A Multilingual Copy Task: Measuring Typing and Motor Skills in Writing with Inputlog

Luuk Van Waes, Mariëlle Leijten, Tom Pauwaert and Eric Van Horenbeeck

Department of Management, University of Antwerp, BE
Corresponding author: Luuk Van Waes (luuk.vanwaes@uantwerpen.be)

Keyboarding is an essential mode of text production. In the context of typing courses, instruments have been developed to gauge typing skills but to our knowledge there is no computerized typing test measuring both motor skills and the influence of different levels of lexicality on typing fluency. That is why, we developed a strictly controlled copy task guiding participants through seven modules in which different prompts are presented, each dealing with complementary levels of lexicality. Fine-grained keystroke logging allows for a range of analyses with Inputlog [16].

The copy task can be used in writing process studies: research shows that typing fluency has an impact in the quality of text production [1, 24]. In previous writing research studies copy tasks have been used to study the relation between writing fluency and text composition (Wallot and Grabowski, 2013). The Inputlog copy task allows researchers to investigate different levels of lexicality in more detail. At the moment the copy task has been developed in ten different languages. The software is open-access and allows researchers to adapt the tasks to their specific needs.

Keywords: copy tasks; typing skills; motor coordination; keyboarding; writing studies; interkey intervals; bigram frequency; tapping task; typing fluency

Funding statement: We thank the department of Management, Faculty of Business and Economics at the University of Antwerp (Belgium) for their financial contribution to this project. We also received support of ELN (European Literacy Network – COST Action IS1401).

Introduction

Until recently typing dynamics have mainly been analyzed in the domain of biometric studies [2, 4, 21]. After the introduction of keystroke logging in writing and translation studies [16, 23, 27], the analysis of process based typing skills became more important, for instance in studying writing development [5] or writing fluency [26]. Keystroke logging is especially useful to identify and understand strategies governing the dynamics of writing [12, 13]. These analyses mainly focus on dysfluencies in writing [17, 19], such as pauses and revisions.

It is shown that typing skills are a factor affecting online text production when generating digital texts [1, 6, 24, 31, 29], comparable to the role of handwriting in handwritten text [7, 10, 11, 15]. The advantage of copy tasks is that they produce writing conditions where the cognitive load is low, eliminating to a large extent higher mental processes such as content planning [9]. The current copy task is designed so that it creates a set of measures allowing a fine-grained analysis of low level typing and motor skills.

A first version of this copy task was developed for research into the typing and writing behavior of young adults, cognitively healthy elderly, and age-matched elderly with mild cognitive impairment (MCI) or mild dementia due to Alzheimer [28]. The task consisted of different subtasks in which specific bigram characteristics were manipulated. This explorative study demonstrates the potential relevance of using a typing copy task in the diagnostic workup of patients with neurodegenerative brain disorders. It was found that the three groups differed significantly from each other in performing the assignment and that their typing speed gradually decreased with age. Moreover, at the methodological level, integrative multilevel modelling shows that all the manipulated bigram characteristics contribute significantly to the model. Based on this experience we further developed the instrument and programmed a user friendly open experimental tool set, facilitating the use of copy tasks in a variety of (multilingual) writing research settings.

Components

The multilingual copy task was developed as part of Inputlog, a keystroke logging program used in writing and translation studies, made available for researchers on www.inputlog.net ([16]; see the download section for more information on the installation procedure). However, the copy task is also directly accessible as a
webtool via: http://inputlog.ua.ac.be/WebSite/copytask/tasks.html. The source code is downloadable via GitHub (see below).

At this moment the copy task has been developed for ten languages (Dutch, English, French, German, Norwegian, Polish, Portuguese, Spanish, Turkish, and Welsh) and three keyboard layouts (Azerty, Qwerty, and Qwertz). The software has been pretested worldwide in a number of experimental studies in ongoing writing and translation research. More than 3000 copy task logs with participants aged between 14 and 82 years have been collected and analyzed. A typical copy task consists of seven typing assignments (Table 1).

**Tapping task**
The first task, the tapping task, intends to measure the fastest motor speed by pressing two keys with alternating hands (viz. ‘d’ and ‘k’, resp. a Left-Right and Right-Left hand combination). Participants are asked to type the ‘d’-‘k’ key combination for 15 seconds [22]. A time circle at the top right corner is used as a time indicator. The time starts after the first key press, allowing the participants to read the instructions at their own pace.

Finger-tapping tasks are commonly used to study the human motor system. They have the advantage of being simple enough to use in the study of both normal control subjects as well as those with neuropathologies affecting the motor system [32].

**Sentence copy task**
The next task is a sentence copy task. It intends to measure the typing skills related to copying a series of – short and high frequent – words presented in a simple sentence. For example, in the English copy task: "the cat was sleeping under the apple tree". Participants are asked to repetitively type this sentence for 30 seconds (without capital letters nor punctuation marks).

The prompted sentence consists of seven to nine high frequent short words (one or two syllables). In developing the task for the different languages, we tried to implement these construction characteristics as consistently as possible to allow for interlanguage comparison (Mean ($M$): 7.9 words; Standard Deviation ($SD$): 0.8 – $M$: 32.3 characters – $SD$: 2.8).

**Three-word combination**
The main part of the copy task consists of four components in which a three-word combination has to be copy-typed seven times: the first three-word combinations target the repetitive production of mainly high frequent bigrams; the fourth contains low frequent bigrams.

Table 2 gives an actual example of the presented word combinations (in the English version), together with their characteristics.

### Table 1: Overview of the seven components of the copy task.

<table>
<thead>
<tr>
<th>Components</th>
<th>Tapping task</th>
<th>Sentence</th>
<th>Word combination 1</th>
<th>Word combination 2</th>
<th>Word combination 3</th>
<th>Word combination 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>press the ‘d’ and ‘k’ key alternatively during 15 s</td>
<td>copy a sentence during 30 s</td>
<td>copy a combination of three words seven times</td>
<td>copy a combination of three words seven times</td>
<td>copy a combination of three words seven times</td>
<td>copy a combination of three words seven times</td>
</tr>
<tr>
<td>Consonant groups</td>
<td>copy four blocks of six consonants once</td>
<td>copy four blocks of six consonants once</td>
<td>copy four blocks of six consonants once</td>
<td>copy four blocks of six consonants once</td>
<td>copy four blocks of six consonants once</td>
<td>copy four blocks of six consonants once</td>
</tr>
</tbody>
</table>

### Table 2: Examples of the word combination prompts in the English_UK (QWERTY) copy task including the controlled characteristics.

<table>
<thead>
<tr>
<th></th>
<th>Word combination 1</th>
<th>Word combination 2</th>
<th>Word combination 3</th>
<th>Word combination 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word 1 (numerical)</td>
<td>four</td>
<td>seven</td>
<td>five</td>
<td>some</td>
</tr>
<tr>
<td>Word 2 (adjective)</td>
<td>interesting</td>
<td>wonderful</td>
<td>important</td>
<td>awkward</td>
</tr>
<tr>
<td>Word 3 (noun)</td>
<td>questions</td>
<td>surprises</td>
<td>behaviors</td>
<td>zigzags</td>
</tr>
<tr>
<td>#characters</td>
<td>24</td>
<td>23</td>
<td>23</td>
<td>18</td>
</tr>
<tr>
<td>High Frequent bigrams (HF – e.g., ‘nt’)</td>
<td>19</td>
<td>18</td>
<td>18</td>
<td>8</td>
</tr>
<tr>
<td>Low Frequent bigrams (LF – e.g., ‘gz’)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Left-Left (LL – e.g., ‘es’)</td>
<td>4</td>
<td>6</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Left-Right (LR – e.g., ‘fo’)</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Right-Right (RR – e.g., ‘bu’)</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Right-Left (RL – e.g., ‘us’)</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Adjacent keys (e.g., ‘io’)</td>
<td>7</td>
<td>6</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Repetitive keys</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
To create the word combinations with high frequent bigrams, we used the following criteria:

1. **Words**: Three words are presented per entry, i.e. a combination of a numeral + an adjective + a noun;
2. **Characters**: Total prompt length is between 19 and 24 characters in total – [3/4] + [10/12] + [8/10] characters (average number of characters per word combination for all the language tasks: M: 22.3 (SD: 1.6);
3. **Word frequency**: Only high frequency words are included (30% highest segment; log lemma frequency of the words or the stem words in case of compound words; based on CELEX/Subtlex [3, 14];
4. **Bigram frequency**: Between 18 to 20 high frequency bigrams are implemented selected from the 30 % most frequent percentile in CELEX/Subtlex [3, 14] or another comparable corpus-based bigram frequency list for languages for which these databases were not available. No low frequent bigrams are included in these three word combinations;
5. **Hand combination and key adjacency**: A mix of hand combinations is addressed with an average of three to five bigrams for each hand combination (Left-Left; Left-Right; Right-Left; Right-Right); five to seven keyboard adjacent bigrams and none (or maximum one) repetitive keys (**Figure 1**).

For the component with *low frequent bigrams* (Word combination 4) we used the same criteria, but instead of selecting high frequency bigrams in the words, low frequency bigrams were presented, e.g. three to four bigrams occurring in the 50% lowest frequency percentiles of our corpus. The total length of the prompt across all languages was slightly shorter than in the high frequency components: average number of characters per word combination: M 19.0 (SD: 1.3).

### Consonant groups

The final component is designed to measure the typing skills in a non-lexical context [9]. Participants are asked to copy four blocks of six consonants once. The blocks are identical for all languages: tjxggl pgkfkq dtdrtt npwdvf.

**Procedure**

The copy task is presented to the participant as one fixed flow in which he or she is taken from one component to the next (**Figure 2**).

---

**Figure 1**: Keyboard scheme (e.g. QWERTY) used to define the four hand combinations: Blue = Left hand – Red = Right hand; midzone is characterized as undefined.

**Figure 2**: Presentation flow of the copy task.
In a first step the participant selects the language for the copy task. The next screen prompts him/her to briefly provide some session information: name/participant code, age and gender; session name (optional) and keyboard used (Azerty, Qwerty or Qwertz, visual help provided). We also included a privacy notice to obtain explicit consent, in line with the privacy policy of the General Data Protection Regulation (GDPR) of the European Union.

The copy task then starts with a brief introduction to instruct and inform the user about the task. For instance: “In this session we ask you to type letters, words and sentences that are shown on the screen. Try to do this as fast and error free as possible.” When pressing ‘start’ the components are shown one by one, always preceded by a brief – and specific – instruction. So as to allow the participants to read the task instructions at their own pace, we explicitly inform them that their time only starts after they have typed the first character.

The prompt needing to be copied is always presented separately and printed in dark red (Figure 3). When switching from the time-based components to the repetition-based components (word combinations that need to be retyped), a separate example screen is presented to further clarify this instruction.

Next, the flow brings the user to a brief closing questionnaire in which the following subjects are addressed: handedness, hardware and browser used for the test, dominant language, reading or writing difficulties, and familiarity with this task. For the handedness test we chose the reduced Edinburgh handedness test [18, 30]. Veale proved his assessment to be reliable with only four questions. Finally, the data are stored on a dedicated server. The participants have the option to download their log files locally.

Copy task creator

A default copy task – based on the principles described above – is made available in ten languages. However, if a researcher wants to translate/transpose the default copy task into another language or wants to construct an additional copy task, he/she can use the so called ‘copy task creator’, made available as an integrated tool in Inputlog but also available as an isolated, stand-alone tool (see Github). The copy task creator consists of several building blocks that can be combined into a custom-made task flow. The following components are defined in the interface (see Figure 4):

1. Tasks: A task component assists the user in defining a specific copy task component. It asks for a brief task
instruction, a user prompt and either a time limit or a predefined repetition number.

2. Examples: The example block builds a part of the copy task requiring no input, only an illustration of what the user is supposed to do in the next task.

3. Text blocks: A text block component providing feedback or feed forward information clarifying the task flow so as to guarantee a fluid transition between the different instructions.

4. Practice: The final option offers the possibility to add a ‘practice’ component. It looks like a normal task, however, typing activities are not included in the log file. This part gives the participant the possibility to get acquainted with the web interface of the copy task.

After having defined and constructed the flow, the copy task file can be uploaded and is ready to start. The opening and closing questions are default.

Copy task analysis
The copy task analysis as provided in Inputlog consists of a carefully designed set of perspectives to explore the log file. The aim is to characterize a participant’s motoric keyboard interaction based on the different copy task components. The output is presented such that it is possible to use the data in further analyses (e.g., as a co-variate in variance analyses, or in regression and multilevel analyses). Two types of output are generated: an aggregated analysis (synthesis) and an event-based data file (extended raw data) including an automated characterization of all the targeted bigrams together with a precise intra bigram time-stamp. Both are XML-files. The analysis of the copy task log files can be performed in batch in the standard Inputlog version 8.X. Automated merging provides a csv file at group level for large data collections enabling further statistical analyses (e.g., in SPSS or R). A visual representation of the task flow, together with a more detailed description, is included in the documentation folder on Github.

The analysis report (styled XML) is preceded by a brief summary of the session information, together with the answers to the initial and final questionnaire. Next, the report addresses the following bigram characteristics (cf. supra):

1. Correctness: Because the participants are challenged to find a trade-off between speed and accuracy, correctness scores are reported, both at the overall and the component level. These scores indicate the number of correctly typed bigrams as opposed to the number of typing mistakes during copying.

2. Synthesis: Results for the Interkey Intervals (IKI) overall and at the component and trial level, including time filtered results to take into account learning effects;

3. Frequency: Results for IKI’s for high and low frequent bigrams;

4. Adjacency: Results for the IKI’s for bigrams that are formed by combining adjacent and non-adjacent keys on the keyboard;

5. Repetition: Results for the IKI’s that are composed by pressing the same key or not;

6. Hand combination: Results for the IKI’s with respect to the hand combination used to type the bigram (location on the keyboard).

For each level within the copy task analysis, several descriptive measures are reported (see Figure 5 for an excerpt). For instance: ‘Count (targeted)’ reports the number of characters typed that correspond to a character that was prompted in one of the respective copy task components; ‘Mean IKI’ is the arithmetic mean (in milliseconds) of the targeted bigrams within the scope of the respective analyses. Characters per minute (CPM): calculation of (theoretical number of) characters per minute based on the mean IKI, specifically 60 000 (ms)/mean IKI (ms).

For a more detailed description (and an example of a full sample analysis output) we refer to the Github repository mentioned below.

Implementation and architecture
The copy task creator and the web execution environment have both been open sourced. Both components are simple applications of low architectural complexity.
written as a C# application and javascript/html single page web application respectively (see Github).

The task creator consists of a single C# WinForm that allows the user to build a custom copy task by adding their own sequence of specific tasks. Their task configuration is then exported to an XML file that contains the copy task definition.

The web application is a JavaScript single page app (web/tasks.html) that reads this copy task definition file and dynamically creates the different screens as defined in definition. Depending on the specified language of the copy task a different language resource file will be loaded to provide localized instructions. The fallback language for a missing resource translation is English. During the execution of the test all the relevant keyboard activities of the user are tracked by hooking into the keypress events of the input areas. Each incoming event is timed and stored in an event list until the very end when the user submits their task. At that point the event list is exported to the IDFX xml format which can be processed by the respective copy task analysis in Inputlog. This hook and the corresponding event serialization code can be found in the scripts/CopyTask.js file on the Github repository.

Quality control

The logging accuracy of the copy task is tested and compared to the Inputlog logging accuracy [20]. The time log accuracy has the same maxdiff of 8 ms [8], which is in line with other keystroke logging programs, like Scriptlog or CyWrite.

In a test-retest experiment the copy task reliability has been measured. Fifty participants completed the task twice (with a time lag of 4 months in between). The results show that the test results are highly consistent: Pearson $r = .932$ ($p < .001$ – overall Median interkey intervals: $M_{t1} = 116$ ms; $SD = 17$; $M_{t2} = 115$ ms; $SD = 19$). Additional analysis with GLM repeated measures Anova at the component level showed a strong main effect of the components, $(F (4,192) = 227.353; p < .001; \eta^2 = .826)$. Post-hoc Bonferroni also shows a significant effect at the individual component level, indicating that the different (main) components address different (lexical) characteristics of motor related typing skills.

Availability

Operating system

The copy task is web based and is tested on most common browsers (Chrome, Internet Explorer, Firefox, Safari and Opera – OSX/Windows). The copy task creator is only available for Windows (version 8 and higher).

Programming language

The copy task is developed in JavaScript (with JQuery 2.1.0) and is web based. The logfile is stored as an XML-file conforming to the IDFX-structure [25]. The file is event based and contains all key strokes with a time stamp (key down time for every key press) and an identification of the active component and of any trials within that component. In addition, the session identification and the answers to the final questionnaire are included as part of the XML-output file.

The copy task creator is written in C# (2015) and can be used as a standalone program. The output generates an XML file with a specific set of tags that are recognized by the copy task JavaScript. A language tag in the header activates the language specific resources for the web interface and the navigation instructions.

Dependencies

The web application contains all the javascript libraries it depends on in the web/scripts/folder. The copy task creator written in C# is built on top of the .NET Framework version 4.6.1.

List of contributors

We thank Sebastian Fierens and Tom Pauwaert for their help in programming the copy task.

Software location

The compiled ready-to-use software is hosted on the Inputlog website and is downloadable from there: http://www.inputlog.net (Inputlog version 8). This website hosts the main application as well as additional materials (manual; presentations; articles; videos). The software is free to use after completing a simple registration procedure.

The source code for the copy task itself and the copy task creator is hosted in the GitHub repository, mentioned below.

Archive and Code repository: GitHub

Name: Inputlog-Copy-Task
URL: https://github.com/lvanwaes/Inputlog-Copy-Task
License: MIT License
Publisher: Luuk van Waes & Mariëlle Leijten
Contact: luuk.vanwaes@uantwerp.be or marielle.leijten@uantwerp.be
Persistent identifier: DOI: https://doi.org/10.5281/zenodo.2908966
Project homepage: http://www.inputlog.net
Version published: 1.0.0
Date published: 30/05/2018

Language

The web-based copy task is currently available in ten languages. The task, the instructions and the interface are rendered in the selected language by the program (JavaScript). A language resource database with translations of the navigation, instruction and error messages has been developed for that purpose.

Reuse potential

The copy task creator makes it easy to develop a standardized copy task in other languages, and to adapt the existing tasks to specific needs. It is also possible to edit the JavaScript used. The modular concept allows the development of new components creating other copy task
functions such as audio-based or graphically prompted copying. The task instructions could be enhanced with video guidelines.

Support
At the moment we are building a Forum that is related to the Inputlog website (www.inputlog.net). This will accommodate easy interactions with the community.

To add a new language that is currently not supported in the default copy task, do not hesitate to contact the authors (see contact information).

Acknowledgements
We thank the following colleagues for their help in developing the different language versions: English: Lise Fontaine (Cardiff University, UK) and Mark Torrance (Nottingham University, UK); French: Thierry Olive (CNRS@University of Poitiers, France); German: Esther Breuer (University of Cologne, Germany); Norwegian: Anne Sætersdal Myklestad & Vibeke Ronneberg (Western Norway University of Applied Sciences & University of Stravanger, Norway); Polish: Olga Witzczak (Adam Mickiewicz University, Poland); Portuguese: Teresa Limpo (University of Porto, Portugal); Spanish: Anna Sala (University of Barcelona, Spain); Swedish: Asa Wengelin (University of Gothenburgh, Sweden) & Victoria Johansson (University of Lund, Sweden); Turkish: Gulay Tiryakioglu (University of Lyon 2, France); Welsh: Lise Fontaine (Cardiff University, UK).

Competing Interests
The authors have no competing interests to declare.

References


29. Van Weerdenburg, M, Tesselhof, M and Van der Meijden, H 2019 Touch-typing for better spelling and narrative-writing skills on the computer. 35: 143–152. DOI: https://doi.org/10.1111/jcal.12323

