Applied Linguistics*, *156*, 203–226.

Wengelin, A. (2006). Examining pauses in writing: Theory, methods and empirical data. In K. P. H. Sullivan & E. Lindgren (Eds.), *Computer key-stroke logging: Methods and applications* (pp. 107–130). Oxford, UK: Elsevier.

Wengelin, A. (2007). The word level focus in text production by adults with reading and writing difficulties. In M. Torrance, L. van Waes, & D. Galbraith (Eds.), *Writing and cognition: Research and applications* (pp. 67–82). Amsterdam, The Netherlands: Elsevier.

Wengelin, A., Torrance, M., Holmqvist, K., Simpson, S., Galbraith, D., Johansson, V., & Johansson, R. (2009). Combined eyetracking and keystroke-logging methods for studying cognitive processes in text production. *Behavior Research Methods*, *41*, 337–351.

Xu, C., & Ding, Y. (2014). An exploratory study of pauses in computer-assisted EFL writing. *Language Learning and Technology*, *18*, 80–96.

Xu, C., & Qi, Y. (2017). Analyzing pauses in computer-assisted EFL writing: A computer-keystroke-log perspective. *Journal of Educational Technology and Society*, *20*, 24–34.

Zhang, M., & Deane, P. (2015). *Process features in writing: Internal structure and incremental value over product features* (Research Report No. RR-15-02). Princeton, NJ: Educational Testing Service.

https://doi.org/10.1017/S027226311900010X Published online by Cambridge University Press