

USING MICRO-CONTEXTS TO DESCRIBE A WRITING PROCESS IN ESTONIAN AS A SECOND LANGUAGE ACROSS PROFICIENCY LEVELS

Olga Pastuhhova

Abstract. This paper aims to describe the writing process of native Russian-speaking students in Estonian as a second language. 34 participants were given the assignment of writing a text in the L2. The written texts were then rated as being at levels from A2 to C1 according to the Common European Framework of Reference for Languages (CEFR). The data were collected by computer keystroke logging and analysed based on the concept of a 'micro-context'. Micro-contexts were analysed according to their frequency and duration and were compared across proficiency levels. The results show that writing in the L2 is not a smooth process. The longest transitions in micro-contexts reveal that the most cognitive effort is made between paragraphs and sentences and when deletions are involved. The growing number of consecutive deletions demonstrates that even with developing proficiency, the linear production text is subject to constant revision, correction and modification.

Keywords: keystroke, ScriptLog, Estonian, Russian

1. Introduction

Computer keystroke logging (Sullivan, Lindgren 2006, Van Waes et al. 2009, Van Waes et al. 2011, Leijten, Van Waes 2013) is a data collection method for recording writing in real time. It helps to reveal the cognitive processes underlying written text production. According to John R. Hayes (2012), the writing model is represented at three levels: the control level (e.g., motivation and goal setting), the process level (writing processes and task environment), and the resource level (e.g., attention, working memory). Writing processes at the process level include a proposer, translator, transcriber, and evaluator component. Mariëlle Leijten and Luuk Van Waes (2013) situate the methodological contribution of keystroke logging mainly

at the process level of Hayes' latest writing model (2012). Keystroke logging as a method is mainly aimed at observing translation (or formulation) and transcription processes, but is not limited to the process level. Since logging data provide information on pausing and revision behaviour, the method is also inferred to the resource level (Leijten, Van Waes 2013: 362).

Keystroke logging has been used in researching writing in the L1 and L2, in translation studies (Jakobsen 2006, Dimitrova 2006), as well as in researching writers with a language impairment (Asker-Árnason et al. 2010). Guido Nottbusch (2010) divides keystroke logging research into single word studies and typewritten composition. Evgeny Chukharev-Hudilainen (2014) links two strands of keystroke logging research to analyse keystroke patterns in spontaneous computer-mediated communication. In L1 and L2 writing research, keystroke logging has been used to investigate, for example, revision (Lindgren 2005, Stevenson et al. 2006), pauses (Spelman Miller 2000, 2006a, 2006b) and fluency (Uppstad, Solheim 2007, Palviainen et al. 2012).

The present study combines writing process and second language acquisition research. Namely, the writing process of native Russian-speaking students in Estonian as an L2 is under investigation. Writing is an important skill of second language proficiency. Previous research in the field of acquisition of Estonian as an L2 has largely focused on analysing the final product, while internationally it is the writing process itself that has become a subject to be explored. The shift to writing process research in L2 Estonian was noted by Olga Pastuhova (2011) in an article which introduced computer keystroke logging as a relevant method for investigating the writing process, previous research based on the method and a case study of the writing process of a native Russian-speaking learner of Estonian.

The participants of the present study are Russian-speaking students at Tallinn University. The computer program ScriptLog was used to collect the data. The program keeps a detailed record of activity on the keyboard during a writing assignment (Strömquist, Karlsson 2002, Strömquist et al. 2006, Johansson 2009). The analysis is based on the concept of a micro-context, defined as 'the context around the certain transition between two keystrokes' (Wengelin 2006: 107). Although each transition between two keystrokes includes a short period of inactivity, it is not reasonable to interpret each transition as a pause. 'A working definition is that a pause is 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' (Wengelin 2006: 111). The criterion of two seconds in the current study is based on previous research that described inactivity between two keystrokes that was longer than two seconds as a pause (Wengelin 2006).

The present study aims to examine what an analysis of micro-contexts can tell us about text translation (or formulation) and transcription processes in terms of Hayes' latest writing model (2012) in general, and about L2 writing proficiency in particular. The study concerns L1 Russian writers of Estonian as an L2. Micro-contexts are analysed based on the frequency of their occurrence and the duration of inactivity between keystrokes. Since text production is investigated in L2 writing, the study aims to compare the findings across proficiency levels, from A2 to C1 according to the Common European Framework of Reference for Languages: Learning, Teaching Assessment (CEFR), to point out similarities and differences

in text production at the different levels. The data enable us to describe a writing process with the following components: text composition; text length; syntactic complexity; and deletions/mouse events (see section 3). The keywords of the data analysis in the present study are micro-contexts, frequency of their occurrence, duration, and proficiency levels.

2. Previous research

Micro-context is a key concept used by Åsa Wengelin (2006) in analysing the distribution of pauses in Swedish online text-writing by university students with no reading and writing difficulties and adults with reading and writing difficulties (dyslexia and deafness). The method was also applied to detect developmental change in the allocation of cognitive effort to different levels of text-writing by native speakers of Swedish across four age-specific groups: 4th, 7th and 10th graders, and university students. Their results indicate that at an early phase of text-writing development, more cognitive resources are allocated to low-level processes, such as spelling and punctuation. Later the focus is shifted to higher-level activities, such as the content, structure, and the text style (Wengelin, Strömquist 2004: 189–190). Pause frequency was observed in six micro-contexts: before word, within word, after word, before sentence, before delimiter¹, and after sentence. The pause frequency profiles of the four age-specific groups, the university students with no reading and writing difficulties, and the adults with reading and writing difficulties were found to be surprisingly similar (Wengelin 2006: 119–120). The 4th graders and the 7th graders differed from the 10th graders and university students in two main aspects only: the 4th graders made many more pauses before words in comparison with the others and both the 4th and the 7th graders made more pauses before major delimiters than before a sentence or after a delimiter. The pause frequency of the adults with reading and writing difficulties resembled a combination of the 4th and 7th graders. The main findings of pause behaviour of all groups of subjects were found to be as follows: ‘micro-contexts around sentence boundaries are more predictive of pauses than micro-contexts around words’, ‘micro-contexts before words are generally more predictive of pauses than micro-contexts within and after words’, ‘micro-contexts before sentences are more predictive of pauses than micro-contexts after sentences’, ‘micro-contexts just before a major delimiter change systematically with age or perhaps writing development’ (Wengelin 2006: 120–121).

Kristyan Spelman Miller (2006b: 141–142) points out that pausological research on writing has mainly focused on native speakers, although recently the situation has changed to include L2 writers as well. The author investigated the pausing of L1 and advanced L2 writers of English and found differences in pause duration and pause frequency between the groups. Pauses are longer and more frequent for L2 writers, especially at lower text (often spelling and morphological) unit levels.

The present study involves an analysis of micro-contexts in L2 writing by L1 Russian learners of Estonian.

¹ Major delimiters refer to a full stop, a question mark or an exclamation mark. Minor delimiters refer to a comma, a colon, a semicolon, a dash or a pair of brackets (Wengelin 2006: 114).

3. Data and method

The data of the study consist of texts written by 34 participants whose L1 is Russian and for whom Estonian is an L2. For the most part (31 out of the 34), the participants were bachelor students at Tallinn University; the other three were studying at the master's level. All participants attended an Estonian language course in different groups according to their proficiency level. Being students, the participants were regularly made to write, whether in the form of essays, examinations or assignments for language courses. The students were asked to participate in a study concerning the writing process in Estonian as an L2. 34 participants agreed to take part in the study and the author met everyone individually. All participants were asked to write a text on a topic related to education for the purpose of the study. The topic was given on the basis of participants' expected proficiency level. Since text type is in some sense a property of proficiency level, more advanced learners wrote an argumentative text on the topic 'Is higher education a requirement for everyone?', while less advanced learners wrote a narrative text on the topic 'What I have gained from university'. All texts were written on a computer using the program ScriptLog. The participants were not allowed to use dictionaries or Internet materials. All were given one hour to express their thoughts on the topic and finish the assignment, but no time pressure was established. Upon completion of the task, the students were asked to comment on their writing in general as well as the task specifically. These interviews were taped. Retrospective data remain outside the scope of the present study.

The written texts were rated by two experts on the basis of the illustrative scales provided for overall written production, creative writing, and reports and essays of the Common European Framework of Reference for Languages (CEFR: 61–62). The data were representative of four language proficiency levels: 4 texts corresponded to A2 (all texts narrative), 13 texts to B1 (12 texts narrative and 1 argumentative), 13 texts to B2 (12 texts argumentative and 1 narrative), and 4 texts to C1 (all texts argumentative).

The data were analysed based on the micro-contexts registered by ScriptLog. The micro-contexts allowed for the description of a text production process according to the following criteria: text composition; text length; syntactic complexity; and deletions/mouse events. Notations and descriptions of micro-contexts recordable by ScriptLog and used in the study are found in Appendix 1.

What is meant by **text composition** is the dividing of the text into paragraphs. It is tracked in terms of two micro-contexts: transition to a new paragraph ('.[^]c') and starting a new paragraph ('c[^]a'). As regards **text length**, ScriptLog provides a statistics output file stating the number of tokens in the linear as well as the final text. Although this is an easy and straightforward way to calculate the length, the data provided with micro-contexts will be discussed here (transitions inside words ('a[^]a'), between words ('a_[^]a', 'a[^]_+a'), and between sentences ('a[^].', '.[^]_+a', '.[^]+a')).

What is meant by **syntactic complexity** is the use of minor delimiters ('a[^].', '.[^]_+a', '.[^]+a'). **Deletions/mouse events** are registered in a number of different surroundings and they give an overview of deletion behaviour during text production: 'D[^]D' represents consecutive deletions; 'a[^]D', 'D[^]a' word level deletions;

‘a^_+D’ the cases where a typer taps the spacebar upon completion of the previous word with the intention to continue producing the text but ends up with a deletion instead; ‘_^D’, ‘D^_’ deletions between words; and ‘D_+^a’ transition to continuing constructing the text after the previous text has been edited. It is worth noting that ScriptLog denotes both deletions and mouse events with the same symbol. For this reason the discussion of these micro-contexts concerns deletions, mouse events or both.

4. Results

The findings of the study are presented in the following order: first, the analysis of micro-contexts based on the frequency of their occurrence; second, the analysis of micro-contexts based on the duration of inactivity between the keystrokes. The frequency and duration of the micro-contexts registered are compared by proficiency level. Finally, a general overview of the text production process across the proficiency levels and potential pause locations is discussed. Since the number of participants varied per level and not all micro-contexts were present in every text, the mean numbers of micro-contexts were used and the mean of the mean transition time in every micro-context is analysed per text.

4.1. Micro-contexts by frequency

Table 1 below presents the most frequent micro-contexts across the proficiency levels.

Table 1. Micro-contexts with mean frequency of occurrence exceeding 10 times per text across the proficiency levels

A2 (N = 4)		B1 (N = 13)		B2 (N = 13)		C1 (N = 4)	
Micro-context	Mean N/ text	Micro-context	Mean N/ text	Micro-context	Mean N/ text	Micro-context	Mean N/ text
a^a	938	a^a	1132	a^a	1299	a^a	1214
a^_+a	180	D^D	333	D^D	340	D^D	391
a_+^a	180	a^_+a	212	a^_+a	220	a^_+a	208
D^D	156	a_+^a	212	a_+^a	220	a_+^a	208
D^a	43	D^a	56	D^a	69	D^a	59
a^D	41	a^D	52	a^D	62	a^D	56
a^.	27	_ ^D	23	,^_+a	31	,^_+a	25
a^,	22	a^,	22	,_+^a	31	,_+^a	25
,^_+a	26	,^_+a	20	_ ^D	27	a^,	25
,_+^a	26	,_+^a	20	a^,	25	_ ^D	18
_ ^D	16	a^.	19	a^_+D	20	a^_+D	14
		a^_+D	17	a^.	17	a^.	14
		D^_	13	_ ^_	22	D^_	13
				D^_	15	D_+^a	12
				D_+^a	11		

The most frequent micro-context across all proficiency levels reflects continuing to construct a word. However, the second most frequent micro-context at B1, B2 and C1 levels demonstrates that text production is not a smooth process and is characterised by a high level of consecutive deletions or mouse movements. Their number exceeds that of transitions between words across all the proficiency levels except for A2 level. The fifth and sixth positions in the frequency order across all levels involve deletions inside a word. The next micro-context at A2 level is associated with finishing a sentence followed by use of minor delimiters. At B1, B2 and C1 levels the opposite is true – the use of minor delimiters per text is more frequent micro-contexts related to sentence finishing. Other micro-contexts with mean frequency of occurrence exceeding 10 times per text include, again, transitions in several surroundings with deletion involved. The full list of micro-contexts across all levels together with their total number, mean number per text and mean of the mean transition time per text is to be found in the appendices (Appendix 2, 4, 6, 8).

The first six micro-contexts demonstrate that text is constructed across proficiency levels similarly. It is no surprise that a text is formulated by constructing and linking words. A tendency towards deletions across proficiency levels shows a requirement for additional research to understand the nature of cognitive difficulties in L2 Estonian writing.

4.2. Micro-contexts by duration

Table 2 below presents micro-contexts across proficiency levels with the mean of the mean transition time exceeding two seconds per text.

Table 2. Micro-contexts with the mean of the mean transition time per text exceeding two seconds across the proficiency levels

A2 (N = 4)		B1 (N = 13)		B2 (N = 13)		C1 (N = 4)	
Micro-context	Mean-mean transition time per text (s)	Micro-context	Mean-mean transition time per text (s)	Micro-context	Mean-mean transition time per text (s)	Micro-context	Mean-mean transition time per text (s)
.^c	7.69	.^c	8.19	c^a	11.45	.^c	10.52
_ ^D	6.16	_ ^D	6.13	._+^a	8.28	_ ^D	7.03
._+^a	3.90	c^a	5.84	.^c	7.61	.^+_a	4.94
D_+^a	3.40	._+^a	3.55	_ ^D	6.87	c^a	4.77
a^D	3.39	a^D	3.15	a^.	3.25	._+^a	3.77
D^a	2.54	a^c	2.70	_ ^_	3.19	a^D	2.48
a_+^a	2.34	D^a	2.54	a^D	3.13	D_+^a	2.47
a^,	2.14	a^,	2.47	.^+_a	2.87		
a^.	2.06	a^.	2.31	D^_	2.04		
		.^+_a	2.27				

Long periods of inactivity between keystrokes across all proficiency levels take place at the end of the paragraph. The mean transition to a new paragraph exceeds seven seconds at each level and the mean transition time is the longest at A2, B1

and C1 levels. Using solely keystroke logging without any additional data collection methods does not provide information on what is going on during these periods of inactivity. It can be assumed that the writer is either rereading what he or she has previously written or planning the following utterance.

The most time-consuming micro-context at B2 level is associated with starting a new paragraph. Likewise, B1 and C1 levels are marked by inactivity preceding the writing of a new paragraph, although the duration of that inactivity is only about half as long as at B2 level. Considering that both transition to a new paragraph and starting a new paragraph cause keystroke inactivity, the duration of inactivity between paragraphs is even longer (14.03 s at B1 level, 19.06 s at B2 level, 15.29 s at C1 level).

A mean period of inactivity of over six seconds across all proficiency levels is inherent to a transition from a space to a deletion. Considering that the preceding word has been finished and the writer has moved on to continue producing the text but ended up with a deletion, it can be assumed that in the given context the writer is rereading what he or she has previously written, revising or planning corrections.

Starting a new sentence is characterised by a mean period of inactivity exceeding three seconds across all proficiency levels and reaching more than eight seconds at B2 level. If we include the transition from a major delimiter at the end of the preceding sentence to the space followed by the first letter of the next sentence, the inter-sentence period of inactivity will be even longer (5.82 s at B1 level, 11.15 s at B2 level and 8.71 s at C1 level). Thus the longest micro-contexts at B2 level demonstrate that the biggest cognitive effort is made at the beginning of a new paragraph, followed by the beginning of a new sentence. Likewise, ending a sentence is characterised by a maximum mean period of inactivity specifically at B2 level.

All proficiency levels are characterised by periods of inactivity within a word between a letter and a delete/backstroke, although the duration of this inactivity decreases with each subsequent proficiency level. Conversely, the transition exceeds two seconds only at A2 and B1 levels, although at B2 level it is very close to the pause criterion.

The continuation of the text composition process after the previously written text has been edited is preceded by a mean period of inactivity of 3.40 s at A2 level and 2.47 s at C1 level. Unlike the more proficient levels, A2 level is also characterised by a mean period of inactivity exceeding two seconds between words.

To sum up, the longest transitions occur between paragraphs, sentences and in the micro-contexts which involve deletions. The full lists of micro-contexts with mean numbers of the mean transition time per text, the mean of the minimum and maximum transition time per text and the mean number of the given micro-contexts per text can be found in the appendices (Appendix 3, 5, 7, 9). Since not all texts included all the contexts, the tables also contain the percentage of texts in which various micro-contexts occurred.

4.3. Text production process across proficiency levels: similarities and differences

The text production process is analysed according to the text composition, text length, syntactic complexity, and deletions and/or mouse events data; this is followed by a discussion of potential pause locations.

Table 3 below summarises the mean numbers per text of micro-contexts and the mean of mean transition times across all proficiency levels describing the text production on the basis of the available data. All numbers represent the linear text production process and not the final text.

Table 3. Text production based on micro-contexts data across proficiency levels: similarities and differences

Micro-context	A2 (N = 4)		B1 (N = 13)		B2 (N = 13)		C1 (N = 4)	
	Mean N/ text	Mean-mean transition time per text (s)	Mean N/ text	Mean-mean transition time per text (s)	Mean N/ text	Mean-mean transition time per text (s)	Mean N/ text	Mean-mean transition time per text (s)
Text composition								
.^c	3	7.69	1	8.19	2	7.61	4	10.52
c^a	0	0.00	2	5.84	3	11.45	5	4.77
Text length								
a^a	938	0.49	1132	0.49	1299	0.34	1214	0.30
a_^a	180	2.34	212	1.58	220	1.46	208	1.47
a^_+a	180	0.55	212	0.50	220	0.48	208	0.50
a^.	27	2.06	19	2.31	17	3.25	14	1.13
._+^a	7	3.90	7	3.55	10	8.28	8	3.77
.^_+a	7	1.96	7	2.27	10	2.87	8	4.94
Syntactic complexity								
a^,	22	2.14	22	2.47	25	1.94	25	1.60
,^_+a	26	0.42	20	0.63	31	0.57	25	0.38
,_^+a	26	1.34	20	1.08	31	0.88	25	1.44
Deletions/mouse events								
D^D	156	1.12	333	0.51	340	0.97	391	0.49
a^D	41	3.39	52	3.15	62	3.13	56	2.48
D^a	43	2.54	56	2.54	69	1.99	59	1.45
a^_+D	10	1.87	17	0.68	20	0.72	14	0.74
_+^D	16	6.16	23	6.13	27	6.87	18	7.03
D^_	10	0.97	13	1.90	15	2.04	13	1.93
D_^+a	6	3.40	8	1.67	11	1.97	12	2.47

If we compare the number of paragraphs across the proficiency levels, a tendency towards growth is noted along with the improvement in proficiency level. However, the given numbers describe only those cases where a paragraph is made with a transition from an enter keystroke to a letter. Several participants divided their texts into

paragraphs by other means, for example by pressing several spacebars after the enter key before a letter. Checking the texts confirms that paragraphs are used also at A2 level, but these cases were not registered by ScriptLog with the given micro-context.

As regards text length, linear texts tend to become longer from A2 level to B2 level and then slightly decrease in length at C1 level. The mean number of sentences based on micro-contexts associated with closing a sentence decreases along with growing proficiency, being almost equal at B1 level and B2 level, but twice as big at A2 level (27) as at C1 level (14). It can be concluded that the mean number of sentences is bigger at A2 level, but the sentences are shorter. At C1 level, conversely, the texts are shorter in terms of the number of sentences, but the sentences are longer. The mean numbers of micro-contexts associated with beginning and closing a sentence differ substantially. One reason for this is that this process is subject to constant modification. Also, several participants did not use spaces between sentences, so these cases were left outside this micro-context.

In the case of syntactic complexity, minor delimiters are used on each level and at a similar frequency, even at levels at the extremes: A2 and C1.

The writing process across all proficiency levels is characterised by deletions and mouse movements. Consecutive deletions grow from A2 level to C1 level, with twice as many at C1 level as at A2 level. Even with increasing language proficiency, linear texts are subject to constant revision. Even deletions inside and between words gradually increase from A2 level to B2 level. Although C1 level is characterised by the maximum number of consecutive deletions and mouse events, B2 level is the level with the most correction and modification inside and between words.

Finally, let us compare the means of mean transition time per text in order to reveal which of the micro-contexts cause potential pause locations. The findings are in line with Åsa Wengelin (2006), Veerle M. Baaijen, David Galbraith and Kees de Glopper (2012), and Kristyan Spelman Miller (2000), confirming that higher level text units are subject to longer pauses and pause length increases for higher-level locations.

Micro-contexts between paragraphs cause the longest inactivity and are subject to extensive pauses across all proficiency levels. So too do micro-contexts before sentences across all levels. Closing a sentence is accompanied with pauses from A2 level to B2 level. Pauses are more often made before sentences than at the end. Micro-contexts before words require more transition time than within and after words. That said, their mean transition times generally do not exceed two seconds, except at A2 level before words.

Pauses are made before minor delimiters at A2 level and B1 level. Pauses are longer before minor delimiters across all levels than pauses before major delimiters, except at B2 level. Transitions to minor delimiters are longer across all proficiency levels than transitions right after minor delimiters.

In the case of deletions and mouse events, the most time-consuming pause locations across all proficiency levels are between words before beginning editing after the preceding word has been finished and the spacebar pressed. The time is required to look back to a previously written text to plan a modification. The second time-consuming pause location is inside words before a deletion across all proficiency levels, and after a deletion at A2 and B1 levels. Transitions to a new word after a deletion and a spacebar cause pauses at A2 and C1 levels. Consecutive deletions are not connected with pauses.

5. Conclusion

The study presented in this article aimed to examine what an analysis of micro-contexts can tell us about text production in the L2 Estonian of L1 Russian students across proficiency levels from A2 to C1.

The writing process across all levels is not a smooth process and a lot of deletions are made during the writing session. The longest transitions in micro-contexts reveal that the most cognitive effort is made between paragraphs and sentences and when deletions are involved.

The study shows that along with proficiency development, the number of paragraphs in a text increases. Linear texts tend to become longer from A2 to B2 level. However, the mean number of sentences decreases along with increasing proficiency. The use of minor delimiters across proficiency levels is quite similar.

The growing amount of consecutive deletions demonstrates that even with developing proficiency, linear production text is subject to constant revision, correction and modification. Deletions inside and between words gradually increase from the A2 level to the B2 level. Although the C1 level is characterised by a maximum of consecutive deletions and mouse events, B2 is the level with the most correction and modification inside and between words.

However, this study involved a number of limitations. The cases in which participants did not follow the writing conventions were not included in the study. For example, some participants wrote partly without spaces between sentences or made a space before a minor or major delimiter. Still, using micro-contexts to describe text production provides valuable data which cannot be collected in any other way.

The results of the study do not lend themselves to generalizations but rather show tendencies in how text is formulated and transcribed. The study highlights the need for future research on writing in Estonian as an L2. One possible direction for future research concerns the allocation of time during writing. Since writing is characterised by making a lot of corrections, text deletions should be looked at more closely. The third option could be combining keystroke logging and retrospective methods to investigate what is occurring during inactivity in writing and what exactly makes the writing process complicated across proficiency levels.

Symbols

^	transition from a given keystroke to a subsequent keystroke
a	character in a word (either upper or lower case) or a digit
—	spacebar
c	enter keystroke
D	delete/backspace keystroke or a mouse event
.	major delimiter, i.e. a full stop, a question mark or an exclamation mark
,	minor delimiter, i.e. a comma, a colon, a semicolon, a dash or brackets
+	one or more of the same character as the preceding one

References

- Asker-Árnason, Lena; Ibertsson, Tina; Wass, Malin; Wegelin, Åsa; Sahlén, Birgitta 2010. Picture-elicited written narratives, process and product, in 18 children with cochlear implants. – *Communication Disorders Quarterly*, 31 (4), 195–212. <http://dx.doi.org/10.1177/1525740109337734>

- Baaijen, Veerle M.; Galbraith, David; Glopper, Kees de 2012. Keystroke analysis: Reflections on procedures and measures. – *Written Communication*, 29 (3), 246–277. <http://dx.doi.org/10.1177/0741088312451108>
- CEFR = Common European Framework of Reference for Languages: Learning, Teaching, Assessment. http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf (9.2.2015).
- Chukharev-Hudilainen, Evgeny 2014. Pauses in spontaneous written communication: A keystroke logging study. – *Journal of Writing Research*, 6 (1), 61–84.
- Dimitrova, Birgitta Englund 2006. Segmentation of the writing process in translation: Experts versus novices. – Kirk P. H. Sullivan, Eva Lindgren (Eds.). *Computer Keystroke Logging and Writing: Methods and Applications*. Oxford: Elsevier, 189–201.
- Hayes, John R. 2012. Modeling and remodeling writing. – *Written Communication*, 29 (3), 369–388. <http://dx.doi.org/10.1177/0741088312451260>
- Jakobsen, Arnt Lykke 2006. Research methods in translation – *Translog*. – Kirk P. H. Sullivan, Eva Lindgren (Eds.). *Computer Keystroke Logging and Writing: Methods and Applications*. Oxford: Elsevier, 95–105.
- Johansson, Victoria 2009. *Developmental Aspects of Text Production in Writing and Speech*. Travaux de l'institut de linguistique de Lund 48. Lund: Lund University.
- Leijten, Mariëlle; Waes, Luuk van 2013. Keystroke logging in writing research: Using inputlog to analyze and visualize writing processes. – *Written Communication*, 30 (3), 358–392. <http://dx.doi.org/10.1177/0741088313491692>
- Lindgren, Eva 2005. *Writing and Revising. Didactic and Methodological Implications of Keystroke Writing*. *Skrifter från moderna språk* 18. Umeå: Umeå universitet.
- Nottbusch, Guido 2010. Grammatical planning, execution, and control in written sentence production. – *Reading and Writing*, 23 (7), 777–801. <http://dx.doi.org/10.1007/s11145-009-9188-4>
- Palviainen, Åsa; Kalaja, Paula; Mäntylä, Katja 2012. Development of L2 writing: Fluency and proficiency. – L. Meriläinen, L. Kolehmainen, T. Nieminen (Eds.). *AFinLA-e Soveltavan Kielitieteen Tutkimuksia*, 4, 47–59.
- Pastuhhova, Olga 2011. Kirjaliku produtseerimisprotsessi uurimise võimalused programmi *ScriptLog* abil: juhtumiuuring. [Research on the writing process using *ScriptLog*: A case study.] – *Lähivõrdlusi. Lähivertailuja* 21, 185–212. <http://dx.doi.org/10.5128/LV21.08>
- Spelman Miller, Kristyan 2000. Academic writers on-line: Investigating pausing in the production of text. – *Language Teaching Research*, 4 (2), 123–148. <http://dx.doi.org/10.1177/136216880000400203>
- Spelman Miller, Kristyan 2006a. The pausological study of written language production. – Kirk P. H. Sullivan, Eva Lindgren (Eds.). *Computer Keystroke Logging and Writing: Methods and Applications*. Oxford: Elsevier, 11–30.
- Spelman Miller, Kristyan 2006b. Pausing, productivity and the processing of topic in online writing. – Kirk P. H. Sullivan, Eva Lindgren (Eds.). *Computer Keystroke Logging and Writing: Methods and Applications*. Oxford: Elsevier, 131–156.
- Stevenson, Marie; Schoonen, Rob; Glopper, Kees de 2006. Revising in two languages: A multi-dimensional comparison of online writing revisions in L1 and FL. – *Journal of Second Language Writing*, 15 (3), 201–233. <http://dx.doi.org/10.1016/j.jslw.2006.06.002>
- Strömquist, Sven; Karlsson, Henrik 2002. *ScriptLog for Windows. User's Manual*. University of Lund: Department of Linguistics, University College of Stavanger: Centre for Reading Research.
- Strömquist, Sven; Holmqvist, Kenneth; Johansson, Victoria; Karlsson, Henrik; Wengelin Åsa 2006. What keystroke logging can reveal about writing. – Kirk P. H. Sullivan, Eva Lindgren (Eds.). *Computer Keystroke Logging and Writing: Methods and Applications*. Oxford: Elsevier, 45–71.

- Sullivan, Kirk P. H.; Lindgren, Eva (Eds.) 2006. Computer Keystroke Logging and Writing: Methods and Applications. Studies in Writing 18. Oxford: Elsevier.
- Uppstad, Per Henning; Solheim, Oddny Judith 2007. Aspects of fluency in writing. – Journal of Psycholinguistics Research, 36 (2), 79–87. <http://dx.doi.org/10.1007/s10936-006-9034-7>
- Waes, Luuk van; Leijten, Mariëlle; Weijen, Daphne van 2009. Keystroke logging in writing research. Observing writing processes with Inputlog. – GFL Journal, 2–3, 41–64.
- Waes, Luuk van; Leijten, Mariëlle; Wengelin, Åsa; Lindgren, Eva 2011. Logging tools to study digital writing processes. – Virginia Wise Berninger (Ed.). Past, Present, and Future Contributions of Cognitive Writing Research to Cognitive Psychology. New York: Psychology Press, 507–534.
- Wengelin, Åsa 2006. Examining pauses in writing: Theory, methods and empirical data. – Kirk P. H. Sullivan, Eva Lindgren (Eds.). Computer Keystroke Logging and Writing: Methods and Applications. Oxford: Elsevier, 107–130.
- Wengelin, Åsa; Strömquist, Sven 2004. Text-writing development viewed through on-line pausing in Swedish. – R. Berman (Ed.). Language Development Across Childhood and Adolescence. Trends in Language Acquisition Research 3. Philadelphia: John Benjamins, 177–190. <http://dx.doi.org/10.1075/tilar.3.12wen>

Olga Pastuhhova's (Tallinn University) main research interests are acquisition of Estonian as second language and writing process.
Narva mnt 25, 10120 Tallinn, Estonia
olgap@tlu.ee

Appendix 1. Types of micro-contexts recordable by ScriptLog and used in the study (Strömqvist, Karlsson 2002, Wengelin 2006: 114)

Notation	Description
a^a	Character, TRANSITION, character Inactivity between two subsequent letters (within a word)
a_+^a	Character, space, TRANSITION, character Inactivity before a letter and after a spacebar preceded by a letter
a^_+a	Character, TRANSITION, space, character Inactivity after the last letter of a word and before a spacebar followed by a letter
._+^a	Major delimiter, space, TRANSITION, character Inactivity before a letter and after a major delimiter followed by a spacebar
a^.	Character, TRANSITION, major delimiter Inactivity after the last letter of a word and before a major delimiter
.^c	Major delimiter, TRANSITION, enter Inactivity after a major delimiter and before an enter keystroke
c^a	Enter, TRANSITION, character Inactivity after an enter keystroke and before a letter
,^+a	Minor delimiter, TRANSITION, space, character Inactivity after a minor delimiter and before a spacebar followed by a letter
,_+^a	Minor delimiter, space, TRANSITION, character Inactivity before a letter and after a minor delimiter followed by a spacebar
.^+a	Major delimiter, TRANSITION, space, character Inactivity after a major delimiter and before a spacebar followed by a letter
D^D	Delete/mouse event, TRANSITION, delete/mouse event Inactivity between two delete/backstrokes
D^_	Delete/mouse event, TRANSITION, space Inactivity between a delete/backstroke and a spacebar
D^a	Delete/mouse event, TRANSITION, character Inactivity between a delete/backstroke and a letter
D_+^a	Delete/mouse event, space, TRANSITION, character Inactivity before a letter and after a delete/backstroke followed by a spacebar
_ ^D	Space, TRANSITION, delete/mouse event Inactivity between a spacebar and a delete/backstroke
a^D	Character, TRANSITION, delete/mouse event Inactivity between a letter and a delete/backstroke
a^_+D	Character, TRANSITION, space, delete/mouse event Inactivity after a letter and before a spacebar followed by a delete/backstroke
a^,	Character, TRANSITION, minor delimiter Inactivity after a letter and before a minor delimiter
_ ^ _	Space, TRANSITION, space Inactivity between two spacebars
a^c	Character, TRANSITION, enter Inactivity after a letter and before an enter keystroke

Appendix 2. Full list of micro-contexts at A2 level according to the frequency of occurrence

Context	Number	Mean N/text	Mean-mean transition time per text (s)
a^a	3752	938	0.49
a^_+a	720	180	0.55
a_+^a	720	180	2.34
D^D	625	156	1.12
D^a	173	43	2.54
a^D	162	41	3.39
a^.	108	27	2.06
a^,	86	22	2.14
,^_+a	77	26	0.42
,_+^a	77	26	1.34
_ ^D	65	16	6.16
a^_+D	39	10	1.87
D^_	38	10	0.97
D_+^a	23	6	3.40
.^_+a	22	7	1.96
._+^a	22	7	3.90
_ ^_	12	6	1.28
.^c	6	3	7.69

Appendix 3. Full list of micro-contexts at A2 level according to the duration

Micro-context	Mean of the mean transition time per text (s)	Mean of the min transition per text (s)	Mean of the max transition per text (s)	% of texts in which micro-context occurred	Mean N/text
.^c	7.69	6.05	8.81	50	3
_ ^D	6.16	0.96	36.61	100	16
._+^a	3.90	1.29	9.14	75	7
D_+^a	3.40	0.79	12.94	100	6
a^D	3.39	0.52	35.78	100	41
D^a	2.54	0.32	17.20	100	43
a_+^a	2.34	0.18	40.23	100	180
a^,	2.14	0.46	7.62	100	22
a^.	2.06	0.52	10.24	100	27
.^_+a	1.96	1.78	2.50	75	7
a^_+D	1.87	0.15	5.96	100	10
,_+^a	1.34	0.28	7.94	75	26
_ ^_	1.28	1.21	1.34	50	6
D^D	1.12	0.11	24.51	100	156
D^_	0.97	0.57	2.01	100	10
a^_+a	0.55	0.12	13.26	100	180
a^a	0.49	0.09	14.91	100	938
,^_+a	0.42	0.23	0.89	75	26

Appendix 4. Full list of micro-contexts at B1 level according to the frequency of occurrence

Context	Number	Mean N/text	Mean-mean transition time per text (s)
a^a	14713	1132	0.49
D^D	4324	333	0.51
a^_+a	2752	212	0.50
a_+^a	2752	212	1.58
D^a	727	56	2.54
a^D	677	52	3.15
_ ^D	303	23	6.13
a^,	281	22	2.47
a^.	240	19	2.31
a^_+D	219	17	0.68
,^_+a	215	20	0.63
,_+^a	215	20	1.08
D^_	169	13	1.90
D_+^a	101	8	1.67
,^_+a	72	7	2.27
,_+^a	72	7	3.55
_ ^_	48	6	0.57
c^a	13	2	5.84
,^c	11	1	8.19
a^c	2	1	2.70

Appendix 5. Full list of micro-contexts at B1 level according to the duration

Micro-context	Mean of the mean transition time per text (s)	Mean of the median transition time per text (s)	Mean of the min transition per text (s)	Mean of the max transition per text (s)	% of texts in which micro-context occurred	Mean N/text
,^c	8.19	7.69	7.55	9.34	69	1
_ ^D	6.13	2.95	0.64	31.72	100	23
c^a	5.84	5.01	3.08	8.41	54	2
,_+^a	3.55	1.31	0.92	10.22	77	7
a^D	3.15	1.32	0.41	31.07	100	52
a^c	2.70	2.70	2.70	2.70	15	1
D^a	2.54	0.82	0.19	20.92	100	56
a^,	2.47	1.35	0.64	14.94	100	22
a^.	2.31	1.54	0.53	8.94	100	19
,^_+a	2.27	0.84	0.69	5.42	77	7
D^_	1.90	0.67	0.29	10.86	100	13
D_+^a	1.67	0.49	0.30	6.44	100	8
a_+^a	1.58	0.69	0.16	24.74	100	212
,_+^a	1.08	0.56	0.22	7.79	85	20
a^_+D	0.68	0.39	0.16	3.35	100	17
,^_+a	0.63	0.44	0.29	1.96	85	20
_ ^_	0.57	0.60	0.44	0.72	62	6
D^D	0.51	0.17	0.05	21.24	100	333
a^_+a	0.50	0.26	0.09	10.75	100	212
a^a	0.49	0.33	0.05	8.40	100	1132

Appendix 6. Full list of micro-contexts at B2 level according to the frequency of occurrence

Context	Number	Mean N/text	Mean-mean transition time per text (s)
a^a	16887	1299	0.34
D^D	4415	340	0.97
a^_+a	2857	220	0.48
a_+^a	2857	220	1.46
D^a	899	69	1.99
a^D	807	62	3.13
,^_+a	397	31	0.57
,_+^a	397	31	0.88
_ ^D	356	27	6.87
a^,	326	25	1.94
a^_+D	258	20	0.72
a^.	224	17	3.25
_ ^_	202	22	3.19
D^_	196	15	2.04
D_+^a	145	11	1.97
.^_+a	128	10	2.87
,_+^a	128	10	8.28
c^a	29	3	11.45
.^c	23	2	7.61
a^c	1	1	0.78

Appendix 7. Full list of micro-contexts at B2 level according to the duration

Micro-context	Mean of the mean transition time per text (s)	Mean of the median transition time per text (s)	Mean of the min transition per text (s)	Mean of the max transition per text (s)	% of texts in which micro-context occurred	Mean N/text
c^a	11.45	4.91	1.06	27.96	68%	3
,_+^a	8.28	4.79	0.93	27.85	100%	10
.^c	7.61	5.99	4.22	11.45	92%	2
_ ^D	6.87	2.29	0.47	43.69	100%	27
a^.	3.25	0.94	0.28	17.47	100%	17
_ ^_	3.19	1.34	1.30	5.77	69%	22
a^D	3.13	1.15	0.39	33.70	100%	62
.^_+a	2.87	0.91	0.37	11.66	100%	10
D^_	2.04	0.73	0.24	15.75	100%	15
D^a	1.99	0.75	0.17	29.42	100%	69
D_+^a	1.97	0.53	0.20	10.51	100%	11
a^,	1.94	0.95	0.27	10.23	100%	25
a_+^a	1.46	0.52	0.10	37.48	100%	220
D^D	0.97	0.16	0.03	44.01	100%	340
,_+^a	0.88	0.40	0.15	7.01	100%	31
a^c	0.78	0.78	0.78	0.78	8%	1
a^_+D	0.72	0.24	0.11	7.95	100%	20
,^_+a	0.57	0.28	0.15	4.85	100%	31
a^_+a	0.48	0.20	0.07	14.26	100%	220
a^a	0.34	0.24	0.05	7.84	100%	1299

Appendix 8. Full list of micro-contexts at C1 level according to the frequency of occurrence

Context	Number	Mean N/text	Mean-mean transition time per text (s)
a^a	4854	1214	0.30
D^D	1565	391	0.49
a^_+a	833	208	0.50
a_+^a	833	208	1.47
D^a	237	59	1.45
a^D	225	56	2.48
,^_+a	101	25	0.38
,_+^a	101	25	1.44
a^,	98	25	1.60
_ ^D	73	18	7.03
a^_+D	57	14	0.74
a^.	55	14	1.13
D^_	52	13	1.93
D_+^a	47	12	2.47
.^_+a	30	8	4.94
._+^a	30	8	3.77
c^a	15	5	4.77
.^c	8	4	10.52

Appendix 9. Full list of micro-contexts at C1 level according to the duration

Micro-context	Mean of the mean transition time per text (s)	Mean of the median transition time per text (s)	Mean of the min transition per text (s)	Mean of the max transition per text (s)	% of texts in which micro-context occurred	Mean N/text
.^c	10.52	10.02	4.10	17.58	50	4
_ ^D	7.03	1.83	0.27	48.22	100	18
.^_+a	4.94	3.39	0.22	11.69	100	8
c^a	4.77	2.91	0.72	12.31	75	5
._+^a	3.77	0.95	0.52	17.43	100	8
a^D	2.48	0.94	0.27	35.61	100	56
D_+^a	2.47	0.67	0.18	15.71	100	12
D^_	1.93	0.65	0.27	10.19	100	13
a^,	1.60	0.64	0.18	7.79	100	25
a_+^a	1.47	0.42	0.12	55.02	100	208
D^a	1.45	0.47	0.12	27.39	100	59
,_+^a	1.44	0.45	0.18	15.48	100	25
a^.	1.13	0.47	0.15	6.83	100	14
a^_+D	0.74	0.34	0.10	3.47	100	14
a^_+a	0.50	0.20	0.06	14.47	100	208
D^D	0.49	0.13	0.03	33.29	100	391
,^_+a	0.38	0.22	0.14	2.45	100	25
a^a	0.30	0.23	0.05	4.08	100	1214

MIKROKONTEKSTID EESTI KEELE KUI TEISE KEELE KIRJUTAMISPROTSESSI KIRJELDAMISEL KEELEOSKUSTASEMETI

Olga Pastuhhova

Tallinna Ülikool

Artikkel kirjeldab venekeelsete eesti keele õppijate kirjutamisprotsessi mikrokonteksti mõiste abil. Uurimismaterjal on kogutud klaviatuuri klahvivajutuste salvestuse meetodil põhineva programmiga ScriptLog. Uuringus osales 34 Tallinna Ülikooli venekeelset üliõpilast, kellel paluti kirjutada arvutis loovkirjutis haridusteemal. Loovkirjutisi hindasid kaks eksperti, kelle hinnangul vastas osalejate eesti keele oskus antud ülesande põhjal “Euroopa keeleõppe raamdokumendi” neljale keeleoskustasemele järgnevalt: 4 teksti vastasid tasemele A2, 13 teksti tasemele B1, 13 teksti tasemele B2 ja 4 teksti tasemele C1.

Mikrokonteksti defineeritakse kui üleminekut ühelt klahvilt teisele teatud ümbruses (Wengelin 2006). Artiklis esitatakse mikrokontekstide analüüs nende esinemissageduse ja ülemineku aja põhjal. Seejärel antakse üldine ülevaade kirjutamisprotsessist ja potentsiaalsetest peatumiskohtadest. Mikrokontekstide andmeid on esitatud ja võrreldud keeleoskustasemeti.

Tulemused näitavad, et kirjutamisprotsess ei ole sujuv ja seda iseloomustab suur hulk parandusi. Pikimad klahvidevahelised üleminekud esinevad lõikude vahel, mikrokontekstides, mis on seotud uue lause algusega ja kustutustega. Järjestikuste kustutamiste ja/või hiireliigutuste arvu kasv näitab, et kirjutamisprotsessi iseloomustab ka keeleoskuse kasvades teksti pidev redigeerimine, parandamine, muutmine. Kustutamised sõna sees ja sõnade vahel suurenevad tasemest A2 kuni tasemeni B2. Kuigi uurimismaterjali põhjal saab järeldada, et tasemel C1 toimub kõige rohkem järjestikuseid kustutamisi, iseloomustab suurim teksti parandamine sõna sees ja sõnade vahel taset B2.

Võtmesõnad: klahvivajutus, ScriptLog, eesti keel, vene keel