Analyzing writing process data

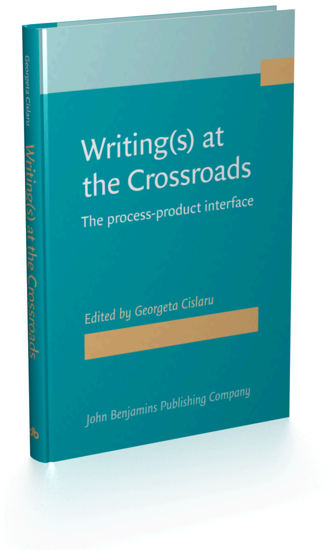
A linguistic perspective

Mariëlle Leijten | University of Antwerp & Research Foundation – Flanders (FWO), University of Antwerp

Luuk Van Waes | University of Antwerp & Research Foundation – Flanders (FWO), University of Antwerp

Eric Van Horenbeeck | University of Antwerp & Research Foundation – Flanders (FWO), University of Antwerp

https://doi.org/10.1075/z.194.14lei



Pages 277–302 of
Writing(s) at the Crossroads: The process–product interface
Edited by Georgeta Cislaru
2015. vi, 304 pp.

John Benjamins Publishing Company

Analyzing writing process data

A linguistic perspective

Mariëlle Leijten[*], Luuk Van Waes[**] & Eric Van Horenbeeck[**]
[*]University of Antwerp & Research Foundation – Flanders (FWO)
[**]University of Antwerp

In this paper we briefly introduce keystroke logging as a research method in writing research, focusing more explicitly on the recently developed linguistic analysis technique. In a case study of two elderly people (healthy versus demented), we illustrate some aspects of this linguistic approach. This analysis aggregates event-based data from the character level to the word level, while taking into account all the revisions that occurred during the composing process. The linguistic process analysis complements the logged process information with results from a part-of-speech tagger, a lemmatizer, a chunker, a syllabifier, and also adds word frequencies. The enriched word level information – together with action time and pause time at the word level – opens up new perspectives in the analysis of process dynamics, once more establishing a closer link between process and product analysis. We thus test the complementary diagnostic accuracy for Alzheimer's disease, mainly focusing on cognitive and linguistic aspects that characterize the process of written language production.

Keywords: keystroke logging; linguistic process analysis; Alzheimer's disease; Inputlog

1. Introduction

Writing research has a rich tradition of research on writing products as well as on writing processes (cf. recent overviews in Berninger 2012; MacArthur, Graham & Fitzgerald 2008; Bazerman 2008; Bazerman et al. 2010). Since the 1980's, there has been a growing focus on process research due to the increase in interest in cognitive psychology. The introduction of the writing model (Flower & Hayes 1981) by Hayes and Flower opened up new areas of research. The first model was based on protocol analysis and mainly provided insights into the internal mechanisms necessary for writing (e.g. memory, planning,

DOI 10.1075/z.194.14lei

problem solving) (Hayes 2012b, a). In the 1996 model the writing medium was the subject of renewed attention (Hayes 1996), mainly due to the fact that the computer gradually became the standard for text production. However, not only have the methods of text production changed considerably, the technical possibilities for studying writing have also evolved. For instance, keystroke logging and eye tracking have been implemented as observation and research tools enabling us to gain a better understanding of the cognitive processes involved in writing. Although there is an increasing interest in and focus on real-time processes, we think it remains very important trying to establish a link between a writer's observed mental processes and the textuality of writing from a product perspective. In this section, we focus on a writing process study in which we use keystroke logging data to specifically examine the crossroads at which the linguistic characteristics of the written product and the writing process itself meet.

2. Keystroke logging

Keystroke logging is a widely used and popular method in writing research. One reason is undoubtedly the fact that it is an unobtrusive method for studying underlying cognitive processes and scarcely interferes with the natural writing process (Sullivan & Lindgren 2006; Leijten & Van Waes 2012; Van Waes et al. 2012). In addition, it is also possible to combine it with complementary observation techniques, like thinking aloud or eye-tracking. Moreover, keystroke logging enables researchers to collect fine-grained pause and revision data and may therefore make it possible to analyze writing processes from a wide range of perspectives. Keystroke logging has been widely used in cognitive writing process research in the broadest sense, for instance in domains like writing development, second language learning, developmental language disorders such as dyslexia, translation, professional writing, on-line writing, etc. An increasing number of studies now report keystroke logging research experiments (e.g. Gunawardhane et al. 2013; Van Waes, Leijten & Remael 2013; Baaijen, Galbraith & de Glopper 2014; Robert & Van Waes 2014; Wininger 2014; Doherty & O'Brien 2014) or describe specific aspects of the research method itself (Ehrensberger-Dow & Perrin 2009; Jakobsen 2011; Baaijen, Galbraith & de Glopper 2012; Galbraith and Baaijen this volume). In addition, there are a number of recent articles focusing on theory development (Leblay & Caporossi 2014; Caporossi & Leblay 2011; Leblay & Caporossi this volume; Macgilchrist & Van Hout 2011; Miller, Lindgren & Sullivan 2008; Van Waes & Leijten 2013; Risku, Windhager & Apfelthaler 2013; Leijten et al. 2014).

In Europe, three free keystroke logging programs are available, each focusing on specific niches: ScriptLog, Translog, and Inputlog.

2.1 Experimental research into writing processes: ScriptLog (www.scriptlog.net)

ScriptLog (Wengelin et al. 2009) was developed by researchers at the universities of Gothenburg, Lund (Sweden) and Stavanger (Norway) for the study of writing processes. It was originally a Macintosh program, then a Windows program, and at the time of writing, a new platform-independent (Windows, MacOS, Linux) version is being tested (ScriptLog 2013: Johansson et al. 2014).

ScriptLog creates a writing environment with a build-in text editor and makes it possible to incorporate frames for different types of elicitation material, such as pictures, texts, movie clips or sounds (for example for dictation experiments). The new version includes extra experimental facilities that enable researchers to set up different writing experiments, for example using dual/triple-task paradigms. The set-up of the environment is controlled in a design module. When activated, ScriptLog keeps a record of all keyboard events, the exact screen position corresponding to these events, and their temporal distribution.

Like other keystroke logging programs, ScriptLog allows the researcher to play back a recorded session – or a selected extract from it – in real time on the basis of the log file. In addition, the analysis module enables the researcher to analyze time distributions across the writing process both for predefined patterns and for user-defined patterns, for example for a particular word string or for a regular expression. Finally, ScriptLog allows researchers with access to an eye tracker to enhance the study of the interplay between writing, monitoring (reading) and revision by integrating eye tracking data. (Currently only SMI eye trackers, more models will be added.) Data on the distribution of visual attention during writing help, for instance, to determine the extent to which pauses are used for monitoring. Data gathered via ScriptLog can now be converted to the Inputlog XML format, thus enabling researchers to conduct Inputlog analyses on ScriptLog data.

2.2 Writing research in translation settings: Translog (www.translog.dk)

Translog was developed at the Copenhagen Business School (Denmark) with the primary aim of studying the writing processes of translators translating a source text from one language into another (Jakobsen 2006). Since 2012, a newly programmed version has been available under the name Translog-II (Carl 2012). It has two interdependent components: a *Translog-Supervisor* component to prepare a logging project and replay it afterwards, and a *Translog-User* component. The latter displays a predefined source text in the top half of the window together with a translation frame in which the target text can be edited. Translog can be used in combination with various eye tracking programs (Tobii 1750 and EyeLink 1000). Since translation research has its own specificities, Translog-II focuses mainly on providing visualization options to show the relation between the source and the target text.

2.3 Writing research in educational and professional settings: Inputlog (www.inputlog.net)

Inputlog was developed at the University of Antwerp (Belgium) to log writing processes in both ecological and experimental settings (Leijten et al. 2014; Leijten & Van Waes 2013). The program logs all keyboard and mouse events in every Windows environment. In the case of texts written in MS Word, extra characteristics relating to the input events are logged to permit fine-grained writing analyses (see below). The program also logs text production with speech recognition systems (Dragon Naturally Speaking, Nuance) and tracks copy-and-paste actions that relate to the use of external digital sources (e.g. the internet).

Inputlog 6.0 features five modules:

1. *Record*: This module logs (keyboard, mouse, and speech) data in Microsoft Word and other Windows-based programs and assigns a unique time stamp (ms) to the data.
2. *Pre-process*: As it is often necessary to prepare and clean up logged data prior to analysis, this module makes it possible to process data from various perspectives: event-based (keyboard, mouse, and speech), time-based or based on changes between Windows (sources: MS Word, Internet etc.). The filter provides an easy way to delete 'noise' at logging session start-up or shut-down. For example, if additional questions are asked at the beginning of the period of observation when the logging session has already started, this pause time (noise) can be excluded from the data analysis.
3. *Analyze*: This module is the heart of the program. It features three process representations (the general and linear logging file and the S-notation of the text) and four aggregated levels of analysis (summary, pause, revision, and source analyses). Additionally, a process graph can be produced. The current version also offers a linguistic process analysis which returns the results from

a part-of-speech tagger, a lemmatizer and a chunker, as well as the syllable boundaries and word frequencies (cf. below).

4. *Post-process*: This module integrates single or multiple log files from Inputlog or other observation tools (Dragon Naturally Speaking and eye tracking data). It is also possible to merge multiple output files for further analysis in, for instance, SPSS, R or MLW in.
5. *Play*: This module allows researchers to play back the recorded session at various levels (time or revision-based). The playback is data-based (not video-based) and the playback speed is adjustable. A logged session can also be reconstructed revision by revision.

The described keystroke logging programs are distributed for free for non-commercial use to researchers and teachers (for a general overview of keystroke logging tools and their characteristics, please see www.writingpro.eu).

3. Linguistic analysis with Inputlog 6.0

Recently Inputlog has implemented a linguistic analysis in which linguistic information is added to the process data (Leijten et al. 2012). As described, the data output from keystroke loggers is mainly based on capturing each character and mouse movement as an isolated event. However, we are now able to aggregate the logged process data from the letter level (keystroke) to the word level. This has allowed us to merge the word-level output with existing lexical and Natural Language Processing (NLP) tools. The logged process data is annotated with different kinds of linguistic information: part-of-speech tags, lemmata, chunking, named entity information, syllabification and frequency information. We will describe these in more detail below. We start with a short overview of the programming challenges we have had to deal with and provide a description of the linguistic analysis flow: what steps are needed to transform character level process analyses into word-level process analyses?

3.1 Aggregating log data from character level to word and sentence level

A number of challenges have to be addressed before the log data of Inputlog can be aggregated to the word level (or higher):

1. First, the concept of a ‘word’ or a ‘sentence’ does not exist in the log file; these items have to be reconstructed because the atomic unit is a key press, a mouse movement, a button click.

2. Second, text production is characterized by revisions to the previously produced text and these may occur either immediately or be initiated at some distance from the word to be revised.
3. Third, additions and deletions can be nested, occurring not in situ but decided on away from the target.

To cope with the non-linearity of writing processes, it is necessary to map the shifting and changing events to the locations where the effects are generated. This can be done using S-notation. S-notation (Severinson Eklundh & Kollberg 2002; Van Horenbeeck et al. 2012) contains information about the types of revision (insertion or deletion), the order of these revisions, and the breaks in the text where the writing process was interrupted.

Consider the following French sentence at the end of a writing process:

(1) "*Des questions sur la science, sur la science et sur l'évolution. Fin.*"

Figure 1 shows the test sentence (1) that we are studying together with all the changes rendered in the S-notation.

Des·questions·sur·la·s[x]^{1}|$_{1}$cience,·sur·la·science·et·sur{·l'évolution}^{4}·[le·progrès]^{2}|$_{3}${.}^{3}|$_{4}$.·Fin.|$_{2}$·

Figure 1. Example of S-Notation

Square brackets indicate a deletion, curly braces an insertion and the vertical pipe symbol, called a 'break', is used to mark the position at which the process was interrupted. The subscript numbers next to the pipe symbol have a corresponding superscript number at either an insertion or at a deletion. In this example: the word 'l'évolution' is surrounded by curly braces indicating that it has been inserted. The insertion is indicated by superscript number 4. This means that it was the 4th revision out of a total of 4 interventions. The vertical pipe symbol with subscript 4 appears before the last word of the sentence and marks the position where the author decided to insert 'l'évolution' instead of 'le progrès', a word that has been deleted as indicated by the square brackets surrounding it.[1]

1. The French sentence is a translation of an English example taken from the Inputlog manual (Leijten & Van Waes 2014).

3.2 Finding words in a stream of events

Inputlog includes a regular expressions-based tokenizer that divides sequences of characters into words and sentences while aggregating the pause time from the individual events. The program uses two data sets: the event log (a file with the 'idfx' extension) used to identify the word boundaries and then, subsequently, a computed list of revisions and edits calculating the pause and the action time once the words have been reconstructed. The main processing method creates a token object that contains the word in its final appearance, its revisions, its position in the final text string, all pause time information (before, within and after words), and the type of processing used for the word (normal production, revised word, deleted word).

For every revision, the process first defines whether it is 'immediate' or 'delayed' (Leijten, De Maeyer & Van Waes 2011). An immediate revision is defined as a revision that takes place before a new word is produced; a delayed revision, in contrast, is characterized by the fact that one or more words are produced between the decision to change the text (the 'break') and the actual deletion or insertion. Initially, every token is assigned the 'normal production' marker which is used for linear text production. However, if the software detects at a later stage that the word contains revisions, the process type is changed to 'revised'. Line feeds, tabs, and other non-characters are replaced with a placeholder to make them visible.

Insertions or deletions may extend over many events, e.g. [le progrès]from the example is a deletion of nine characters and one space. Because the characters are presented as the separate words 'le' and 'progrès', opening or closing symbols inserted by the S-notation are missing. In this case the first word has the opening bracket: ['le' and the second word the closing bracket: 'progrès']. The program adds the missing symbol and returns [le] and [progrès]. Another issue is that the linguistic analysis expects all punctuation symbols to be detached from the front or the back of a word and to be saved as separate tokens. Hence, the word 'Fin.' with the end-of-sentence dot at the end of the example sentence is rendered as 'Fin .' with a space between the word and the endpoint.

Finally, each word is accompanied by timing information such as start time and end time, the word action time and pause time (Figure 2). When the final word is the result of one or more revisions (deletions, insertions) then the pause and action times of the editing events are added to the word production time. Revisions are presented at the word level in combination with the before-word and after-word pauses and the word production time.

Revisions	S-notation	#Chars	Token	Part of Speech	Chunk	Lemma	Syllable	BfrWord-2	BfrWord-1	Word Prod	Within WordPause	AftWord+1
	Des	3	Des	DT	B-NP	de	de	0	0	608	530	141
	questions·	9	questions	NNS	I-NP	question	question	141	312	1731	1653	172
	sur·	3	sur	IN	B-PP	sur	sur	172	187	421	343	156
	la·	2	la	DT	B-NP	la	la	156	250	250	172	234
1-I	s[x]cience··	9	science	NNP	I-NP	science	science	234	218	5304	5242	187

Figure 2. Example of the timing information on word level for a part of the example French sentence considered here

3.3 Flow of the linguistic analysis

S-notation makes a vital contribution by moving the logging process from character to word level. Once words and sentences are recognized, all the tools developed for the NLP framework become available to conduct linguistic analyses on the process data, including word-level revisions and deleted fragments. To support the linguistic analysis of the writing process, a client-server version of Inputlog has been developed (Leijten et al. 2012). Although the logging process and most analyses can be performed using a desktop version of Inputlog, the linguistic modules need to interact with specialized programs, sometimes backed up by very large data files. By decoupling data capture from the analytics, it is possible to add in any NLP module that receives Inputlog data via a communication layer. A workflow procedure presents the data in sequence to the different NLP packages and collects the final output. Because all data traffic is performed using simple text files, cooperation between different software packages is conceivable. The extension has been developed for English and Dutch in order to provide a proof of concept, but it is largely language-independent.

The LT3 Linguistic Preprocessing toolkit developed by the Language and Translation Technology Team of the University of Ghent (LT3serv.ugent.be) is currently used as the main toolkit (Figure 3).[2] The different linguistic tools are installed on an LT3 web server. The output is a tab-delimited string in UTF-8 containing the following fields: token, part-of-speech, probability of the part-of-speech tag, lemma, probability of the lemma, chunk information, named-entity, probability of the named entity, absolute frequency, relative frequency, and

2. These toolkits are also available for French and German.

syllabification (The manual belonging to Inputlog 6.0 has more details on the different components and the tags used for the part of speech tags and the chunks: Leijten & Van Waes 2014).

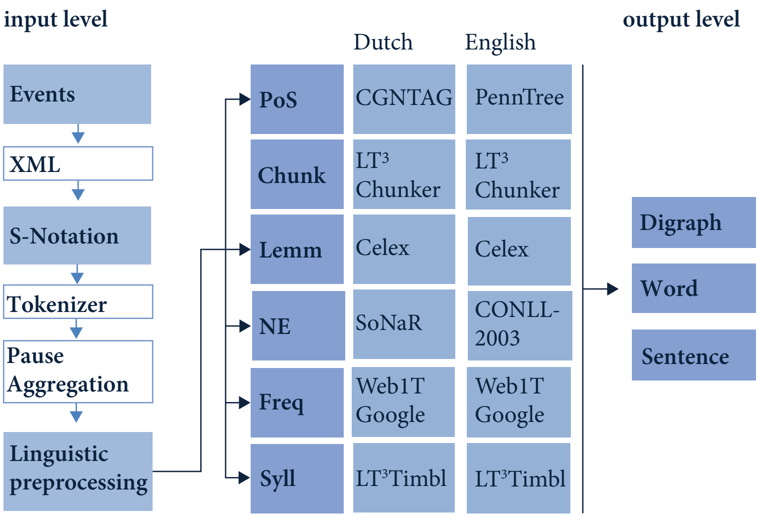


Figure 3. Schematic representation of the flow used in the linguistic analysis performed by Inputlog 6

3.4 Part-of-speech (PoS) tagger

Part-of-speech tagging (PoS tagging), also called grammatical tagging or syntactic word-category disambiguation, is a piece of software that reads text and assigns an appropriate part of speech to a word such as noun, verb, adjective, etc. based on both its definition and its relationship with adjacent words (e.g. Part of Speech column in Figure 2). Because many words have more than one syntactic category, the tagger tries to determine which of the syntactic categories is the most likely for a particular use of a word in a sentence (Manning & Schütze 1999). The English PoS tagger uses the Penn Treebank tag set which contains 45 distinct tags. The Dutch part-of-speech tagger uses the CGN tag set codes which are characterized by a high level of granularity (Van Eynde, Zavrel & Daelemans 2000).

3.5 Chunker

Text chunking combines syntactically related consecutive words into non-overlapping, non-recursive chunks on the basis of a fairly superficial analysis. The LT3 chunkers are rule-based and contain a small set of constituency and distituency rules. Constituency rules define the part-of-speech tag sequences that can occur within a constituent (such as preposition + noun) while distituency rules define the part-of-speech tag sequences that cannot be adjacent within a constituent (such as noun + preposition). The chunks are represented by means of IOB tags (See Figure 2). In the IOB-tagging scheme, each token belongs to one of the following three types: I (inside), O (outside) and B (begin); the B and I tags are followed by the chunk type, e.g. B-VP, I-VP (Jurafsky & Martin 2009).

3.6 Lemmatizer

The base form (lemma) for each orthographic token is generated during lemmatization. For verbs, the base form is the infinitive. For most other words, the base is the stem, i.e. the word form without inflectional affixes. The lemmatizers make use of the predicted PoS codes to disambiguate ambiguous word forms. For instance 'Paris' can be a city or a person. It is classified as a city, for instance, when it is preceded by a preposition of place (bought in) and not by a preposition of possession (bought from). The lemmatizers were trained on the English and Dutch parts of the Celex lexical database, respectively.

3.7 Named entity recognizer

Named entity recognition (NER) sequences of words in a text that belong to predefined NER categories such as names of persons, products, or locations are identified in a text. The Dutch NER system is trained on the 1-million-word subset of SoNaR and identifies the following six NER categories: person, organization, location, product, event, and miscellaneous. The English NER system is trained on the CONLL-2003 shared task data and discerns four NER categories: person, organization, location, and miscellaneous.

3.8 Frequency

Word-frequency information for English and Dutch is retrieved from frequency lists derived from the Web1T Google corpus which is available from LDC.[3] The frequency lists contain the 2 million most frequent words in Dutch and English. The word frequencies are presented both as absolute frequencies and relative frequencies (expressed as percentages).

3.9 Syllabification

Syllabification was approached as a classification task: a large instance base of syllabified data was presented to a classification algorithm which automatically

3. Details about the data set can be found at www.ldc.upenn.edu/Catalog/docs/LDC2006T13/readme.txt.

learned the patterns needed to syllabify unseen data. The syllabification tools were trained on Celex using Timbl as classification algorithm.

We will illustrate the concept of linguistic analysis on the basis of a case study taken from a writing research project investigating the cognitive characteristics of people with Alzheimer's disease.

4. Cognitive writing process characteristics in Alzheimer's disease

A large variety of neuropsychological tests are available for the diagnosis of Alzheimer's disease (AD). In some of these tests, linguistic processing – both oral and written – is an important factor. Language disturbances might serve as a strong indicator of an underlying neurodegenerative disorder like AD. However, the current diagnostic instruments for language assessment mainly focus on product measures, ignoring the importance of the process that leads to written or spoken language production. A more process-oriented approach should allow researchers to describe and analyze writing data from a temporal perspective, focusing on motor, cognitive, and linguistic aspects. Keystroke logging data (writing), potentially complemented by eye tracking data (reading while writing), provides an excellent basis for an adequate description of these processes.

To our knowledge, this is the very first project to test whether cognitive and linguistic aspects that characterize the process of written language production could provide *a complementary and accurate diagnosis* of AD.

4.1 Participants

Three groups of participants were involved in the study:

1. Patients with mild dementia due to AD (n = 5),
2. Patients with mild cognitive impairment (MCI) due to AD (n = 8),
3. A group of age-matched cognitively healthy elderly persons (n = 20).

The patients were recruited from the Memory Clinic of the Antwerp, Middelheim and Hoge Beuken Hospital Network (ZNA), Belgium. All the patients were diagnosed by Prof. Dr. Engelborghs and underwent an extensive neuropsychological examination (Van der Mussele et al. 2012).[4]

4. The AD patients met the NINCDS-ADRDA criteria of probable AD (McKhann et al. 1984). Their mini-mental state examination (MMSE) score (Folstein et al. 1975) was above 20 (mild dementia). MCI patients met the criteria of Petersen et al. (2004), and also complied with the new diagnostic criteria of 'MCI due to AD' (Sperling et al. 2011).

4.2 Task

The three groups of participants were instructed to write two short descriptive texts on a computer. We opted to use two figurative elicitation tasks (see Figure 4a and b) which are part of standardized aphasia test batteries (Goodglass, Kaplan & Barresi 1983; Mesulam et al. 2003; Visch-Brink et al. 2014). On the basis of this picture, the participants produced a brief text in which they described the scene presented to them. To evaluate consistency of task execution, we used two comparable scene pictures, while picture elicitation was counterbalanced to avoid order effects.

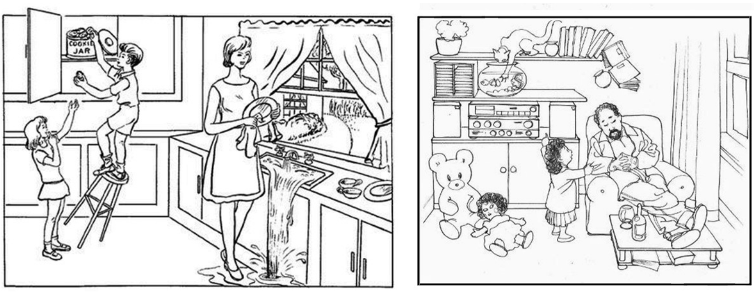


Figure 4. (a-left) 'Kitchen' task by Goodglass and Kaplan (1983); (b) 'Living room' by Visch-Brink et al. (2014)

4.3 Case study

In this study we describe the cognitive processes that characterize the text production of the participants in a controlled task environment. We will first describe the results of some more general process analyses using standard measures employed in keystroke logging research. These include, for example, time taken to perform the task, active writing time, number of pauses and mean length of pauses at different levels, as well as the product/process ratio (i.e. proportion between product and process measures. We will then introduce certain linguistic and product measures in order to control more precisely for word and phrase characteristics that might influence pausing behavior. Since this kind of automated linguistic data analysis has not yet been fully tested for difficulties in aggregating and filtering, we will present a case study in which we selected two participants. We selected one healthy elderly woman (Elise*, 81 years old) and one woman with dementia (Mary*,[5] 79 years old). The participants were matched on three levels: age, education, and career. Both women were about 80 years old, had attended school until

5. *The names of the participants were changed for privacy reasons.

they were 19 years old and they had worked in jobs requiring them to type texts. Readers should note that the main aim of this paper is not to identify differences between the two participants. Instead, the main reason for presenting this case study is because we want to explore the potential value of adding a linguistic perspective to writing process research, and pause analyses in particular, and investigate whether the two approaches can complement one another.

In the same way as in spoken language, we expected that cognitively impaired elderly persons would take longer to produce a (shorter) picture description. Consequently, we expected the proportion of active writing time relative to pausing time to decrease between the healthy elderly and the cognitively impaired elderly (Schilperoord 1996; Van Waes & Schellens 2003). Table 1 gives an overview of some process indicators characterizing the writing processes of the two participants.

Table 1. Mean product, process, and pause characteristics of both picture-depicting tasks

	Elise (healthy)	**Mary (demented)**
Product information		
Number of words in final text	56	41
Number of words in final text (per minute)	11.76	6.73
Process information (pause threshold: 2000 ms)		
Process time	0:04:46	0:05:58
Total pause time	0:02:03	0:03:45
Percentage active writing time (%)	56.65	36.73
Mean number of pauses	24.50	38.50
Mean pause duration (in seconds)	5.08	5.86
Median pause duration (in seconds)	3.23	3.82
Number of characters produced (incl. spaces)	328.5	236
Number of characters produced per minute (incl. spaces)	68.86	38.77
Product/process ratio	0.95	0.99
Mean words produced per sentence	24.17	42.00
Mean word length per sentence	4.59	4.68

The results indicate that Mary (demented – d) took about a minute longer to write the descriptive texts and that her final texts were on average 15 words shorter than Elise's (healthy – h). Thus, compared to Elise, she produced about half the number of words per minute (Elise: 11.76 vs Mary: 6.73). This was due mainly to the amount of pausing time: if we consider the pause analysis based on a threshold of 2 seconds, then Elise(h) paused 25 times on average in both writing tasks, while

Mary(d) paused about 39 times. Consequently, Elise(h) exhibited 20% more active writing time than Mary(d). The average length of their pauses was about 5–6 seconds. The fact that the product/process ratio was close to 1 shows that both writers performed almost no revision. The data also show that the number of words produced per sentence is in itself not a very reliable measure. The number of words produced was about the same as the total text length, indicating that Mary(d) did not use sentence markers. Therefore, pauses within and between words will be a more reliable metric.

In addition to the general pausing behavior, we expected that the mean pause length *within words* and *between words* would help us to further discriminate the healthy elderly from the cognitively impaired elderly (Wengelin 2006; Kellogg 2008; Lindgren et al. 2011). Table 2 (top – Threshold of 2 seconds) shows that Mary(d) made almost twice as many pauses within words as Elise(h) and that the pauses were on average 3 seconds longer. If we aggregate the pauses between words (pause after a word + pause before a word; Leijten & Van Waes 2014) then Elise(h) paused about 43 times and Mary(d) about 29 times at the between-word level. Individual pauses might be below the chosen threshold, but taken together they might exceed the threshold and become relevant (See Figure 5: AW: after words; BW: before words; ww: within words).

However, if we focus only on pauses before words, then Mary(d) made twice as many individual pauses of longer than 2 seconds than Elise(h). The length of individual pauses was about 4 seconds.

translation		the			cakebox									
inputlog events	–	d	e	–	k	o	e	k	e	n	d	o	o	s
pause time (in ms)	***374***	***297***	312	***343***	***1810***	218	343	203	203	249	500	312	296	8877
pause location	***AW***	***BW***	ww	***AW***	***BW***	ww	ww	ww	ww	ww	ww	ww	ww	ww
summed pauses	*sum (671)*			*sum (2153)*										

Figure 5. Example of aggregated between-word pauses for Elise(h) in boxes (AW = after-word pause; BW = before – word pause)

Although in writing research a pause threshold of 2 seconds is quite common, we should be aware of the fact that this causes a lot of data loss, especially at the between and within-word level. When focusing on higher level processes, this is not problematic. However, in the current study we are also interested in more low level processes. Therefore, we lowered the threshold to 200 ms and recalculated the same parameters. This added about 20% more data for the healthy elderly and 25% for the cognitively impaired elderly participant. Following this manipulation,

Table 2. Description of mean pause characteristics in both picture-depicting tasks

	Elise (healthy)	Mary (demented)
Pause information (threshold 2s)	0:02:03	0:03:45
Number of pauses within words	8.00	14.50
Mean pause duration within words (s)	4.67	7.66
Number of pauses between words	43.00	29.00
Mean pause duration between words (s)	6.08	4.97
Number of pauses before words	8.50	20.00
Mean pause duration before words (s)	3.98	4.61
Number of pauses after words	5.50	3.00
Mean pause duration after words (s)	3.10	3.96
Pause information (threshold 0.2s)	0:04:40	0:05:57
Mean number of pauses	349	246
Mean pause duration (in seconds)	0.80	1.59
Median pause duration (s)	0.38	0.66
Number of pauses within words	209.5	153
Mean pause duration within words (s)	0.63	1.47
Median pause duration within words (s)	0.37	0.61
Number of pauses between words	53.5	39
Mean pause duration between words (s)	1.74	2.99
Number of pauses before words	66	52.5
Mean pause duration before words (s)	0.99	2.45

Elise(h) had about 100 more pauses than Mary(d), but her pauses within words were of a mean duration of 600 ms while the pauses made by Mary(d) lasted about twice as long (1470 ms). About 15% of these pauses above the threshold of 200 ms were between words. Again the mean pause duration for Mary(d) was more than 1 second longer.

The above-mentioned measures are common in writing process research (mean pause length within and between words, burst length, process/product ratios). However, using the data from the linguistic analyses we can further refine the concept of 'pause location', especially at the between-word level. The general pause data revealed a difference in the way the two participants dealt with pauses before and after words. We expect that focusing on the pause behavior associated with specific word categories will reveal useful additional features enabling us to further differentiate our observations relating to pre- and post-word pauses. The related literature tells us, for instance, that the elderly in general find it more difficult to choose the correct verb than the correct noun (Yi, Moore, & Grossman 2007).

In the linguistic analysis, pauses are represented in three different ways: BeforeWordPause2 (i.e. the pause immediately following the previous word: technical term 'after word pause'), BeforeWordPause1 (i.e. the pause immediately preceding the word), and AfterWordPause (i.e. the pause immediately after the last character of the word). The 'between word pauses' are therefore calculated as the sum of the BeforeWordPause2 and BeforeWordPause. To a certain extent, this resembles the definition of between-word pauses in handwriting, which are defined as the time it takes to lift the pen when ending a word and starting a new one.

translation	I			see				a			
inputlog events	i	k	–	z	i	e	–	e	e	n	–
pause time (in ms)	0	358	1124	6364	546	1061	1155	1310	312	437	1341
pause location	*BW*	ww	*AW*	*BW*	ww	ww	*AW*	*BW*	ww	ww	AW
summed pauses			*sum (7488)*				*sum (2465)*				

Figure 6. Example of general analysis including pause information (Mary(d)) in boxes (AW = after-word pause; BW = before-word pause)

In Figure 6 we see a coded fragment in which Mary paused for 7488 ms between the words '*ik*' [I] and '*zie*' [See]. Mary paused for 1124 ms after the production of '*ik*', and she paused for 6364 ms before she started to produce the word '*zie*'. One of the main challenges in keystroke logging research is to better understand these kinds of pause patterns. Are both pauses related to the production of the next word? Is there a gradual shift of attention, e.g. from the evaluation of the previous word to the planning (and translation) of the next word (Maggio et al. 2012)? The distribution of between-word pauses is also partly determined by personal preference. Some people systematically seem to pause for longer immediately after a word than before a word. By contrast, Mary exhibited a pattern of longer pauses immediately preceding the word compared to the length of her after-word pauses (see Figure 6 and Figure 7).

S-Notation (Dutch)	S-notation (translation)	BeforeWord 2	BeforeWord1	AfterWordPause
ik-	*I*	0	0	1124
zie-	*see*	1124	6364	1155
een-	*a*	1155	1310	1341
kind-	*child*	1341	2824	499

Figure 7. Example of linguistic analysis including pause information (Mary(d))

Table 3 presents the basic pausing information from the linguistic analysis. This analysis complements the pause analysis data previously presented in Table 2. In Table 2 we reported an average of 53.5 pauses between words for Elise(h) and 39 for Mary(d) for the 0.2 ms pause threshold. However, if we fine-tune the pause analysis for the conduct of our linguistic analysis, we can look in greater detail to the 99 pauses for Elise(h) and 60 for Mary(d) in both writing tasks. Since we decided to focus on those pausing times that clearly indicate cognitive effort related to *producing* a word, we excluded revisions from the current evaluation since they disrupt the data by introducing cognitive effort of a different kind. We also removed extremely long pauses of more than 10 seconds (2 in the case of BFW-1 and 5 for variable BFW-2). Finally, we had to manually correct the automated word reconstruction of Inputlog in a few instances. Examples of such corrections are incorrectly connected words (*halende* ~ *halen de*) and grossly misspelled words (*kantwkanteken* ~ *kantelen*). As a result of this intervention, the number of pauses in Table 3 differs slightly from the numbers and means mentioned in Table 2.

Table 3. Mean pause duration before words (–1 and –2)

	Elise (healthy)	Mary (demented)
Pause information (threshold 0.2s)		
Total number of pauses	96	43
Mean duration Before Word Pauses (sum of –1 and –2)	1718	2661
Mean duration Before Word Pause (–1)	817	1958
Mean duration Before Word Pause (–2)	901	704
Mean duration After Word Pause (–2)	837	749

The pauses between words (before-word pauses –1- and –2) were about 1 second shorter for Elise(h) than for Mary(d). The summed pauses for Elise(h) consisted of two pauses of comparable length, whereas the pauses for Mary(d) were more than twice as long as the preceding pause (–2) just before a new word was produced (–1).

Figure 8 shows the number and mean of the most frequently used word categories (The information on pausing times is presented in Table 5 in the Appendix). By selecting word categories that were used at least 5 times, we provide an overview of more than 90% of the data for each participant (Elise(h): 93.75%, and Mary(d): 90.70%). The difference between the two participants is due to the fact that Elise(h) regularly used connectives (4) and adjectives (7) in her text, whereas only one adjective occurred in Mary's text. The remainder of the infrequently used word categories were adverbs and unspecified tokens (spec). (An overview of the word categories identified by the linguistic analysis is provided in the Inputlog manual (Leijten & Van Waes 2014)).

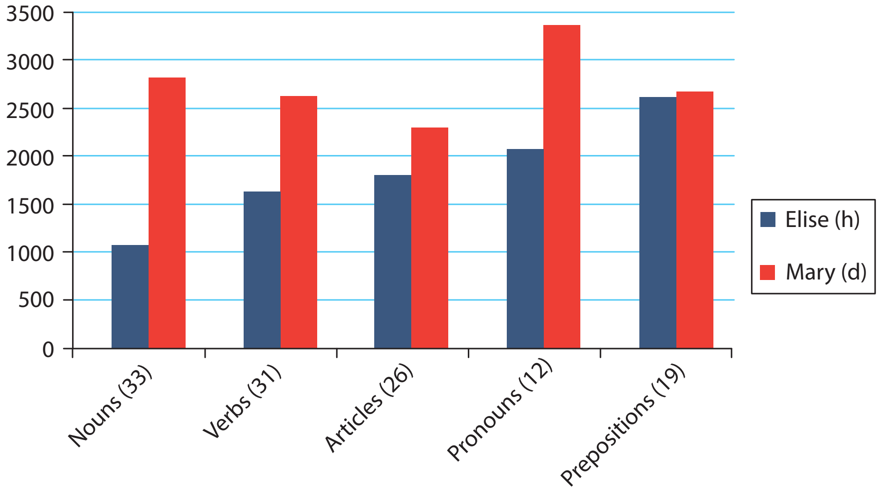


Figure 8. Number of between word pauses and mean pause duration before words per word category

The least demanding word category for Elise(h) seems to have been nouns (1077 ms), with the pause length lengthening gradually from verbs to articles and then on to pronouns. On average, Elise reflected for longest (2630 ms) in the case of prepositions, which often introduced more extensive prepositional phrases including articles. This same hierarchy is not reflected in Mary's data. The differences between the word category-related before-word pauses fluctuated less but were still in all cases longer than those produced by Elise. In particular, nouns, verbs, and pronouns seem to be more cognitively demanding for the participant with dementia, since the mean pause durations on these items were about 1 second longer than for the healthy elderly participant, Elise. The data shows that producing a pronoun required the most effort for the demented participant.

Importantly, the pattern of mean pause lengths before articles and nouns differed between Elise(h) and Mary(d). Mary(d) required a lengthy pause before articles and an even longer pause before nouns (as shown in Figure 5), while Elise(h) required a longer pause before articles than before nouns.

Figure 9 shows that to write the noun phrase '*the kitten*', Elise paused for 3229 ms before the article *the*, and 1030 ms before the noun *kitten*. Pauses after the production of an article were in general relatively short (437 ms). A similar pattern can be found before the production of the more complex noun phrase '*the* *goldfish* (*in the bowl*)'. In this case, the initial pause was longer than 4 seconds. These examples clearly demonstrate the importance of, and the added value conferred by, linguistic diversification in between-word pausing patterns. The extra layer to the pause analysis refines the interpretation of cognitive pauses to a large extent. However, they also show that further fine-tuning of the data is undoubtedly needed in order to better explain the complexity of these pausing patterns, both relative to one another and as a function of the syntactic structure.

S-Notation (Dutch)	S-Notation (translation)	BeforeWord1+2	AfterWordPause
dat-	that	1482	2683
de-	the	3229	437
poes-	kitten	1030	3073
de-	the	4071	359
goudvis-	goldfish	827	327
in-	in	811	281
de-	the	562	562
kom-	bowl	1030	2980
wil-	will	4150	608
vangen-	catch	889	3135

Figure 9. Partial sentence showing pausing times before articles and nouns (Elise(h)). [translation at word level]

To a certain extent, this information is already provided by the chunker since this returns information allowing us to discriminate between pauses at the beginning of a chunk and those inside a chunk. Table 4 summarizes the data for both persons from this perspective.

Table 4. Mean pause duration before words (–1 and –2) at the beginning of a chunk and inside a chunk

	Elise (healthy)	Mary (demented)
	Mean	**Mean**
Beginning	2061	2600
Inside	1049	2821

The mean pause length of the healthy elderly participant Elise was twice as long at the beginning of a chunk as inside a chunk. By contrast, Mary(d) exhibited a pause length of 2600 ms at the beginning of and about 2800 ms inside a chunk. In combination with the pausing data from Table 5 (Appendix), this suggests that

Mary's efforts were more fragmented and occurred at a lower level. It seems that her text production evolved as a staccato word-by-word sequence. Every word required an almost equal amount of effort: at the beginning of a phrase, within a phrase, at the beginning of a chunk, or inside a chunk. Elise's pattern, on the other hand, seems to reflect more diversification, probably due to the fact that she was able to plan larger text sections.

5. Conclusion and discussion

Keystroke logging has become instrumental to observe and analyze writing processes. This chapter summarizes the use of keystroke logging as a research technique in general. It also reviews three freely available research tools: ScriptLog, Translog and Inputlog.

To date, (automated) keystroke logging analyses have been mainly based on data obtained at the character level. Although it is clear that this fine-grained, low-level approach leads to very interesting insights, a long tradition of product analysis has taught us that more high-level analyses could also open up new avenues of research. Therefore, Inputlog has been extended by a so-called linguistic analysis in which data is aggregated through to the word level. This module facilitates linguistic process analysis by taking account of the dynamics of writing as the text unfolds. The linguistic module has been developed in English and Dutch, but can potentially also be used for other (Western) languages thanks to the generic approach adopted during its development.

This chapter explains the operation of the module and provides a case study by way of example. In this case study, we show that it is very important to connect the general mental processes observed in writers, on the one hand, with the linguistic features of the text, on the other hand. The case study clearly shows that 'a pause' is too broad a concept, even when we subdivide pauses into different levels (character – word – sentence etc.). We contend that in order to better understand the underlying cognitive processes, the concept of 'pause' needs to be further defined.

In the case study, we described the cognitive processes characteristic of the text production of two elderly people in a controlled task environment. We selected a healthy elderly woman (Elise) and a demented woman (Mary) whose profiles matched in terms of age, education and working career. The product data showed that the healthy elderly participant was able to produce a longer text (about 10 more words) to describe the picture presented to her. When production time is taken into account, it took the demented participant about 2 minutes longer to produce the texts. Moreover, her texts were shorter and she composed about 7 words per minute. In contrast, the healthy elderly participant produced almost twice as many words per minute (about 12 words). Mary(d) paused about 39 times, whereas Elise(h) paused 25 times. However, as stated above, comparing pausing behavior based on a 2 s pause threshold is perhaps not the best approach if we also wish to address lower-level differences (cf. average pause length of 5.08 s for the healthy elderly participant compared to 5.86 s for the demented elderly woman). Our further results, involving an analysis of within – and between – word pauses using a lower threshold of 200 ms, showed that the pauses were twice as long for the demented participant than for the healthy participant (i.c. within words: 0.80 versus 1.59 s; between words: 1.74 versus 2.99 s).

Furthermore, the new automated linguistic analysis showed that the demented participant took about three times as long to produce nouns (difference of 1750 ms) and twice as long to produce verbs (difference of 1000 ms). By contrast, the pause time before articles differed by about 400 ms. The combined results of the various levels of pause analysis as a function of linguistic feature showed that Mary(d) struggled throughout the writing process as she moved from word to word and that this occurred both at the beginning of a phrase and during a phrase. Elise(h) seemed to produce phrases more fluently and in longer bursts. These production units reveal a pausing behavior with a quite considerable within-participant variance and seem to be defined, to a large extent, by linguistic and syntactic characteristics.

We hope to have demonstrated that automated linguistic analysis provides a large volume of rich data that opens up new avenues for writing process analyses based on keystroke logging. The added value brought about by the further differentiation between different types of between-word pauses undoubtedly merits further exploration and will hopefully lead to a better understanding of the underlying cognitive processes that characterize pause behavior. It is important to remember, however, that – despite the use of sophisticated NLP tools – this type of analysis is more sensitive than, e.g. a general pause analysis. Process data are much more complex than product data, and therefore a certain degree of 'noise' occurs. A typical example is the case in which an unfinished word is deleted during the process, and is presented as such to the linguistic analysis. For instance, when analyzing Mary's data, we had to deal with data loss of about 25% due to complexities in the data, mainly in the form of unrecognized (non-existent or misspelled) words. Adding linguistic features to pauses at the word level has proved to be a first step and is certainly worth further exploration. Moreover, although we believe that adding linguistic features to the pause analysis is an important first step in further diversifying the analysis of cognitive processes, it should be remembered that 'a pause' is still a complex construct that needs to be defined in greater detail and from other theoretical perspectives. For instance, pauses between words are made up of *before* and *after*-word pauses and individuals deal with these in different ways, as they do in the case of pauses before and after a full stop (Van Waes & Leijten 2011; 2014). Consideration of this type of interpersonal difference – perhaps in combination with the study of individual motor and typing skills – constitutes an avenue that is clearly worthy of further exploration.

As stated in the introduction, the present research project combines process information with linguistic characteristics. Future analyses will focus on the richness of the written output relative to the cognitive effort invested by writers in order to produce these texts. The process measures can be matched to product measures (final text), including *word diversity* and *expressivity*.

During the remainder of this research project, it is our goal to describe, on a larger scale, the changes that occur during the different stages of AD development, on the one hand, and to test the diagnostic potential for discriminating AD sufferers from controls, on the other. Furthermore, by linking writing process data to lexica and by using NLP tools, we will be able to analyze the data on a higher, more complex level, while also using more advanced statistical techniques that take into account the hierarchical character of the data and the underlying patterns. In this way, we hope to stimulate interdisciplinary research at the crossroads of product and process analysis.

Acknowledgements

The linguistic analysis was partially funded by a research grant from the Flanders Research Foundation (FWO 2009–2012; in collaboration with Véronique Hoste and Lieve Macken – http://inputlog.ua.ac.be/WebSite/). Mariëlle Leijten received a grant for post-doctoral researchers from the Research Foundation – Flanders (FWO) to conduct the described research project. We are very grateful to Prof. Dr. Sebastiaan Engelborghs and Dr. Stefaan Van Der Mussele for enabling access to the patients at the Memory Clinic of the Antwerp Middelheim and Hoge Beuken Hospital Network (ZNA). Finally, we thank the Master's students in Multilingual Professional Communication for their help in gathering the data (Magali Colemont, Ester Coppieters, Astrid Danau, Aline De Weerdt, Anna-Catherina Rossaert, Daniël Ter Laan, Marie-Claire Van Heeswijk, and Evelien Wouters).

References

Baaijen, Veerle M., David Galbraith, and Kees de Glopper. 2012. "Keystroke Analysis Reflections on Procedures and Measures." *Written Communication* 29(3): 246–277. DOI: 10.1177/0741088312451108

Baaijen, Veerle M., David Galbraith, and Kees de Glopper. 2014. “Effects of writing beliefs and planning on writing performance.” *Learning and Instruction* 33(0): 81–91. DOI: 10.1016/j.learninstruc.2014.04.001

Bazerman, C., ed. 2008. *Handbook of Research on Writing: History, Society, School, Individual, Text*. New York and London: Routledge, Taylor & Francis Group.

Bazerman, Charles, Robert Krut, Karen Lunsford, Susan McLeod, Suzie Null, Paul Rogers, and Amanda Stansell. 2010. *Traditions of Writing Research*. New York and London: Routledge, Taylor & Francis Group.

Berninger, Virginia 2012. *Past, Present, and Future Contributions of Cognitive Writing Research to Cognitive Psychology*. New York and London: Routledge, Taylor & Francis Group.

Caporossi, Gilles, and Christophe Leblay. 2011. “Online writing data representation: a graph theory approach.” *Advances in Intelligent Data Analysis* X:80–89. DOI: 10.1007/978-3-642-24800-9_10

Carl, Michael. 2012. “Translog-II: a Program for Recording User Activity Data for Empirical Reading and Writing Research.” Paper read at *LREC*.

Doherty, Stephen, and Sharon O’Brien. 2014. “Assessing the Usability of Raw Machine Translated Output: A User-Centered Study Using Eye Tracking.” *International Journal of Human-Computer Interaction* 30(1): 40–51. DOI: 10.1080/10447318.2013.802199

Ehrensberger-Dow, Maureen, and Daniel Perrin. 2009. “Capturing translation processes: a multi-method approach.” *Across Languages and Cultures* 20(2): 275–288. DOI: 10.1556/acr.10.2009.2.6

Folstein, Marshall, Susan E. Folstein, and PR McHugh. 1975. “Mini-mental state. A practical method for grading the cognitive state of patients for the clinician.” *Journal of Psychiatric Research* 12: 189–198.

Flower, Linda, and John R. Hayes. 1981. “A cognitive process theory of writing.” *College Composition and Communication* 32: 365–387. DOI: 10.2307/356600

Goodglass, Harold, Edith Kaplan, and Barbara Barresi. 1983. *Boston Diagnostic Aphasia Examination* (*BDAE*). Philadelphia: Lea and Febiger. DOI: 10.1002/ana.410160524

Gunawardhane, Suranga DW, Pasan M De Silva, Dayan SB Kulathunga, and Shiromi MKD Arunatileka. 2013. “Non invasive human stress detection using key stroke dynamics and pattern variations.” Paper read at *Advances in ICT for Emerging Regions* (ICTer), 2013 International Conference on. DOI: 10.1109/icter.2013.6761185

Hayes, John R. 1996. “A new framework for understanding cognition and affect in writing.” In *The science of Writing: Theories, Methods, Individual Differences, and Applications*, ed. by C.Michael Levy, and Sarah E. Ransdell, 1–27. Mahwah: New Jersey: Lawrence Erlbaum Associates.

Hayes, John R. 2012a. “Modeling and remodeling writing.” *Written Communication* 29(3): 369–388. DOI: 10.1177/0741088312451260

Hayes, John R. 2012b. “My Past and Present as Writing Researcher and Thoughts About the Future of Writing Research.” In *Past, Present, and Future Contributions of Cognitive Writing Research to Cognitive Psychology*, ed. by Virginia Berninger, 3–26. New York: Taylor and Francis Group, Psychology Press. DOI: 10.4324/9780203805312

Jakobsen, Arnt. L. 2006. “Translog: Research methods in translation.” In *Computer Keystroke Logging and Writing: Methods and Applications*, ed. by Kirk P. H. Sullivan, and Eva Lindgren, 95–105. Oxford: Elsevier.

Jakobsen, Arnt L. 2011. “Tracking translators’ keystrokes and eye movements with Translog.” *Methods and Strategies of Process Research: Integrative Approaches in Translation Studies* 94: 37. DOI: 10.1075/btl.94.06jak

Johansson, Victoria, Åsa Wengelin, Johan Frid, and Roger Johansson. 2014. "ScriptLog 2013 state of the art." In *Training school on keystroke logging*. University of Antwerp, Belgium.

Jurafsky, Daniel S., and James H. Martin. 2009. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Vol. 3. New Jersey: Pearson Education Inc. DOI: 10.1162/089120100750105975

Kellogg, Ronald T. 2008. "Training writing skills: A cognitive developmental perspective." *Journal of Writing Research* 1(1): 1–26.

Leblay, Christophe, and Gilles Caporossi. 2014. *Temps de l'écriture: Enregistrements et représentations*. Vol. 12: Academia/L'Harmattan.

Leijten, Mariëlle, Sven De Maeyer, and Luuk Van Waes. 2011. "Coordinating sentence composition with error correction: A multilevel analysis." *Journal of Writing Research* 2(3): 331–363. DOI: 10.1007/s11145-009-9190-x

Leijten, Mariëlle, Lieve Macken, Veronique Hoste, Eric Van Horenbeeck, and Luuk Van Waes. 2012. "From character to word level: Enabling the linguistic analyses of Inputlog process data." In *European Association for Computational Linguistics, EACL – Computational Linguistics and Writing (CL&W 2012): Linguistic and Cognitive Aspects of Document Creation and Document Engineering*, ed. by Michael Piotrowski, Cerstin Mahlow, and Robert Dale. Avignon. http://aclweb.org/anthology/W/W12/W12-03.pdf.

Leijten, Mariëlle, and Luuk Van Waes. 2012. "Inputlog 4.0: Keystroke Logging in Writing Research." In *Learning to Write Effectively: Current Trends in European Research*, ed. by Mark Torrance, Denis Alamargot, Montserrat Castelló, Franck Ganier, Otto Kruse, Anne Mangen, Liliana Tolchinsky, and Luuk Van Waes, 363–366. Emerald Group Publishing Limited.

Leijten, Mariëlle, and Luuk Van Waes. 2013. "Keystroke logging in writing research: Using Inputlog to analyze and visualize writing processes." *Written Communication* 30(3): 358–392. DOI: 10.1177/0741088313491692

Leijten, Mariëlle, and Luuk Van Waes. 2014. *Manual Inputlog 6.0*. Antwerp: University of Antwerp.

Leijten, Mariëlle, Luuk Van Waes, Karen Schriver, and John R. Hayes. 2014. "Writing in the workplace: Constructing documents using multiple digital sources." *Journal of Writing Research* 5(3): 285–336.

Lindgren, Eva, Mariëlle Leijten, and Luuk Van Waes. 2011. "Adapting to the reader during writing." *Written Language & Literacy* 14: 188–223.

MacArthur, Charles A., Steve Graham, and Jill Fitzgerald. (Eds.). 2008. *Handbook of Writing Research*. New York, NY: The Guilford Press. DOI: 10.1111/j.1467-873x.2008.00423.x

Macgilchrist, Felicitas, and Tom Van Hout. 2011. "Ethnographic discourse analysis and social science." Paper read at *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*.

Maggio, Severine, Bernard Lété, Florence Chenu, Harriet Jisa, and Michel Fayol. 2012. "Tracking the mind during writing: Immediacy, delayed, and anticipatory effects on pauses and writing rate." *Reading and Writing* no. 25 (9): 2131–2151.
DOI: 10.1007/s11145-011-9348-1

Manning, Christoper, D., and Hinrich Schütze. 1999. *Foundations of Statistical Natural Language Processing*. Cambridge, MA: The MIT Press. DOI: 10.1162/coli.2000.26.2.277

McKhann Guy, David Drachman, Marshall Folstein, et al. 1984. "Clinical diagnosis of Alzheimer's disease." Report of the NINCDSADRDA work group under the auspices of the Department of Human Services Task Force on Alzheimer's disease. *Neurology* 34: 939–944.

Mesulam, M-Marsel, Murray Grossman, Argye Hillis, Andrew Kertesz, and Sandra Weintraub. 2003. "The core and halo of primary progressive aphasia and semantic dementia." *Annals of Neurology* 54(5): 11–14. DOI: 10.1002/ana.10569

Petersen, Ronald C. 2004. “Mild cognitive impairment as a diagnostic entity.” *Journal of Internal Medicine* 256(3): 183–194.

Risku, Hanna, Florian Windhager, and Matthias Apfelthaler. 2013. “A dynamic network model of translatorial cognition and action.” *Translation Spaces* 2(1): 151–182. DOI: 10.1075/ts.2.08ris

Robert, Isabelle S., and Luuk Van Waes. 2014. “Selecting a translation revision procedure: do common sense and statistics agree?” *Perspectives*: 1–18. DOI: 10.1080/0907676x.2013.871047

Schilperoord, Joost. 1996. *It’s about time. Temporal aspects of cognitive processes in text production*. Amsterdam/ Atlanta: Rodopi.

Severinson Eklundh, Kerstin, and Py Kollberg. 2002. “Studying writers’ revision patterns with S-notation analysis.” In *Contemporary tools and techniques for studying writing*, ed. by Thierry Olive, and C. Michael Levy, 89–104. Dordrecht: Kluwer Academic Publishers. DOI: 10.1007/978-94-010-0468-8_5

Spelman Miller, Kristyan, Eva Lindgren, and Kirk P. H. Sullivan. 2008. “The psycholinguistic dimension in second language writing: Opportunities for research and pedagogy using computer keystroke logging.” *TESOL Quarterly* 42(3): 433–454.

Sperling, Reisa A., Paul S. Aisen, Laurel A. Beckett, David A. Bennett, Susanne Craft, et al. 2011. “Toward defining the preclinical stages of Alzheimer’s disease: recommendations from the National Institute on Aging-Alzheimer’s Association workgroups on diagnostic guidelines for Alzheimer’s disease.” *Alzheimers Dement* 7: 280–292.

Sullivan, Kirk P.H., and Eva Lindgren. 2006. *Computer Key-Stroke Logging and Writing*. Edited by G. Rijlaarsdam, *Studies in Writing*. Oxford: Elsevier Science.

Van Eynde, Frank, Jakub Zavrel, and Walter Daelemans. 2000. “Part of speech tagging and lemmatisation for the Spoken Dutch Corpus.” In *Proceeding of the Second International Conference on Language Resources and Evaluation*, ed. by M. Gavrilidou et al., 1427–1433. Athens.

Van Horenbeeck, Eric, Tom Pauwaert, L. Van Waes, and M. Leijten. 2012. *S-notation: S-notation markup rules* (Technical Description). Antwerp: University of Antwerp.

Van Waes, Luuk, and Mariëlle Leijten. 2011. “Observing and analysing digital writing processes with Inputlog.” In *Antwerp Summer School on Writing Process Research: Keystroke logging and Eyetracking*. Antwerp. DOI: 10.1177/0741088313491692

Van Waes, Luuk, and Mariëlle Leijten. 2013. “Vlot schrijven-Een multidimensioneel perspectief op ‘writing fluency’.” *Tijdschrift voor taalbeheersing* 35(2): 160–182. DOI: 10.5117/tvt2013.2.waes

Van Waes, Luuk, and Mariëlle Leijten. 2014. Inputlog 6.0: Pause and fluency analysis.” In *Keystroke logging training school*. Antwerp.

Van Waes, Luuk, Mariëlle Leijten, and Aline Remael. 2013. “Live subtitling with speech recognition. Causes and consequences of text reduction.” *Across Languages and Cultures* 14(1): 15–46. DOI: 10.1556/acr.14.2013.1.2

Van Waes, Luuk, Mariëlle Leijten, Åsa Wengelin, and Eva Lindgren. 2012. “Logging tools to study digital writing processes.” In *Past, Present, and Future Contributions of Cognitive Writing Research to Cognitive Psychology*, ed. by Virginia Wise Berninger, 507–533. New York/Sussex: Taylor & Francis.

Van Waes, Luuk, and Peter Jan Schellens. 2003. “Writing profiles: The effect of the writing mode on pausing and revision patterns of experienced writers.” *Journal of Pragmatics* 35(6): 829–853.

Visch-Brink, Evy, Dorien Vandenborre, Hyo Jung De Smet, and Peter Mariën. 2014. *The Comprehensive Aphasia Test-NL, Pearson*. Amsterdam.

Wengelin, Åsa. 2006. “Examining pauses in writing: Theories, methods and empirical data”. In *Computer Keystroke Logging and Writing: Methods and Applications*, ed. by Kirk P. H. Sullivan, and Eva Lindgren, 107–130. Oxford, UK: Elsevier.

Wengelin, Åsa, Mark Torrance, Kenneth Holmqvist, Sol Simpson, David Galbraith, Victoria Johansson, and Roger Johansson. 2009. "Combined eye-tracking and keystroke-logging methods for studying cognitive processes in text production." *Behavior Research Methods* 41(2): 337–351. DOI: 10.3758/brm.41.2.337

Wininger, Michael. 2014. "Measuring the evolution of a revised document." *Journal of Writing. Research* 6(1): 1–28.

Yi, Hyon-Ah, Peachie Moore, and Murray Grossman. 2007. "Reversal of the Concreteness Effect for Verbs in Patients with Semantic Dementia." *Neuropsychology* 21(9): 9–19. DOI: 10.1037/0894–4105.21.1.9

Appendix

Dutch example of final text produced by healthy elderly woman: Elise(h) (81)

"De ene ramp na de andere: de afwasbak van de mama loopt over (is de kraan geblokkeerd ?) zoonlief wil heimelijk koekjes uit de koekendoos halen, zijn stoel kantelt en hij zal waarschijnlijk op de grond vallen. Wil kleine zus ook een koekje of lacht zij hem uit ?Antwoord op het volgende plaatje."

Dutch example of final text produced by elderly woman with dementia Mary(d) (79):

"ik zie een kind dat een bord iot de kast wenst te halende moeder is eeen bord aan jet afdrogen. het stoeltje waarop de jongen staat is aanhet kantwkanteken; ik denk fat er verscheidene bit borden zullensneuvelenmm moeder is aan het afdeogen er valt warze p op de gron, grond xus zie ik nog andere ongelukkengebeuren."

Table 5. Number of pauses and mean pause duration before words (–1 and –2) per word category

	Elise (healthy)		Mary (demented)	
	Number	Mean	Number	Mean
Articles (26)	19	1812	6	2288
Nouns (33)	23	1077	10	2827
Verbs (31)	20	1639	11	2625
Prepositions (19)	13	2630	6	2681
Pronouns (12)	8	2084	5	3385