Help documentation (in progress)

Draft 03 May 2019

Coordinators
Mariëlle Leijten & Luuk Van Waes

Technical coordinator
Eric Van Horenbeeck

This document has been produced with the assistance of: Robbe Block, Ming Chang, Milou de Smet, Tom Pauwaert, Alexander Prinsier, Joeri Rammelaere en Joris Roovers.

University of Antwerp
Prinsstraat 13
BE 2000 Antwerp
Belgium

www.inputlog.net

Disclaimer

Self publishing

(marielle.leijten@uantwerpen.be, luuk.vanwaes@uantwerpen.be, eric.vanhorenbeeck@uantwerpen.be

Contents

Introduction to Inputlog (general description)

The program enables researchers to precisely register and accurately reconstruct the writing processes of writers who compose texts at the computer. In particular, Inputlog allows researchers to:

- **Record:** This module logs (keyboard, mouse and speech) events in Microsoft Word and other Windows based programs together with a unique time stamp (ms). Moreover, in MS Word this module also logs character position, actual document length and copy/paste/move actions. During the writing process these basic data are continuously and unobtrusively stored for later processing. This continuous data storage does not delay or interfere with the normal use of the computer.
- **Pre-process:** It is often necessary to refine logged data prior to analysis. This module allows us to process data from various perspectives:
  - event based (keyboard, mouse and speech),
  - time based or based on window changes (sources: MS Word, Internet etc.).

    Example. The Filter is also convenient for isolating different writing tasks logged in one session or for deleting logging session start-up or deactivation 'noise'. For example, when additional questions are asked in the beginning of the observation and the logging session has started already, this pausing time (noise) can be excluded from the data analyses.
- **Analyze**: This module is the heart of the program and features three process representations (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 is produced. The general logging file and the aggregated analyses will be discussed in more detail at the end of this section.
- **Post-process:** This module integrates single or multiple log files from Inputlog or other observation tools (Morae, Dragon Naturally Speaking, and Eye tracking data). It is also possible to merge multiple output files for further analysis in, for instance, SPSS or MLWin.
- **Play:** This module allows researchers to play back the recorded session at various levels (time or revision based). The replay is data based (not video based) and the play speed is adjustable. A logged session can also be reconstructed revision by revision.

Inputlog provides data for research on:

- cognitive writing processes
- writing strategies of novice and expert writers
- writing development of children with and without writing difficulties
- professional and creative writing
- first and second language writing
- spelling research
- specialist skill areas such as translation and (live-)subtitling
- biometric measures

Not only can keystroke logging be used in research specifically on writing, it can also be integrated in educational domains for second language learning, programming skills, and typing skills.

Philosophy Inputlog

Inputlog is a software program that is available for free for research purposes. The software has been developed gradually with very limited resources. Whenever there was money available, we tried to adapt and upgrade the program to the growing needs of the writing research communities. Many people directly and indirectly contributed to the program.

Please note that Inputlog is a research instrument that is free to the research community. If you publish research in which inputlog has been used please refer to:

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

We strongly believe in sharing information like research techniques etc. Therefore, we would like to strongly encourage you to share information on your research project via WritingPro. WritingPro is a knowledge center for writing process research where researcher can share information and search for information related to writing process research; project descriptions of other researchers, data collection techniques, data analyses techniques, references, presentations, ...

History

The development of Inputlog started in 2003. Because the existing keystroke logging programs at that time could not be used to log the writing processes of (professional) writers creating their texts in a commercial word processor (e.g., MS Word), we decided to build a new keystroke logger. Another impetus for the development of Inputlog has been the emergence of speech recognition as a new writing mode. Therefore, we decided to integrate speech input in the logging, more specifically, Dragon Naturally Speaking (Nuance).

Year	Version	Explanation
2014	Inputlog 6	linguistic analyses (additional server and user accounts) token analysis bigram analysis fluency analysis segmentation of idfx files database of users via Inputlog website
2013	Inputlog 5.2	extended pause analysis file level conversion merging with Tobii eye tracking merging with Dragon Naturally Speaking recode of sources pre-processing continue logging flow use existing documents
2011	Inputlog 5	reprogrammed in C# 64 bit compatible independent of Windows and Office version used S-notation and revision analysis focus analysis identification of copy/paste contents batch processing event Filter integrated replay (basic functionalities) bug fixes: e.g., character count and position
2009	Inputlog 4	XML output full integration in MS Word identification of window title (focus) revision analysis replay (basic functionalities) data merging with EyeWrite
2008	Inputlog 3	automated integration with Dragon Naturally Speaking introduction of parsing techniques more sophisticated algorithms underlying the analyses
2007	Inputlog 2	reprogrammed in C++ more elaborate analyses

2003	Inputlog 1	programmed in Basic

Contributors

A lot of programmers have been working on Inputlog since 2003.

- Joeri Rammelaere - programmer server linguistic analyses & progress graph - University of Antwerp-Master in computer science - summer 2011, 2012, 2013: 20% in 2013-2014
- Robin Verschoren - programmer Pause Location Markers - University of Antwerp-Master in computer science - summer 2013
- Tom Pauwaert - programmer revision module/focus analyses - University of Antwerp-Master in computer science - 2011-October 2012, March-May 2013
- Robbe Block - programmer revision/play module - University of Antwerp-Master in computer science - 2010-2011
- Joris Roovers - programmer advanced analyses - University of Antwerp-Master in computer science - 2010-2011
- Alexander Prinsier - job student - 2008-2010
- Mathia Van De Poel - internship student/job student - 2006-2007
- Maikel Bral - internship student - 2008
- Tom Druyts - internship student - 2008
- Wim Claessens - internship student - 2006
- Ahmed Essahli - internship student KdG - 2003
- Bart Van de Velde - job student - 2003-2006
- Wesley Cabus - initial programmer - 2003-2005
- Nico Verlinden has been technical coordinator from 2005-2009

Team

Inputlog has been developed since 2003 by **Mariëlle Leijten** & **Luuk Van Waes**

Mariëlle Leijten received her PhD from Utrecht University on the topic of **Writing and Speech Recognition** in 2007. She currently conducts a research project on "**Cognitive writing process characteristics in Alzheimer's disease**" as a post-doc research fellow of the Flanders Research Foundation (FWO). Furthermore, she is Assistant Professor at the University of Antwerp. Her line of research here is "**Writing from multiple (digital) sources**". In her research, she focuses on writing and digital media, more specifically analyzing on-line writing processes in professional organizations via keystroke logging. Mariëlle teaches business communication, communication research methods, and new media at the University of Antwerp. She also co-edited the book, Writing and Digital Media in the Elsevier's series on Studies in Writing.

More information can be obtained from her personal website.

Luuk Van Waes is a full professor in **Professional Communication** at the University of Antwerp (Department of Management), Belgium. He teaches business communication and research methods in the Master in Multilingual Professional Communication and at the Antwerp Management School. He has been involved in different types of writing research, with a focus on the influence of digital media on the organization of (professional) writing processes.
He published papers in several international peer reviewed journals like: Reading and Writing; Written Communication; IEEE Transactions in Professional Communication, Journal of Pragmatics, Journal of Sociolinguistics, Interacting with Computers, Computers and Composition. He is also the editor of the Journal of Writing Research.

More information can be obtained from his personal website.

Since 2010 **Eric Van Horenbeeck** is the technical coordinator of Inputlog.

Eric Van Horenbeeck received his PhD in 2008 from the University of Antwerp with a thesis on **Topical Facets: Semantic Patterns between Documents and the Vocabulary**. This subject is about unsupervised information discovery in unstructured texts by a computer program. Eric's experience with algorithms and software in the linguistic domain allows him to support the research demands of the Inputlog user.

Since 2012 **Joeri Rammelaere** is part-time programmer of Inputlog.

Joeri Rammelaere has a degree as Master in Computer Science (2013, University of Antwerp). At the moment he studies Artificial Intelligence at the KULeuven.

Acknowledgements

Inputlog has been developed in different stages. Because of the very limited financial resource, most of the programming has been done by enthusiastic students during internships or as a part-time job. We thank all the contributors for their excellent work in programming (and reprogramming) Inputlog. We hereby thank the **University of Antwerp** and the **Flemish Research Foundation** for their support. Thanks to the received grants and investments of researchers we could gradually develop the program.

The grants that we have received to invest in Inputlog include:

Conversion and analysis of Translog files
Promoter: Luuk Van Waes, University of Antwerp
Co-promoters: Rikke Hartmann Haugaard (Aarhus University) & Eric Van Horenbeeck (University of Antwerp)
Period: 2013
Funding Department of Business Communication, School of Business and Social Sciences, Aarhus University (Denmark)
Budget: 1 000 euro

Merging of pause and summary analyses
Promoter: Mariëlle Leijten, Research Foundation-Flanders / University of Antwerp
Co-promoters: Milou de Smet (Open University) & Tom Pauwaert (University of Antwerp)
Period: 2011
Funding: Open University, CELSTEC (The Netherlands)
Budget: 200 euro

Merging writing process data with lexica
Promoter: Luuk Van Waes, University of Antwerp
Co-promoters: Véronique Hosten & Lieve Macken (University College Ghent) | Mariëlle Leijten (University of Antwerp)
Period: 2010-2012
Funding Flemish Research Foundation (FWO), Belgium
Budget: 27 000 euro

Revision analysis in Inputlog
Promoter: Mariëlle Leijten
Period: 2010-2012
Funding: Legacy of Matthieu
Budget: 2000 euro

Development of a logging tool for writing research
Promoter: Luuk Van Waes, University of Antwerp
Academic staff members: Nico Verlinden & Mariëlle Leijten
Period: 2005-2008
The project was funded by the Karel de Grote Hogeschool, Antwerp
Budget: 15 000 euro

The development of a multilingual corpus of writing processes
Promoter: Luuk. Van Waes, Dept. Management, University of Antwerp
Researchers: Nico Verlinden, Luuk Van Waes & Mariëlle Leijten
Period: 2007-2009
Funding: BOF, University of Antwerp
Budget: 7 500 euro

In this project, we were happy to collaborate and discuss ideas with the developers of other keystroke loggers. We would like to thank:

- Kerstin Severinsson Eklundh (TraceIt and Jedit)
- Asa Wengelin, Victoria Johansson, Roger Johansson, Johan Frid, Sven Stromqvist (Scriptlog)
- Eva Lindgren
- Arnt Lykke Jakobsen (Translog)

We appreciated their open and warmhearted cooperation in the different stages of the development of the program.

We would also like to thank Stijn van Even and Hans Geuns of Nuance for all their efforts in making the logging facility of Dragon Naturally Speaking available to us.

Getting started

In this section you can find information about which system requirements are needed, how to install and uninstall Inputlog, how to use Inputlog the easy way, and in full option.

System Requirements

The minimum requirements to run Inputlog are:

- Microsoft Windows XP and >
- Microsoft Word XP and >

Remark: a full version of MS Word needs to be installed.

Installing Inputlog

The latest version and previous releases of Inputlog are freely downloadable for research purposes.

Step 1: register and obtain a password

In order to use the software, please, complete the on-line registration form first. You will receive a personal password in your mailbox to unlock the installation software.

Step 2: Install Inputlog

Run the installation file (Inputlog*.exe) and follow the instructions on the screen to install Inputlog properly.

Remark: if you would like to install a new version of Inputlog, please uninstall on request the previous version.

1. Run the file Inputlog*.exe.
2. Click Next to read the license agreement.
3. Click I agree to indicate that you have read and agree with the license agreement. (if you do not agree with the license click cancel)
4. Fill in the password you received by mail and click Next to continue.
5. Browse for a destination folder where you would like to install Inputlog or make a new folder and click Next to continue.
6. Click install to start the actual installation.
7. At the end of the installation you will be prompted to adjust certain settings of Inputlog.
8. Click Next to end the installation.
9. Click finish to leave the installation wizard.

Uninstalling Inputlog

Simply follow the instructions of the **Uninstall Wizard** to uninstall Inputlog properly from your computer. (Note: Inputlog is also automatically uninstalled when installing a new version of Inputlog).

1. Click the **Start** button.
2. Select **Programs** > **Inputlog** > **Uninstall Inputlog** to remove Inputlog from your computer

Starting Inputlog (the easy way)

How to record, analyze and replay a writing session in **3 basic steps**.

Step 1: Record

1. Click the **Record tab**.
2. Indicate where you want to save the logging file.
3. Name the file.
4. Fill in the session identification.
5. Click record to start logging a writing session.
6. Microsoft Word will start automatically.

Step 2: Analyze

1. Click the **Analyze tab**.
2. Search the *.idf file you want to generate data files from.
3. Indicate which data output you want to generate.
   Preferably, you first log a general analysis and use this for further **Filter** of the data.
4. Click generate to start generating the chosen data files.
5. The generated files will be saved automatically.

Step 3: Play

1. Click the **Play tab**.
2. Search the *.idf file you want to replay.
3. Indicate at which speed you want to play the file.
4. Click play to start replaying the writing session.

Using Inputlog (the full option)

The interface of Inputlog consists of 5 different menus: record, pre-process, analyze, post-process and play.

element	function
record	This module logs all input (keyboard, mouse and speech) in Microsoft Word, and other Windows based programs together with a unique time stamp.
pre-process	This module allows to process data from various perspectives: event based, time based, mode based or based on window changes (sources).
analyze	This module is the heart of the program and features six aggregated levels of analysis: summary, general, linear, pause, source, and revision. The resulting XML files contain specific and/or statistical information about a particular writing session. Several process characteristics are presented based on algorithmic processing of the raw logging data. Remark: All analyses are based on theoretical and empirical findings, recent research and discussions with experts in the field.
post-process	This module integrates single or multiple log files from Inputlog or other observation tools (Morae, Dragon Naturally Speaking, Eye tracking data). Secondly, it is possible to merge multiple output files - horizontally or vertically - for further analysis in, for instance, SPSS or MLWin.
play	The play module allows you to play back the recorded session at various levels (time or revision based). The replay is data based (not video based) and the play speed is adjustable. You can also reconstruct a logged session revision by revision.

Route 1 is a situation in which your logging is flawless; the beginning and the end of your logging session does not contain noise. No additional Pre-processing is necessary and you can use the flow: Record > Analyze (basic in Inputlog) > Post-process > Analyze (extended via Excel/SPSS). This option is also described in Figure 1.

Starting Inputlog

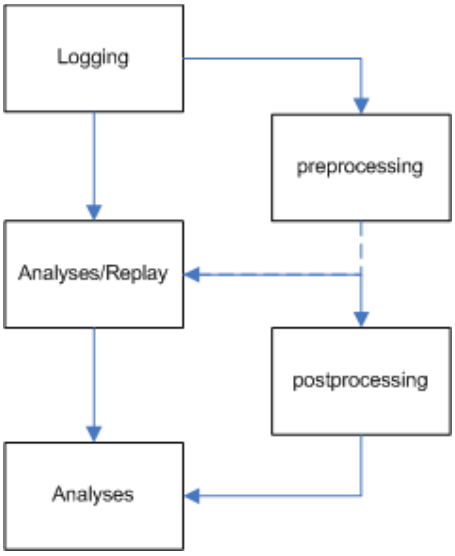


Figure 1. Flow of Inputlog

Route 2 is a situation in which you need to refine logged data prior to analyses.

Example. When analyzing writing from multiple digital sources, researchers can choose to select only MS Word events. Each event in MS Word is included in the analysis, and all the other activities are left out. The Filter is also convenient for isolating different writing tasks, or for deleting logging session start-up or deactivation 'noise'.

The described situation needs the flow: Record > Analyze (basic in Inputlog) > Pre-process > Analyze (basic in Inputlog) > Post-process > Analyze (extended via Excel/SPSS).

The detailed flow of logging, Pre-processing, analyzing, Pre-processing and replaying is shown in Figure 2.

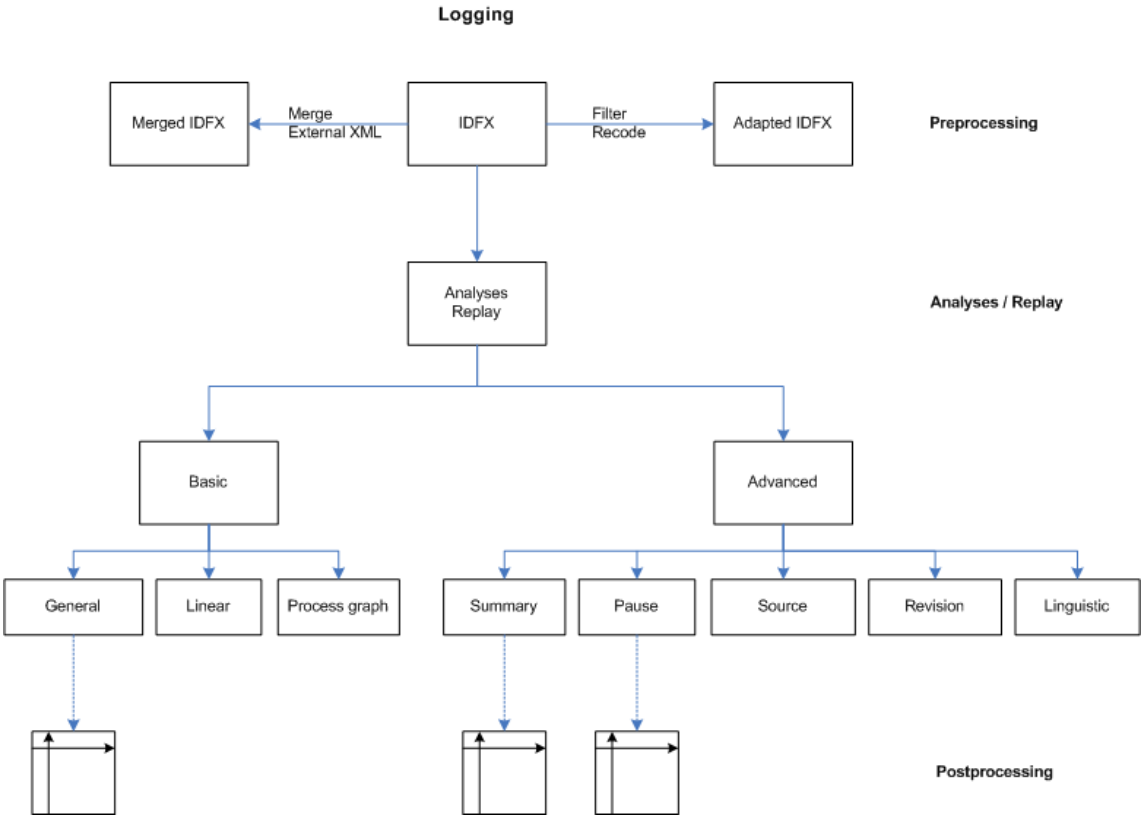


Figure 2. Detailed flow of Inputlog.

Compatibility

To be added

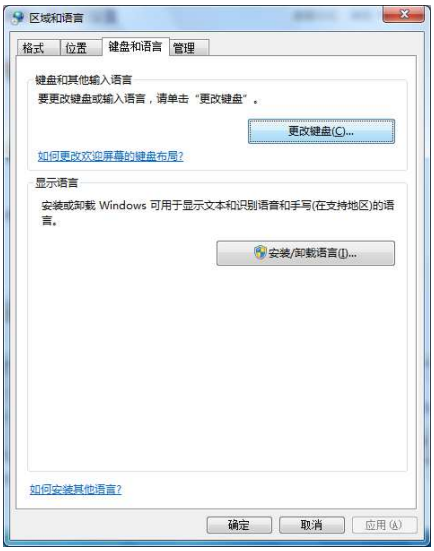
Keyboard Supported Layouts

Latin & Greek (sami)
Inputlog can only log Western Scripts like Latin & Greek.
Belgian (Comma)
Belgian French
Canadian French
Canadian French (Legacy)
Canadian Multilingual Standard
Danish
Dutch
Faeroese
Finnish
French
Gaelic
German
German (IBM)
Icelandic
Irish
Italian
Italian (142)
Latin American
Norwegian
Portuguese
Portuguese (Brazilian ABNT)
Portuguese (Brazilian ABNT2)
Spanish
Spanish Variation
Swedish
Swiss French
Swiss German
United Kingdom
United States-Dvorak
United States-Dvorak for left hand
United States-Dvorak for right hand
United States-International
US
Chinese

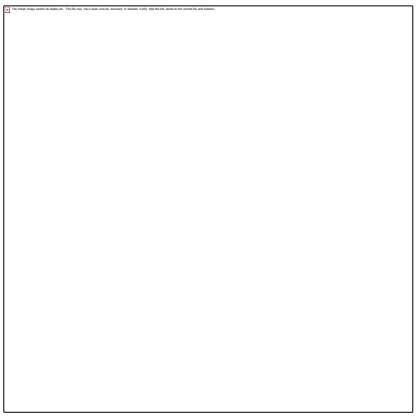
To log **English** texts on **Chinese Windows systems** with Inputlog 5.1 (and higher) you need to adjust the language settings.

Important: Change the settings as described below before you start Inputlog.

1. Go to **Start** in the Taskbar and select **Control Panel**



2. Select **Clock, Languages and Region**
3. Select **Region and Language**
4. Select **Keyboards and Languages** > Change **Keyboards**

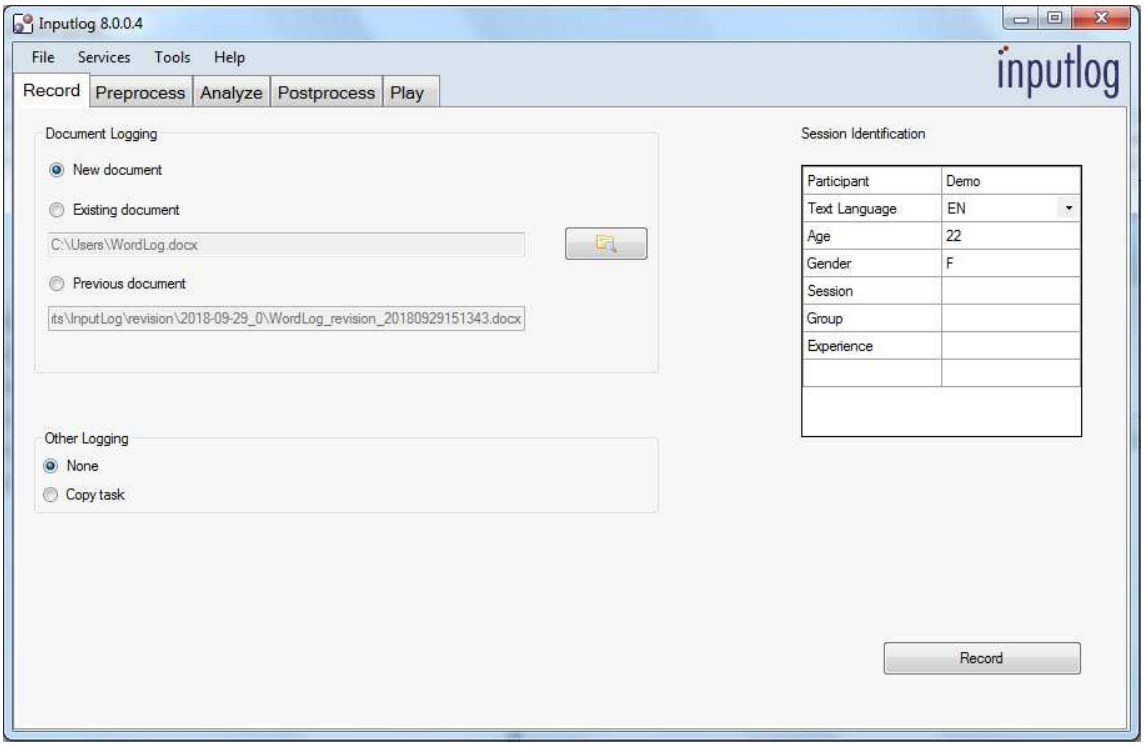


5. Select **General Tab** and change the **Language Settings**

(Thanks to Xu Cuiqin for the screenshots, Nanjing University | China)

Record

Before you start a new logging session you must provide a name and a location for the logging file, as well as identification data about the session. The information that you provide will be included within all files that you generate from that session. This will enable you to easily identify each writing session.



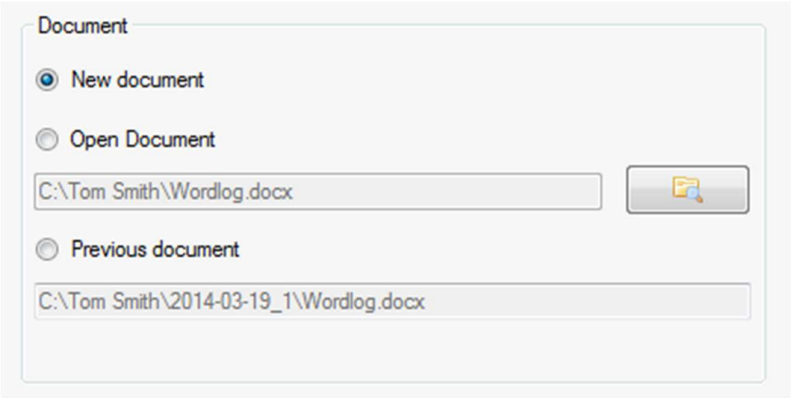
The following gives a functional description of each element.

element	function
document	You have four options to start logging from either: 1. A new document, 2. An open document, 3. A previous document, 4. A copy task By default 'New document' is selected.
session identification	Use the session identification to add set values to each file that you generate. These session identifiers are included in the analyses. You can use them to describe/identify your participants. After each session this information is kept in the interface. If you have multiple sessions by one participant the personal details are already filled in.

record	After you have provided all the necessary information, you click 'record'. The selected MS Word document will open automatically and the recording of the session starts immediately.

Document

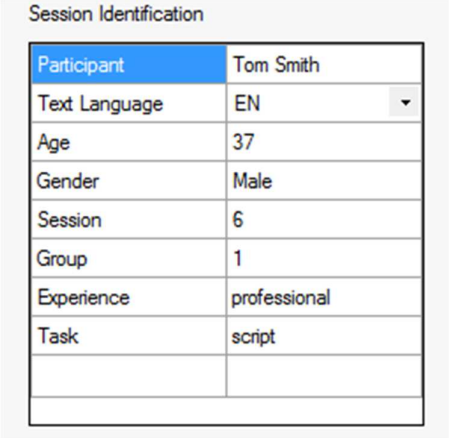
You can start a logging session from three types of documents.



label	description
new document	This options starts an empty document in MS Word. The file is automatically named 'Wordlog.docx'
open document	Via this option you can select a document that is already in progress. This option is also suitable to start from a predefined template.
previous document	Via this option Inputlog automatically opens the previous document that the participant has been working in. This option is very suitable for professional writing settings.

Session Identification

You can identify your writing session using ten variables or "labels". These session identification labels are saved in 'headers' in the XML-files. You can convert this file to SPSS, or another statistical program. As this data is written to facilitate statistical analyses, it is important that you define your labels and variables in advance. Inputlog has 6 defined and 4 undefined variables.



label	description
participant	Give each participant a unique value (e.g., Mike=1, Janet=2, Chris=3). Note: You cannot change this name in the idfx-files, so if you need to be anonymous you better select a unique code for each participant.
text language	Identify the language of the session. This setting is of importance to run the correct Linguistic Analyses. At the moment (4/12/2013) the languages Dutch and English are supported.
age	Insert the age of the participant (value).
gender	Give a unique value per gender (e.g., male=0, female=1).
session	Indicate to which general session the particular logging session belongs (if necessary, e.g., session 1, session 2, etc.)
group	Give a unique value to each group (e.g., students=0, academics=1).
experience	Give a unique value to each experience level (e.g., low-proficiency=0, high-proficiency=1).
additional labels	You can add additional labels and their values (e.g., label = task, values 1 to 4).

Copy Task

In the creation of digital texts, typing skills are a factor that could influence text production. Therefore, the current copy task in Inputlog is designed to create a set of measures that allow a fine grained analysis of low level typing and motor skills.

We have opted for a java based web interface that stepwise guides the participants through the different components of the copy task. A progress bar at the right of the screen indicates the task stages. The interface has been adapted for Chrome, Internet Explorer, Firefox and Safari.

The copy task can either be activated via the Record tab in Inputlog, or directly via the web using following URL: http://inputlog.ua.ac.be/Website/copytask/tasks.html

The copy tasks are coded as XML-files and can be created by using the 'Inputlog copy task creator'. It is our intention to develop a comparable copy task for different languages, based on the same underlying principles.

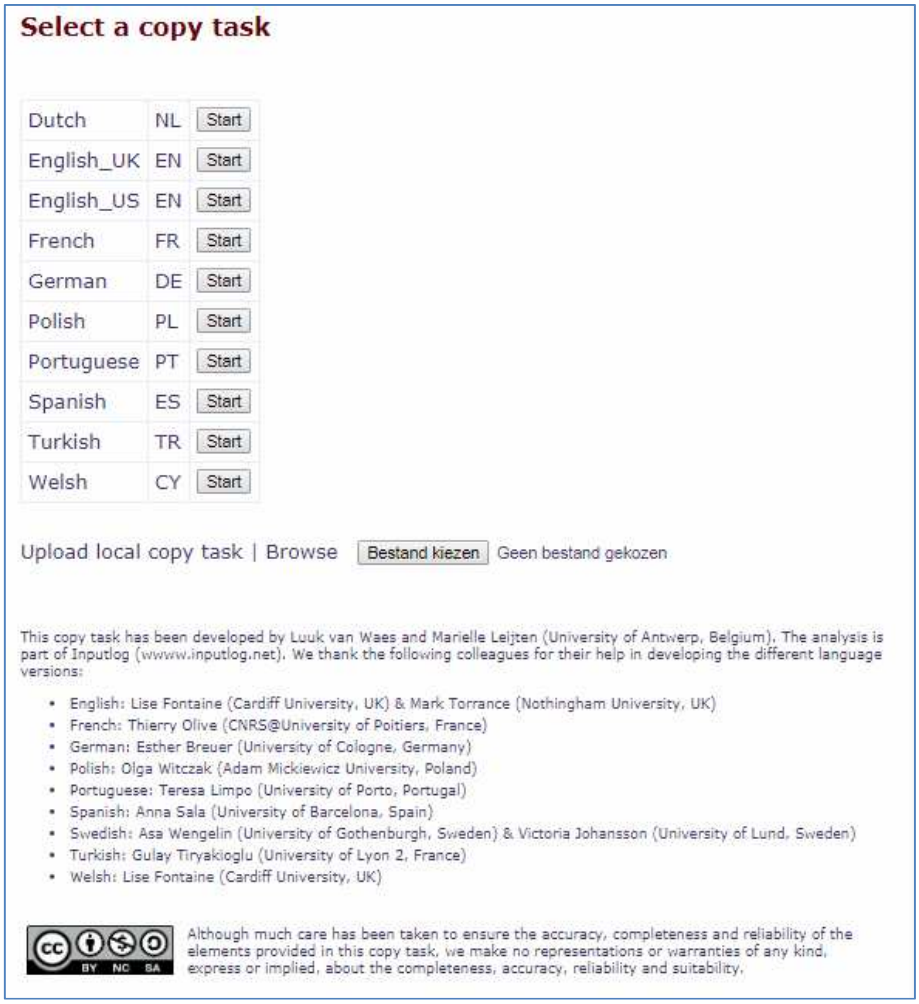
Components

The copy task (e.g., English UK - Qwerty) consists of the following components:

Selection	selection of copy task (language specific)
Identification	participant identification (name, age, gender, session, keyboard)
Introduction	general task instruction
Tapping task	press the 'd' and 'k' key alternatively during 15 s
Sentence	copy a sentence
Example	dummy item explaining word combination tasks
Word combination 1	copy a combination of three words seven times
Word combination 2	copy a combination of three words seven times
Word combination 3	copy a combination of three words seven times
Word combination 4	copy a combination of three words seven times
Consonant groups	copy four blocks of six consonants
Extra information	extra participant information (handedness, hard and software used, education, L1, learning disabilities)
End	thank you (and download)

Language selection

The participants can select a predefined (default) copy task or can upload a specific task. The task can be accessed via the 'Record Tab' in Inputlog or via the following URL:
`http://inputlog.ua.ac.be/WebSite/copytask/tasks.html`

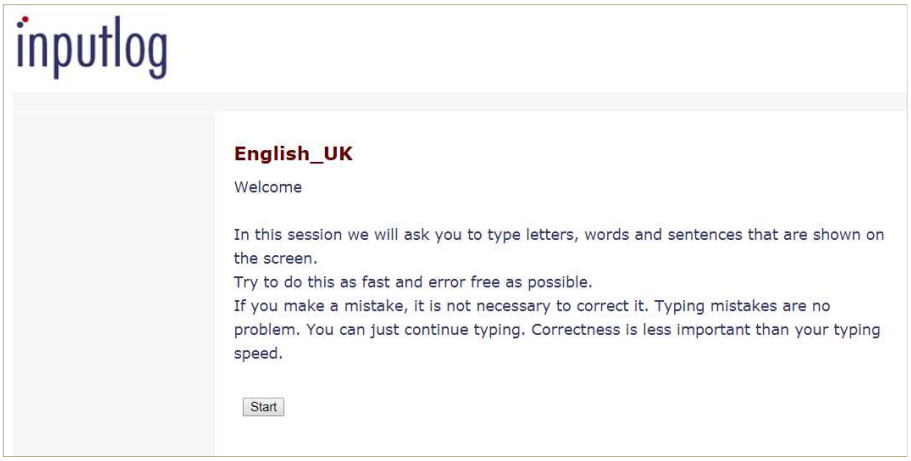


Participant identification

The participants are invited to provide some basic identification information: Name, age, gender, session name (if applicable) and keyboard layout. We also included a privacy notice to obtain consent from the participant, in line with the privacy policy stated in the General Data Protection Regulation (GDPR) of the European Union.

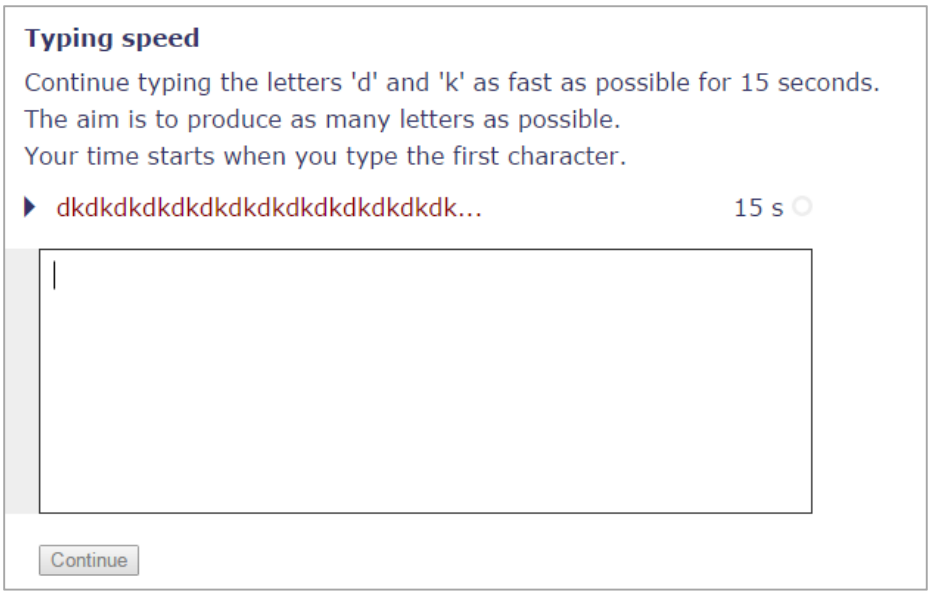
Introduction

The copy task is explained in a few lines.



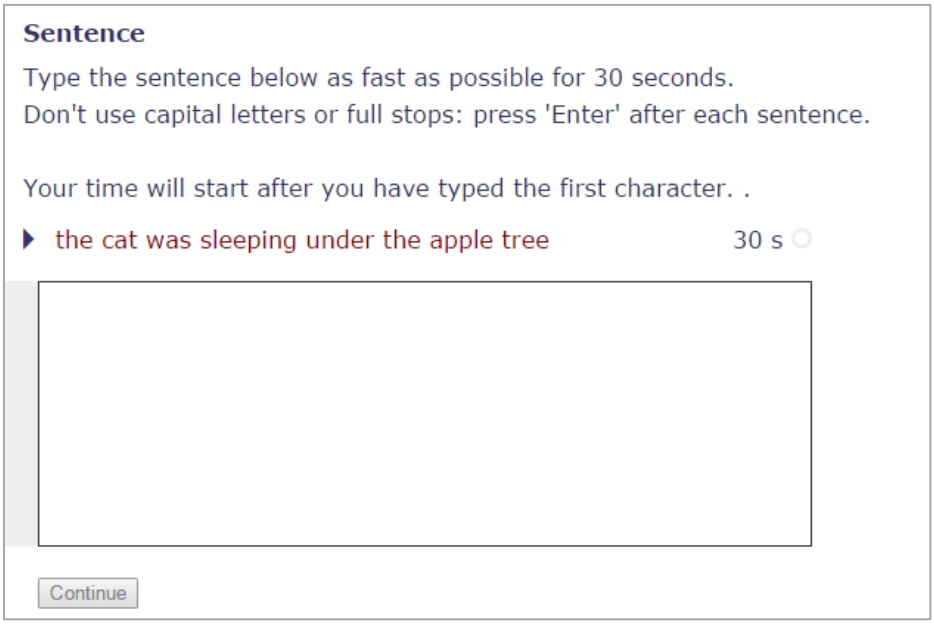
Tapping task

The tapping task intends to measure the fastest motor speed of pressing two keys with alternating hands (viz. 'd' and 'k', resp. a LeftRight and RightLeft hand combination). Finger-tapping tasks are commonly used to study the human motor system. Tapping tasks 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 (Witt et al., 2008).
In this copy task a bimanual, self-paced tapping task is opted for. Participants are asked to type the 'd'-'k' key combination for 15 s. (Salthouse, T.A., 1984). A time circle at the top right corner is used as a time indicator.



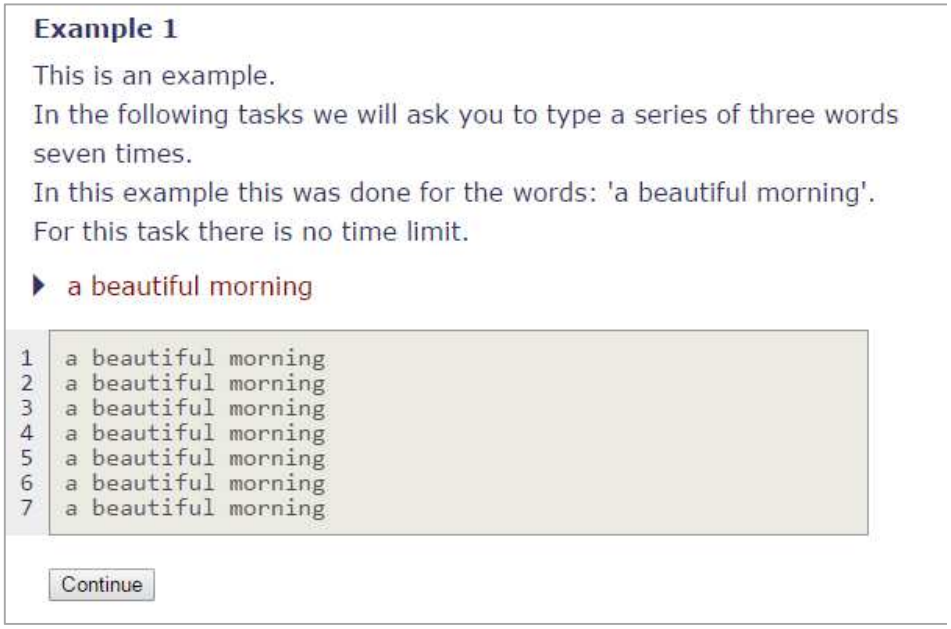
Sentence task

The sentence copy task intends to measure the typing skills related to copying a series of - short high frequency - words in a sentence context. Participants are asked to repetitively type this sentence for 30 seconds.



Example

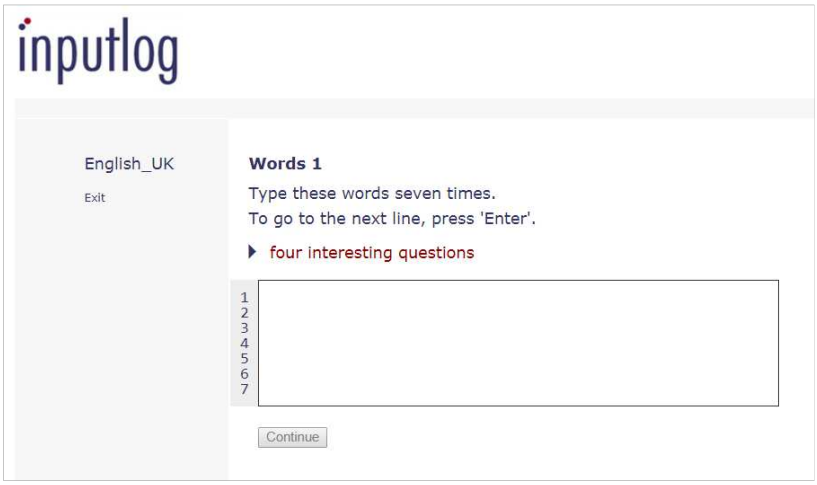
An example is used to explain the next typing tasks, i.e. the repetitive (seven times) typing of a three word combinations.



Word combinations 1 to 3 (HF)

Three sets of word combination are presented. In each word combinations high frequent bigrams are implemented spread over three words. Participants are asked to type these word combination seven times.
To create these word combinations we used the following criteria:

- three words per entry
- combination of a numeral + an adjective + a noun
- word length: between 19 and 24 characters in total - [3/4] + [10/12] + [8/10] characters (average number of characters per word combination: *M*: 22.3 (*SD*: 1.6)
- high frequency words (30% highest segment; lemma frequency)
- high frequency bigrams only highest 30 % percentile in CELEX/Subtlex or another comparable corpus based bigram frequency list)
- 18 to 20 high frequent bigrams in total (no LF bigrams)
- on average 3 to 5 bigrams for each hand combination (LL; LR; RL; RR)
- 5 to 7 keyboard adjacent bigrams
- no (or maximum one) repetitive keys



Word combination 4 (LF)

This word combination copy task intends to measure the typing skills related to low frequent bigrams in a three word noun phrase context. Participants are asked to type these word combination seven times.
To create word combination 4 we used the following criteria:

- three words entry
- combination of indefinite determiner/countable + adjective + noun
- word length: between 17 to 21 characters - [3/4] + [7/10] + [6/8] characters (average number of characters per word combination: *M* 19.0 *SD*: 1.3)
- 3/4 low frequent bigrams (< 50% percentile in CELEX/SUBTLEX)
- no (or maximum 1) repetitive keys

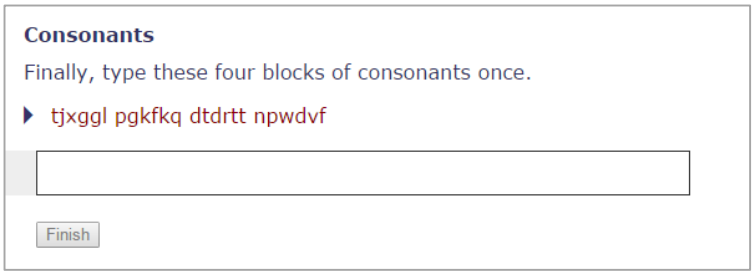
Example

English_UK (Qwerty) "Copy task"	word combination 1	word combination 2	word combination 3	word combination 4
Word 1 (numerical)	four	seven	five	some
Word 2 (adjective)	interesting	wonderful	important	awkward
Word 3 (noun)	questions	surprises	behaviours	zigzags

#characters	24	23	23	18
High Frequent bigrams (HF)	19	18	18	8
Low Frequent bigrams (LF)	0	0	0	4
Left-Left (LL)	4	6	1	5
Left-Right (LR)	4	6	2	3
Right-Right (RR)	4	2	5	1
Right-Left (RL)	3	4	2	2
Adjacent keys	7	6	3	4
Repetitive keys	0	0	0	0

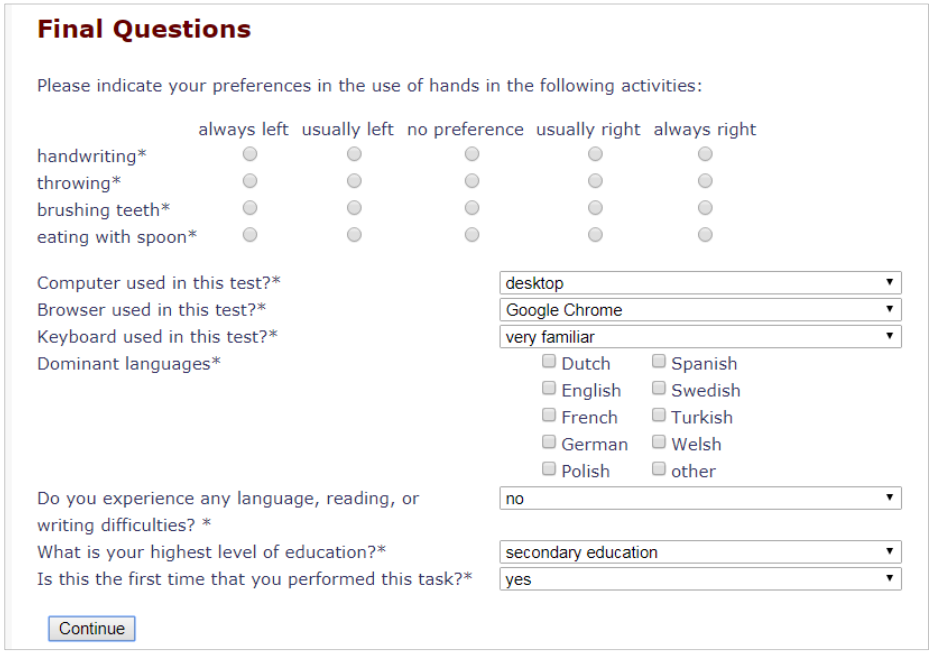
Consonant groups

The consonant copy task intends to measure the typing skills in a non-word context. Participants are asked to copy four blocks of six consonants once.



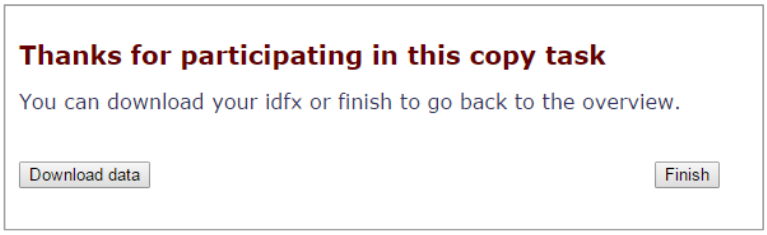
Extra information

We end the task flow with a final set of questions addresses respectively, handedness, hardware and software used, dominant language, reading or writing difficulties, and familiarity with this task.



Data storage

The final screen shows the data storage: online transfer to web server is done automatically and the off-line storage is possible by selecting the 'download' button.



Tools

- **CELEX bigram frequency**
  https://catalog.ldc.upenn.edu/LDC96L14
  Baayen, R, R Piepenbrock, and L Gulikers. CELEX2 LDC96L14. Web Download. Philadelphia: Linguistic Data Consortium, 1995.
- **Subtlex word frequency**
  http://crr.ugent.be/programs-data/subtitle-frequencies
  Walter J. B. van Heuven, Pawel Mandera, Emmanuel Keuleers & Marc Brysbaert (2014) SUBTLEX-UK: A new and improved word frequency database for British English, The Quarterly Journal of Experimental Psychology, 67:6, 1176-1190, DOI:10.1080/17470218.2013.850521
- **WordGen**
  http://users.ugent.be/~wduyck/wwgman.htm
  Duyck, W., Desmet, T., Verbeke, L., & Brysbaert, M. (2004). WordGen: A Tool for Word Selection and Non-Word Generation in Dutch, German, English, and French. Behavior Research Methods, Instruments & Computers, 36(3), 488-499. (full text available here)
- **Coded Excel file:** Excel constructed to design a language specific copy task based on the criteria described above (contains frequency list, finger combinations, adjacency).
  Bigram_"Language"_coded_hand_freq_item construction.xlsx
- **Inputlog Copy task Creator**
  Part of the 'Service' menu in Inputlog 7.1 + manual
- **Inputlog: copy task analysis**
  Part of the 'Record' and 'Analysis' tab in Inputlog 7.1

6. URL: http://inputlog.ua.ac.be/WebSite/copytask/tasks.html

- **Keyboard lay-out**
  AdjacentCharacters.exe (Eric Van Horenbeeck)
  Script that identifies hand combination, repetion and key adjacency on a keyboard
- **Python script to extract bigram frequency**
  count_char_bigrams.py (by Maximiliana Behnke <maximiliana.behnke@ed.ac.uk>)
  A small python script that extracts frequency of character bigrams in subtlex corpora.

References

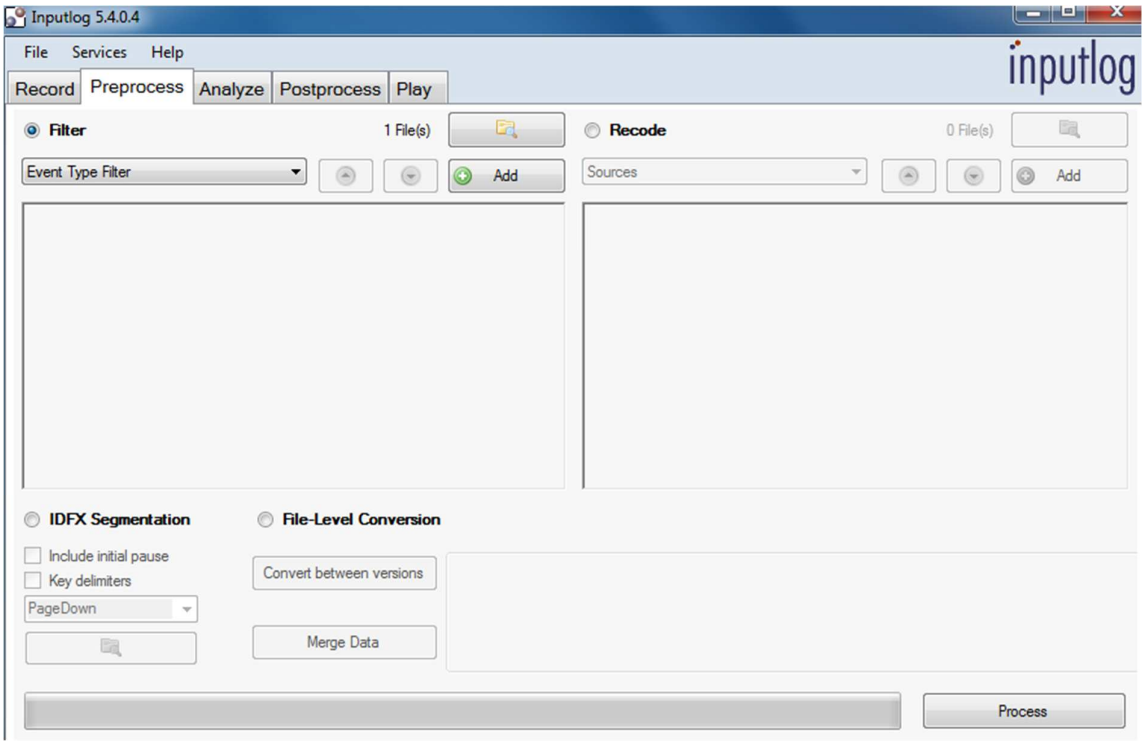
Salthouse, T. A., Rogan, J. D., & Prill, K. A. (1984). Division of attention: Age differences on a visually presented memory task. *Memory & Cognition, 12*(6), 613-620. doi:10.3758/bf03213350

Witt, S. T., Laird, A. R., & Meyerand, M. E. (2008). Functional neuroimaging correlates of finger-tapping task variations: an ALE meta-analysis. *Neuroimage, 42*(1), 343-356.

Pre-process

The ***Pre-process*** module offers four options:

option	description
Filter	You can Filter logging data: 1. event based, 2. time based, 3. based on window changes (sources). You can use these Filters single or multiple.
recode	You can recode the sources of an idfx-file
IDFX Segmentation	You can segment an original idfx-file into multiple idfx-files.
file-Level conversion	You can convert files in two ways: 1. Conversion between Inputlog version (Note: this option is only possible with idfx-files logged with Inputlog 5.*) 2. Merge data of Inputlog with a. Dragon Naturally Speaking b. Tobii TX 300



Via Filter, Recode, Segmentation and Conversion you create a **new IDFX** (IDFX_bis in Figure 3). This idfx is the input for further analyses.

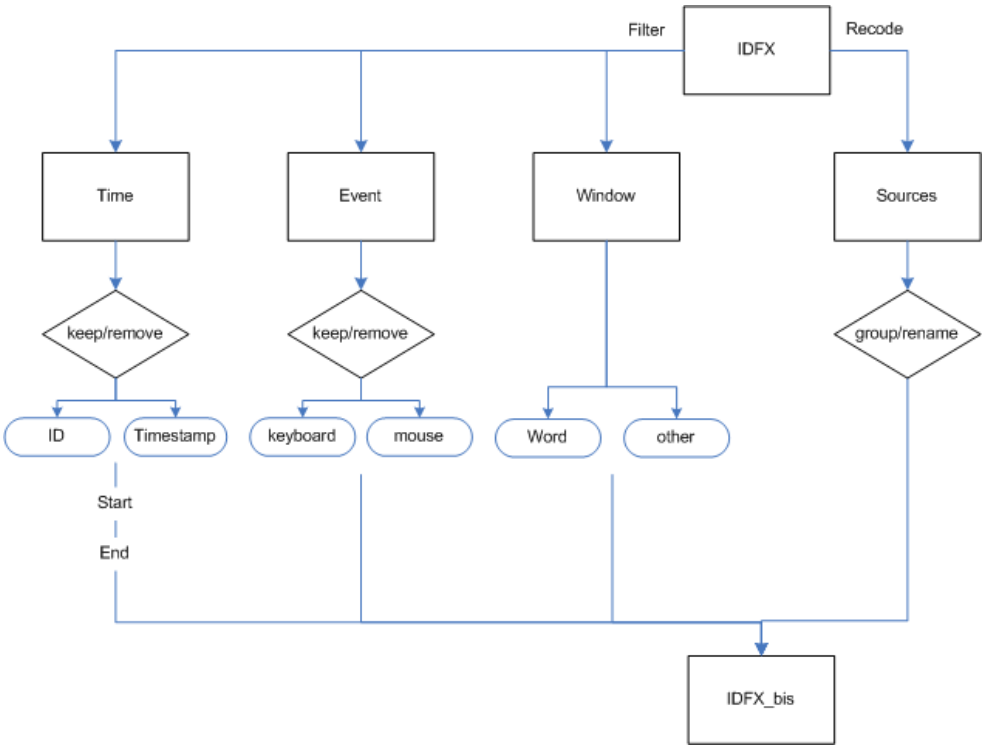
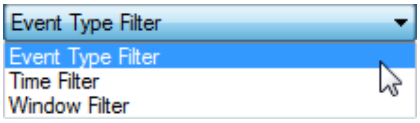


Figure 3. Detailed flow of Pre-processing actions Filter and Recode.

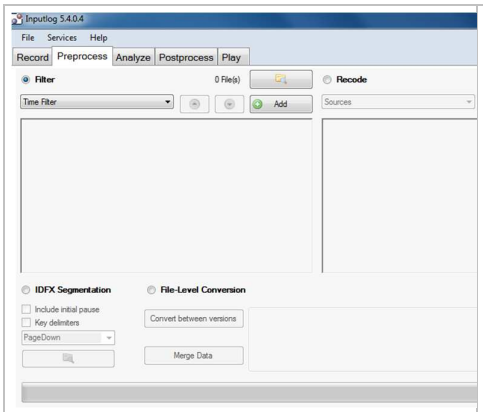
Filter

You can select three Filters. You can use one Filter at the time, or use multiple Filters simultaneously. Be aware that the selection of a Filter might be of influence to another Filter.

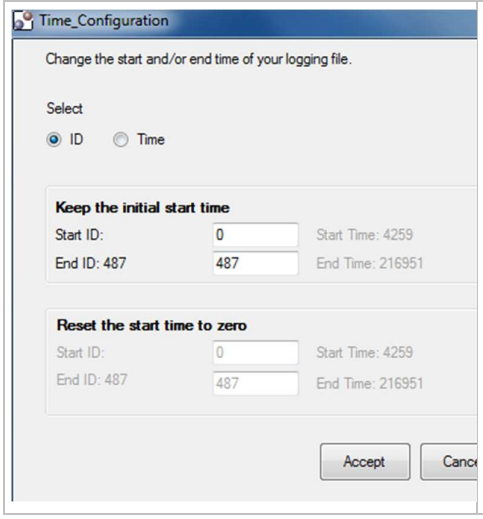


Time Filter

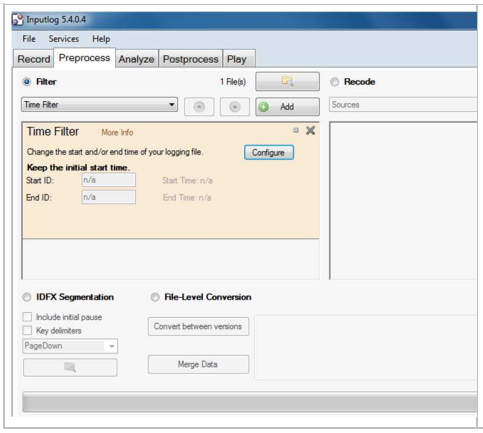
Via the time Filter you can change the start end/or end time of you logging.



1. Select the radiobutton Filter
2. Browse to the idfx file you would like to Filter
3. Select the Time Filter
4. Click Add
5. Click configure
6. The time configuration wizard opens



7. Select ID or Time. You have now two options:
   a. Keep the initial start time of the logging session
   b. Reset the start time to zero
8. Click Accept

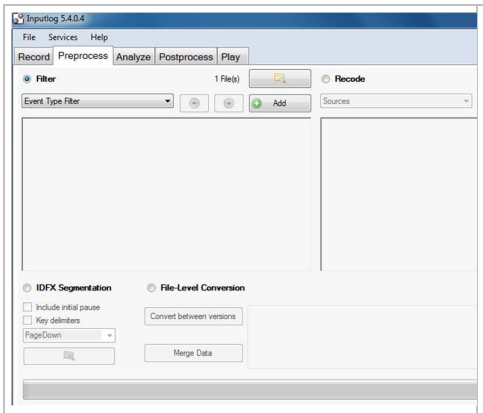


9. Click Process and save the new idfx-file at a new location

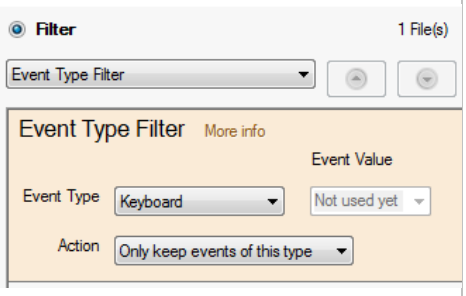
Remark: you use this idfx-file for further analyses.

Event type Filter

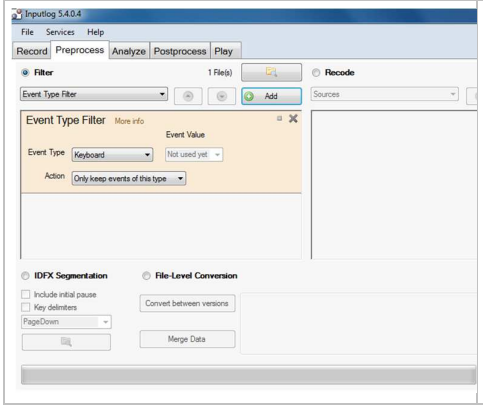
Via the Event Filter you can select the type of event that you want to analyze.



1. Select the radiobutton Filter
2. Browse to the idfx file you would like to Filter
3. Select the Event Type Filter
4. Click Add



5. Select the Event Type. You can select two event types:
   a. keyboard
   b. mouse
6. Select the Action. You can select two Actions:
   a. Only keep events of this type
   b. Remove events of this type

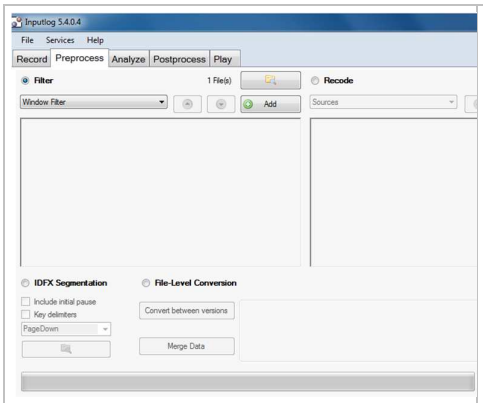


7. Click Process and save the new idfx-file at a new location

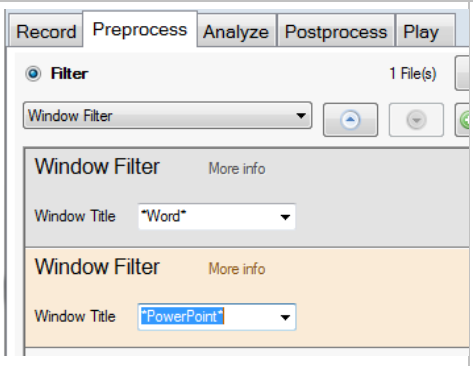
Remark: you use this idfx-file for further analyses.

Window type Filter

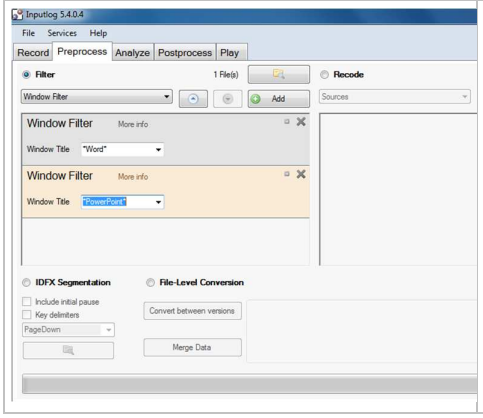
Via the Window Type Filter you can select one or multiple Windows that you want to include in further analyses.



1. Select the radiobutton Filter
2. Browse to the idfx file you would like to Filter
3. Select the Window Filter
4. Click Add
5. Note: You can select one or multiple windows to include in your analyses. You can opt to analyze data from MS Word and PowerPoint, but excluding information from other windows (internet explorer, etc.).



6. You can select predefined windows:
   a. *word*
   b. *firefox*
   c. *chrome*
   d. *internet explorer*
   The asterisk (*) allows for additional information in the naming, e.g., 'Wordlog.docx'
7. You can also add a new window name, e.g., PowerPoint.

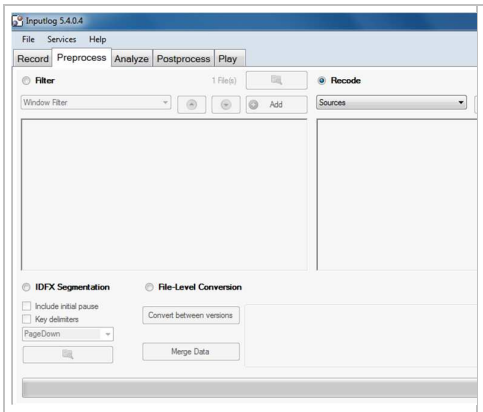


8. Click Process and save the new idfx-file at a new location

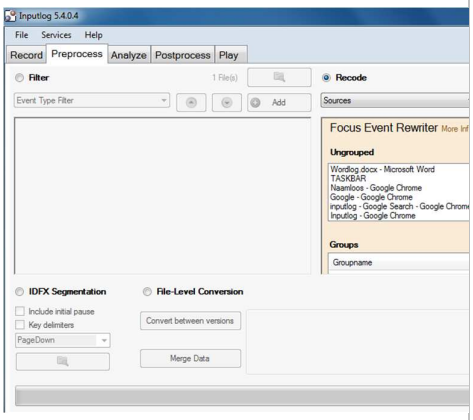
Remark: you use this idfx-file for further analyses.

Recode

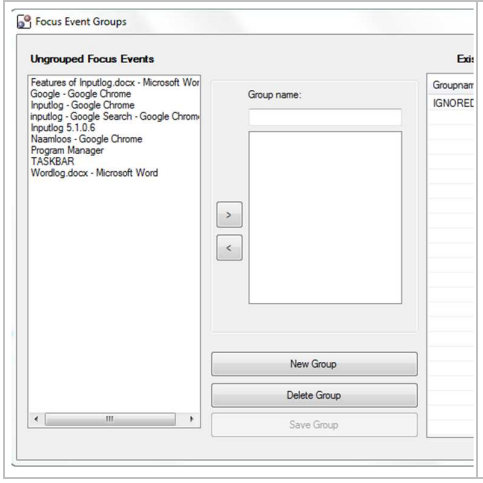
In Inputlog 6 you can recode Sources into groups: e.g., various internet pages that you would like to recode into 'internet', or the use of dictionaries into 'on-line dictionaries'.



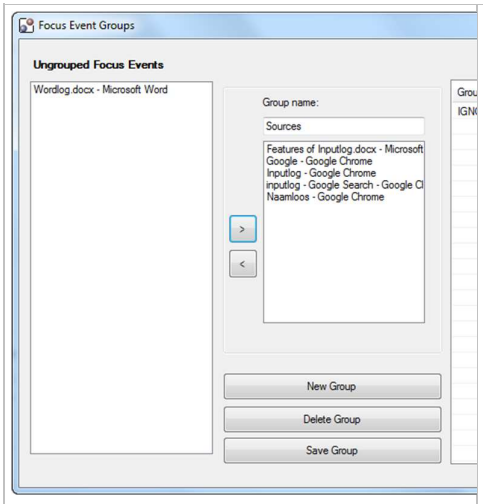
1. Select the radiobutton Recode
2. Browse to the idfx file you would like to recode
3. Select the Time Filter
4. Click Add
5. The ungrouped sources are shown in the Focus Event Rewriter



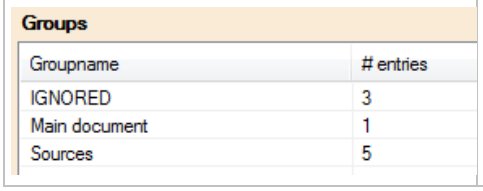
6. Select any ungrouped source by double-clicking a source or click New Group.
7. The Source wizard opens



8. Select Ignored under Excisting Groups
9. Add sources that you want to ignore: e.g., Inputlog, Taskbar.
10. Click Save Group



11. Click New Group
12. Type a Group Name
13. Add Ungrouped Focus Events via > into the group
14. Save Group
15. Repeat this for each group of sources
16. Click the Close Wizard cross (X)



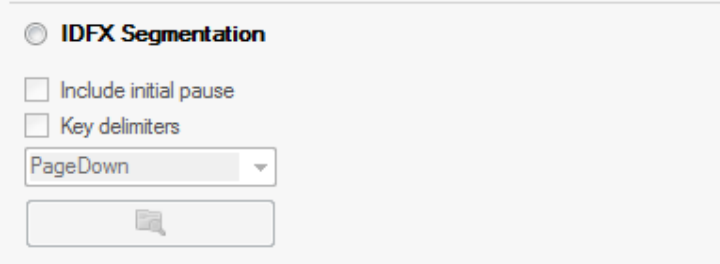
Groups

Groupname	# entries
IGNORED	3
Main document	1
Sources	5

17. The recoded sources (groups) are shown in the interface.
18. Click Process and save the new idfx-file at a new location.

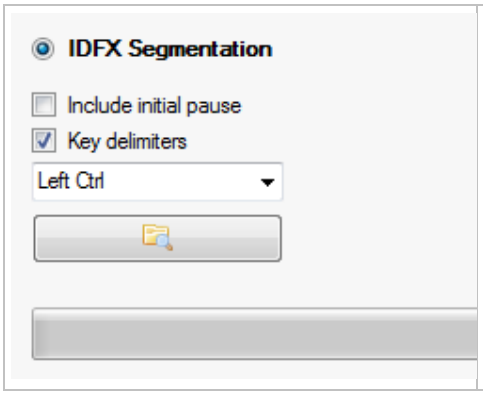
IDFX segmentation

The IDFX segmentation allows you to automatically split the IDFX file from a logging session into several subparts, separated by a specific keystroke. For example, if you choose to segment on the PageDown key and this key is pressed 2 times during the session, 3 subparts will be found: from the beginning to the first PageDown, between the two PageDown presses, and from the second PageDown till the end of the session. The output directory will contain a complete *.idfx file, and n *_segment_n.idfx files. Segmentation can also be performed after the session is completed (see Preprocess module). For IDFX segmentation to be activated, the "key delimiters" checkbox must be checked.

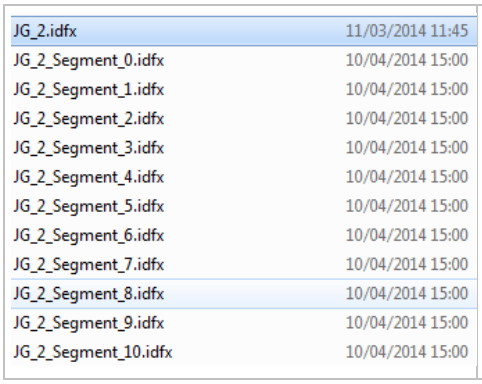


label	description
include initial pause	By default, events before the first keystroke in every segment are discarded (e.g., time taken to read the assignment). Use this checkbox to preserve these events when desired.
key delimiters	Select which key will act as separator between session subparts. Currently available: PageDown, =, Pause/Break, Left and Right Ctrl

You can follow this procedure to segment and idfx-file.

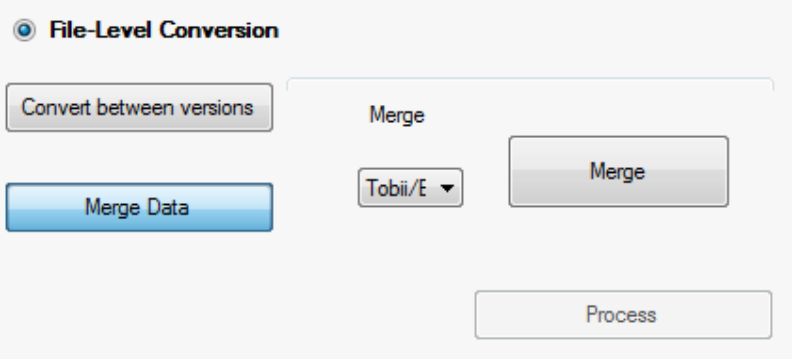


1. Select the radiobutton IDFX Segmentation
2. Browse to the idfx file you would like to segment
3. Select if you would like to include a delimiter
4. Select the Key delimiter
5. Click Process
6. The segmented idfx-files are shown at the location of the original idfx-file

	7. This will result in a number of new segmented idfx-files Each file can be used for new analyses. Each idfx starts with ID 1

File Level Conversion

Via File level conversion you can convert between previous versions of Inputlog (5.0 and higher) and merge data of Tobii eye tracking and Dragon Naturally Speaking speech recognition.

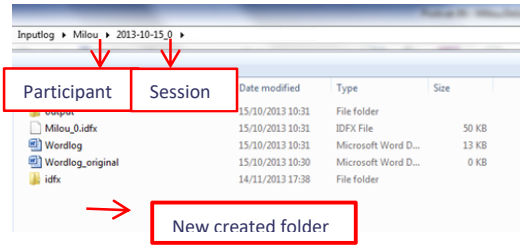


Convert between versions

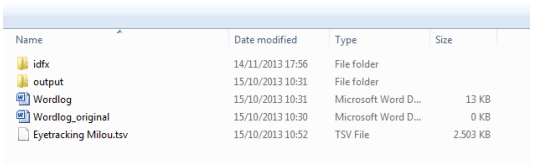
To be added.

Merge data with Tobii Eye tracking

Before you start merging you need to make sure the data is at the correct location.

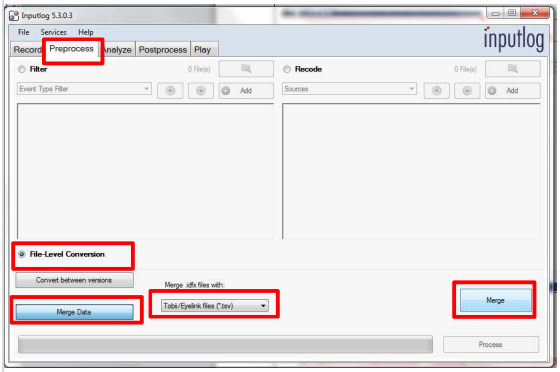


1. Within the general Inputlog folder you find a folder with the participant's name, as indicated during the recording session.
2. Open the participant's folder to find the sessions(s) folder(s) with the date and session as title (YYYY-MM-DD_#).
3. Create a new folder within this session folder. This new folder should be named 'idfx'.
   Note: Any other name won't work, so please be sure to use this title.

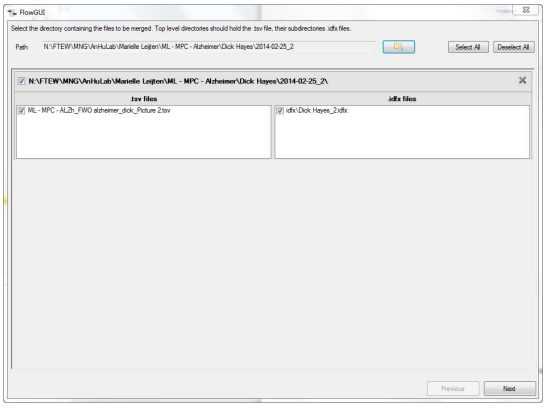


Replace (don't copy) the .idfx file into the new idfx folder.

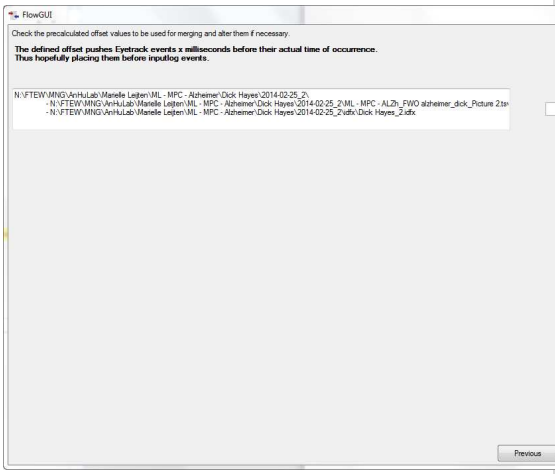
4. Save the .tsv eye tracking file in the session folder, i.e., one level higher than the .idfx file. The folder structure should then be as pictured.



5. Open the **Preprocess** menu in Inputlog
6. Select the radiobutton **file-level conversion.**
7. Select **Merge data**
8. Select merge idfx files with **Tobii/Eyelink files (*.tsv)** from the drop-down menu
9. Select the **Merge** button.

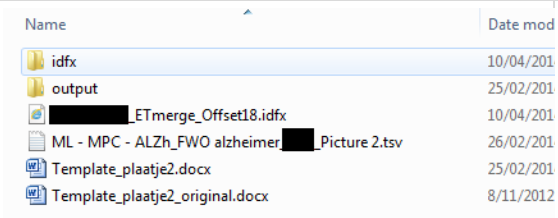


10. Browse to the general session folder (YYYY-MM-DD_#). Inputlog automatically selects the *.tsv and *.idfx file that are stored in this folder. and puts then next to each other.
11. Click Next



12. Inputlog shows the calculated offset* time.
13. Click Nex and Finish.
14. Inputlog now creates a new *.idfx file merging the keystroke logging and eye tracking data. Inputlog uses an offset for this new file to match both data files simultaneously. This new file can now be found within the general session folder.

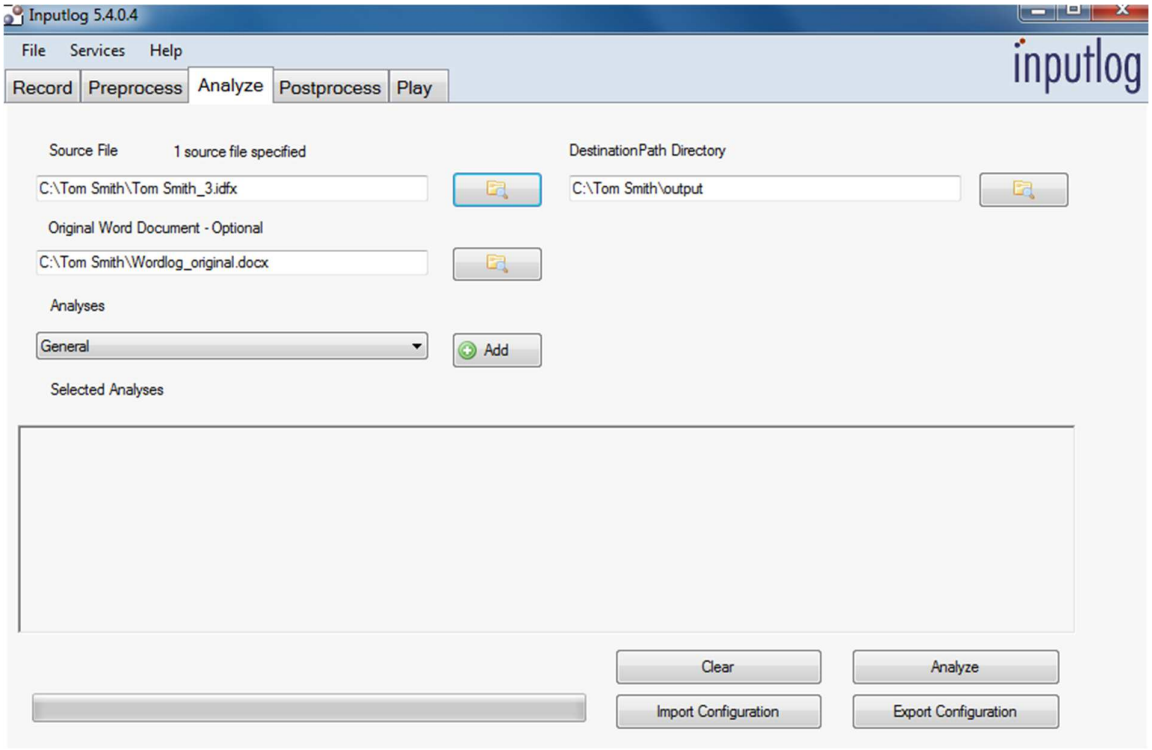
    *Offset: Inputlog uses the first five A-E-I keyboard events to place Tobii before the Inputlog-event. It uses the biggest shift between the Tobii and Inputlog data +3ms extra. When there are fewer than five matches, the offset is set to +0ms.



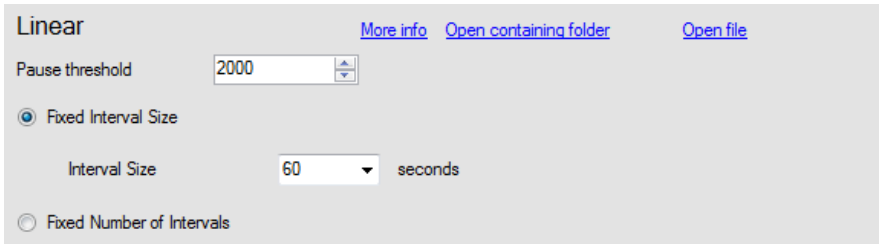
15. Open the Analyze tab of Inputlog
16. Browse to the new idfx file with underscore Offset#. This analysis includes both the keystroke logging and eye tracking data.
17. Select General or General - Condensed Eyetrack.

Analyze

Via the analyze tab you can analyze your logged data. In this section you can find how to perform analyses and how to interpret these analyses.



element	function
source file	You select the idfx file that you would like to analyse. 1. If you have just recorded a session, the last idfx will automatically appear as source file 2. You can select any previously recorded source file.
original Word document - optional	This element has two ways of operating: 1. This field will be automatically filled in if you continue with an idfx-file that you have just finished recording. 2. If you would like to perform a **Error! Not a valid result for table.** you need to identifiy the original Word document that is connected to this writing task. You have two options: a. You started from an empty document: the file will a normal WordLog_original file (empty). b. You started from a document: the file will be a WordLog_original file based on the file you have selected (template/draft/...) Note: the Data Structure needs to be similar as when you recorded it.

destinationpath directory	You can indicate at which location you would like to store you output-files.
analyses	You can select via a dropdownlist which analyses you would like to generate: 1. General 2. Linear 3. Process Graph 4. Summary 5. Pause 6. Source 7. Revision 8. S-notation 9. Token analyzer 10. Fluency 11. Bigram 12. Linguistic
selected analyses	In the field of selected analyses you see which analyses you have selected. In this field you can choose which settings you prefer for each analysis. These settings are furhter explained for each **Error! Not a valid result for table.**.
export configuration	If you would like to conduct similar analyses for various datasets, you can opt to create a configuration.
import configuration	You can import a predefined configuration of analyses.
clear	You can clear the selection you have made at once via **Clear**. Note: you can also clear single analyses via the close button in the top right corner. 
analyze	After you have selected all the necessary analyses, you click 'analyze'. The selected analyses will be generated. You have two options to open the analyses: 1. You can open them via the interface via **Open File**. 2. You can open them via the folder in which you have saved the output files via **Open containing folder** or via your Windows explorer. 

Inputlog generates basic and advanced analyses.

basic	**advanced**
1) General 2) Linear 3) Process Graph	1) Summary 2) Pause 3) Source 4) Revision 5) S-notation 6) Token analyzer 7) Fluency 8) Bigram 9) Linguistic

The data represented in the **Basic analyses** are directly based on the logging data via Inputlog. No additional interpretation occurred. You can rely on these analyses.

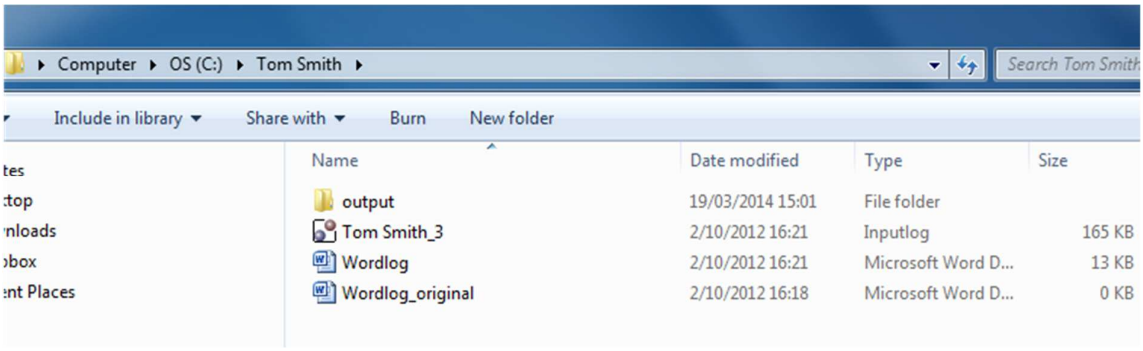
The **Advanced analyses** on the other hand are based on algorithmic processing of the raw logging data. These are based on theoretical and empirical findings, recent research, and discussions with experts in the field. Note that this implies that decisions have been made that might not be suitable for your research question. E.g., a sentence is based on the algorithm that a Sentence starts with a capital letter and ends with a full stop. Children and elderly do not comply with this algorithm. Be aware about this when reporting these results.

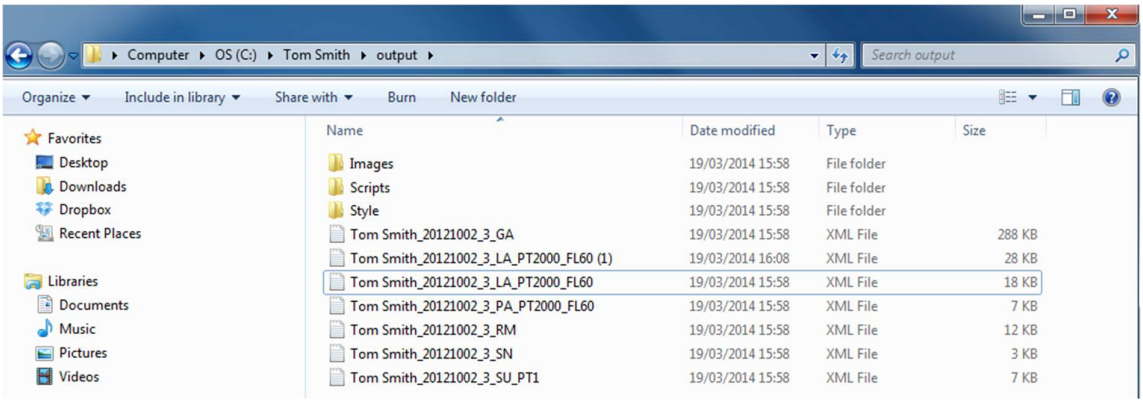
Original Word Document

Data Structure

Apart from and prior to the specific analyses that appear in the 'data output' section, Inputlog also produces a number of other files that do not appear in the data output section. These files are the source file (idf file) that prepares the data for further analyses.

A description of each type of file or analysis follows.





Datasets included in the example:

- Tom Smith_20121002_3_GA.xml: General analysis
- Tom Smith_20121002_3_SU_PT1.xml: Summary analysis (pause threshold 1 seconds)
- Tom Smith_20121002_3_PA_PT2000_FL60.xml: Pause analysis (fixed intervals 60 seconds)
- Tom Smith_20121002_3_LA_PT2000_FL60.xml: Linear analysis
- Tom Smith_20121002_3_LA_PT2000_FL60(1).xml: Linear analysis including condensed analyses
- Tom Smith_20121002_3_RM.xml: Revision matrix
- Tom Smith_20121002_3_SN.xml: S-notation

Each File begins with Firstname Lastname_date_number

Firstname_Lastname	Session identification details of participant
date	date of the logging: year - month - date

_number	consecutive number of the logging session under this session identification (Firstname Lastname)

The final part of the file name varies for each analyses.

_GA.xml	XML file in which every line represents an input action. See: **Error! Not a valid result for table.** GA General Analyses
_LA_PT2000_FL60.xml	XML file with the complete linear production of the text including the mouse movements and other activities of the writing process, divided into a number of periods of a fixed length. See: **Error! Not a valid result for table.** LA Linear Analyses PT2000 Pause Threshold = 2000 milliseconds FL60 Fixed length = 60 seconds You have the option to add a condensed analyses which only shows the keystrokes and no other activities.
_LA_PT2000_FN5.xml	XML file with the complete linear production of the text, divided into a fixed number of intervals. See: **Error! Not a valid result for table.** LA Linear Analyses PT2000 Pause Threshold = 2000 milliseconds FN5 Fixed number of intervals = 5
_PG_PT2000.PNG	Graphical file of the complete process graph. **Error! Not a valid result for table.** PG Progress Graph PT2000 Pause Threshold = 2000 milliseconds You can activate or disactivate the variables *;*;*.
_SU_PT2000.xml	XML file with an overview of basic statistics about the produced words and sentences, pausing behaviour, and the use of keyboard versus speech, amongst others. See: **Error! Not a valid result for table.** SA Summary Analyses PT2000 Pause Threshold = 2000 milliseconds FL60 Fixed length = 60 seconds
_PA_PT2000_FL60.xml	XML file with analyses of every non-scribal period. See: **Error! Not a valid result for table.**

	PA Pause Analyses PT2000 Pause Threshold = 2000 milliseconds FL60 Fixed length = 60 seconds
_PA_PT2000_FN5.xml	XML file with analyses of every non-scribal period of either 1, 2 and 5 seconds or any other - user-specified - pause length. See **Error! Not a valid result for table.** PA Pause Analyses PT2000 Pause Threshold = 2000 milliseconds FN5 Fixed number of intervals = 5
_SO.xml	XML file with analyses of the sources (focus events). See **Error! Not a valid result for table.** Analyses. SO Source Analyses
_SN	XML file with the S-notation of the writing process. See **Error! Not a valid result for table.**. SN S-notation
_RM	XML file with the Revision Matrix. RM Revision Matrix

Analyses

The **analyses section** enables you to specify the different analyses you want to generate. After you have selected the file and the type of analysis of your choice, click the 'Add' button. The file that can be generated for the output will then appear in the section 'data output'. A description of each possible analysis follows.

Each analysis begins with a section of Meta Information, Session Identification and Parameters.

Meta Information	The meta information consists of the following labels.	
	Logfile	Name of the idfx-file
	Log Creation	Date and time of the logging session
	Log GUID	A globally unique identifier or GUID is a special type of identifier used in software applications to provide a unique reference number. In other words, the Log GUID uniquely defines a logging session.
	Logging Program Version Number	Version number of Inputlog used to record the logging session (e.g., 5.1.0.19)
	Analysis Creation	Date and time of the analyses
	Analysis GUID	A globally unique identifier or GUID is a special type of identifier used in software applications to provide a unique reference number. In other words, the Log GUID uniquely defines a logging session.
	Analysis Program Version Number	Version number of Inputlog used to record the logging session (e.g., 5.1.0.21; this number may vary form the Logging Program Version Number)
Session Identification	The session identification contains all the information that you have entered in the Record Tab.	
	Participant	Give each participant a unique value (e.g., Mike=1, Janet=2, Chris=3).
	Text Language	Identify the language of the session.
	Age	Insert the age of the participant (value).
	Gender	Give a unique value per gender (e.g., male=0, female=1).
	Session	Indicate to which general session the particular logging session belongs (if necessary, e.g., session 1, session 2, etc.)
	Group	Give a unique value to each group (e.g., students=0, academics=1).
	Experience	Give a unique value to each experience level (e.g., low-proficiency=0, high-proficiency=1).
	Additional labels	You can add additional labels and their values

	(e.g., label = task, values 1 to 4).
Parameters	Each analysis shows the parameters that you have chosen for this particular analysis. The parameters will be further explained for each type of analyses.

General

The general analysis provides you with an XML file with a basic log file of the writing session in which every line represents an input action. The XML file can be converted to an Excel file or can be used as input for further analysis by for instance SPSS.

#Id	Event Type	Output	Position	DocLength	Character Production	StartTime	StartClock	EndTime	EndClock	ActionTime	PauseTime	PauseLocation	X	Y
0	focus	Wordlog.docx - Microsoft Word			0	4259	00:00:04	4259	00:00:04	0	4259	UNKNOWN PAUSE		
3	keyboard	D	0	1	1	8050	00:00:08	8175	00:00:08	234	7941	INITIAL PAUSE		
4	keyboard	e	1	2	2	8455	00:00:08	8580	00:00:08	125	405	WITHIN WORDS		
5	keyboard	m	2	3	3	8767	00:00:08	8861	00:00:08	94	312	WITHIN WORDS		
6	keyboard	o	3	4	4	8986	00:00:08	9157	00:00:09	171	219	WITHIN WORDS		
7	keyboard	n	4	5	5	9079	00:00:09	9251	00:00:09	172	93	WITHIN WORDS		
8	keyboard	s	5	6	6	9329	00:00:09	9516	00:00:09	187	250	WITHIN WORDS		
9	keyboard	t	6	7	7	9625	00:00:09	9688	00:00:09	63	296	WITHIN WORDS		
10	keyboard	r	7	8	8	9797	00:00:09	9891	00:00:09	94	172	WITHIN WORDS		
11	keyboard	a	8	9	9	9891	00:00:09	10047	00:00:10	156	94	WITHIN WORDS		
12	keyboard	t	9	10	10	9984	00:00:09	10140	00:00:10	156	93	WITHIN WORDS		
13	keyboard	i	10	11	11	10078	00:00:10	10265	00:00:10	187	94	WITHIN WORDS		
14	keyboard	o	11	12	12	10125	00:00:10	10296	00:00:10	171	47	WITHIN WORDS		
15	keyboard	n	12	13	13	10234	00:00:10	10421	00:00:10	187	109	WITHIN WORDS		
16	keyboard	SPACE	13	14	14	10421	00:00:10	10593	00:00:10	172	187	AFTER WORDS		
18	keyboard	I	14	15	15	10842	00:00:10	10920	00:00:10	156	343	BEFORE WORDS		
19	keyboard	n	15	16	16	11061	00:00:11	11217	00:00:11	156	219	WITHIN WORDS		
20	keyboard	p	16	17	17	11217	00:00:11	11357	00:00:11	140	156	WITHIN WORDS		
21	keyboard	u	17	18	18	11310	00:00:11	11404	00:00:11	94	93	WITHIN WORDS		
22	keyboard	t	18	19	19	11404	00:00:11	11529	00:00:11	125	94	WITHIN WORDS		

For every input action (letter, function, mouse click or movement, voice input) the session information is stored together with an identification (ID) of the input, the time stamps (in clock time and in seconds), the action time, the pausing time that followed it, and - for a mouse operation - the xy-value of the screen position.

32	mouse	LEFT Click			27	17488	00:00:17	17597	00:00:17	109	109	BEFORE SENTENCES	224	1029
33	focus	TASKBAR			27	17597	00:00:17	17597	00:00:17	0	0	UNKNOWN PAUSE		
34	mouse	Movement			27	17644	00:00:17	21591	00:00:21	3947	47	UNKNOWN PAUSE	639	519
35	focus	Naamloos - Google Chrome			27	17831	00:00:17	17831	00:00:17	0	0	UNKNOWN PAUSE		
36	focus	Google - Google Chrome			27	20358	00:00:20	20358	00:00:20	0	0	UNKNOWN PAUSE		
37	keyboard	i			27	22183	00:00:22	22293	00:00:22	110	592	UNKNOWN PAUSE		
38	keyboard	n			27	22402	00:00:22	22542	00:00:22	140	219	WITHIN WORDS		
39	keyboard	p			27	22542	00:00:22	22698	00:00:22	156	140	WITHIN WORDS		
40	keyboard	u			27	22667	00:00:22	22792	00:00:22	125	125	WITHIN WORDS		
41	keyboard	t			27	22823	00:00:22	22995	00:00:22	172	156	WITHIN WORDS		
42	keyboard	l			27	22932	00:00:22	23010	00:00:23	78	109	WITHIN WORDS		
43	keyboard	o			27	23088	00:00:23	23166	00:00:23	78	156	WITHIN WORDS		
44	keyboard	g			27	23166	00:00:23	23260	00:00:23	94	78	BEFORE SENTENCES		
45	keyboard	RETURN			27	23775	00:00:23	23868	00:00:23	93	609	AFTER SENTENCES		
46	mouse	Movement			27	24555	00:00:24	25272	00:00:25	717	780	UNKNOWN PAUSE	524	272
47	focus	inputlog - Google Search - Google Chrome			27	24555	00:00:24	24555	00:00:24	0	0	UNKNOWN PAUSE		
48	mouse	LEFT Click			27	25397	00:00:25	25491	00:00:25	94	125	BEFORE SENTENCES	524	272
49	mouse	Movement			27	25974	00:00:25	28486	00:00:28	2512	483	UNKNOWN PAUSE	438	244
50	focus	Inputlog - Google Chrome			27	25974	00:00:25	25974	00:00:25	0	0	UNKNOWN PAUSE		
51	mouse	LEFT Click			27	28720	00:00:28	28845	00:00:28	125	234	UNKNOWN PAUSE	438	244

ID	unique number of each event (consecutively) Note: Not all IDs from the IDFX are represented in the general analyses: e.g., mousemovements can be merged.
event type	All possible events that are logged:

<table>
<tr><td></td><td>keyboard
mouse
speech
focus
insert
replacement

Note: Focus, insert and replacement events do not have time stamps.</td></tr>
<tr><td>output</td><td>The output shows various types of output, related to the input

keyboard & speech Keystroke and dictated segment
mouse movements and clicks (with related XY value)
focus window name
insert text inserted from other source or other location in same document
replacement selection of text that can either be pasted at a different location, or that is replaced afterwards by new text (See Illustration "Inserts - replacements.xlsx")</td></tr>
<tr><td>position (Full)</td><td>Cursor position (starts at zero and starts counting while writing).
Consecutive writing shows contiguous number. Any interruption in the contiguous counting shows a backwards movement and is an indication of a revision.
Note: in the general analysis you can toggle this view.</td></tr>
<tr><td>doc length (Full)</td><td>Length of the document (starts at one)
Deletions etc. are substracted here
Note: you can toggle this view by clicking on the heading
<table>
<tr><th>#Id</th><th>Event Type</th><th>Output</th><th>Position</th><th>DocLength</th></tr>
<tr><td>0</td><td>focus</td><td>Wordlog.docx - Microsoft Word</td><td></td><td></td></tr>
<tr><td>3</td><td>keyboard</td><td>D</td><td>0</td><td>1</td></tr>
<tr><td>4</td><td>keyboard</td><td>e</td><td>1</td><td>2</td></tr>
<tr><td>5</td><td>keyboard</td><td>m</td><td>2</td><td>3</td></tr>
<tr><td>6</td><td>keyboard</td><td>o</td><td>3</td><td>4</td></tr>
</table>
<table>
<tr><th>#Id</th><th>Event Type</th><th>Output</th><th>Position Full</th><th>DocLength Full</th></tr>
<tr><td>0</td><td>focus</td><td>Wordlog.docx - Microsoft Word</td><td>0</td><td>0</td></tr>
<tr><td>3</td><td>keyboard</td><td>D</td><td>0</td><td>1</td></tr>
<tr><td>4</td><td>keyboard</td><td>e</td><td>1</td><td>2</td></tr>
<tr><td>5</td><td>keyboard</td><td>m</td><td>2</td><td>3</td></tr>
<tr><td>6</td><td>keyboard</td><td>o</td><td>3</td><td>4</td></tr>
</table>
</td></tr>
<tr><td>character Production</td><td>Characters produced (all keystrokes)</td></tr>
<tr><td>startTime</td><td>Time of key in: in milliseconds</td></tr>
<tr><td>startClock</td><td>Time of key in: in clock time</td></tr>
<tr><td>end Time</td><td>Time of key up: in milliseconds</td></tr>
<tr><td>endClock</td><td>Time of key up: in clock time</td></tr>
<tr><td>actionTime</td><td>Keyboard: Time between key in and key up: in milliseconds
Speech: Time between beginning of dictated segment and end of dictated segment</td></tr>
<tr><td>pauseTime</td><td>Time between key in and key in: in milliseconds</td></tr>
</table>

	Speech: Time between beginning of dictated segment and end of dictated segment
pauseLocation	The location of the pauses are subdivided into 10 categories. **within words** each pause within words **before words** each pause before a word **after words** each pause after a word Note: in the pause analyses after word and before word pauses are grouped together. **between sentences** each pause before a sentence **before sentences** each pause after a sentence **after sentences** Note: in the pause analyses after sentence and before sentence pauses are grouped together. **between paragraphs** each pause before a paragraph **before paragraphs** each pause after a paragraph **after paragraphs** Note: in the pause analyses after paragraph and before paragrap pauses are grouped together. **miscelaneous** **initial pause** ERIC **end pause** ERIC **unkown** ERIC
x value	Location of the mouse on x-axis
y value	Location of the mouse on y-axis

Note: The order of these variables may vary (In the future we strive to have a fixed order)

Note: Only the last two columns contain the values of the mouse. The locations of the keystrokes are not detailed.

Session identification

The session identification contains all the information that you have entered in the

Record. If you import your data in Excel or SPSS this information is shown in the first columns.

Output (labels and description)

In the column output you see the following labels.

label	action
BACK	backspace
DELETE	delete
RSHIFT / LSHIFT	Shift button: capital letter is visible in output
CAPS LOCK	capsloc: capitalization is visible in output
LCTRL / RCTRL	control + key
LALT	alt + key
UP	up: arrow up
DOWN	down: arrow down
LEFT	left: arrow to left
RIGHT	right: arrow to right
TAB	tab: indentation
ENTER	enter: Next line
HOME	home: beginning of document
END	end: end of document
PAGE UP	page up
PAGE DOWN	page down
INS	insert + character
Left Click	mouse click left
Movement	mouse movement (+ the start and end value of the xy-axis)
Scroll	scroll with the mouse scroll wheel & the Windows scrollbar + the start and end value of the xy-axis)
Right Click	mouse click right

General - condensed eyetrack

The general analysis provides you with an XML file with a basic log file of the writing session in which every line represents an input action (including eye tracking). The XML file can be converted to an Excel file or can be used as input for further analysis by for instance SPSS.

General - added with detailed event based eye tracking information

The general analyses is complemented with Tobii eye track information. Each Tobii event is inserted in the General analyses.

Event Type	Output	Position	DocLength	Character Production	StartTime	StartClock	EndTime	EndClock	ActionTime	PauseTime	PauseLocation	X	Y	Fixation Index	Saccade Index	Gaze Type	Gaze Event Duration
keyboard	T	0	1	2	18003	00:00:18	18050	00:00:18	78	983	WITHIN WORDS						
keyboard	h	1	2	2	18674	00:00:18	18720	00:00:18	46	671	WITHIN WORDS						
keyboard	r	2	3	3	18861	00:00:18	18892	00:00:18	31	187	WITHIN WORDS						
keyboard	e	3	4	4	18861	00:00:18	18892	00:00:18	31	0	WITHIN WORDS						
eyetrack	[Unclassified: 10 ms]			5	19539	00:00:19	19546	00:00:19	10	0	EYETRACK					Unclassified	10
eyetrack	[Saccade: 3 ms]			5	19549	00:00:19	19549	00:00:19	3	0	EYETRACK				83	Saccade	3
eyetrack	[Fixation: 253 ms]			5	19552	00:00:19	19803	00:00:19	253	0	EYETRACK			38		Fixation	253
eyetrack	[Saccade: 43 ms]			5	19806	00:00:19	19846	00:00:19	43	0	EYETRACK				84	Saccade	43
eyetrack	[Fixation: 133 ms]			5	19849	00:00:19	19979	00:00:19	133	0	EYETRACK			39		Fixation	133
eyetrack	[Saccade: 33 ms]			5	19982	00:00:19	20012	00:00:20	33	0	EYETRACK				85	Saccade	33
eyetrack	[Fixation: 93 ms]			5	20016	00:00:20	20106	00:00:20	93	0	EYETRACK			40		Fixation	93
eyetrack	[Saccade: 23 ms]			5	20109	00:00:20	20129	00:00:20	23	0	EYETRACK				86	Saccade	23
eyetrack	[Fixation: 237 ms]			5	20132	00:00:20	20366	00:00:20	237	0	EYETRACK			41		Fixation	237
eyetrack	[Saccade: 13 ms]			5	20369	00:00:20	20379	00:00:20	13	0	EYETRACK				87	Saccade	13
eyetrack	[Unclassified: 3 ms]			5	20382	00:00:20	20382	00:00:20	3	0	EYETRACK					Unclassified	3
eyetrack	[Saccade: 20 ms]			5	20386	00:00:20	20402	00:00:20	20	0	EYETRACK				88	Saccade	20

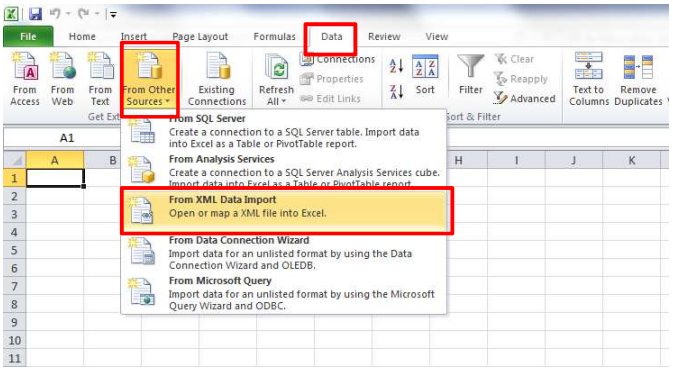
General - Condensed Eyetrack - general analyses with aggregated eye track information

The General - condensed eyetrack analysis shows the Tobii eye track information in an aggregated form. The keystoke log forms the basis and the eye track information is shown in an aggregated way. In the example below the writer writes the word "thre" and before he presses the "BACK" key the writer has 5 gaze events on the screen.

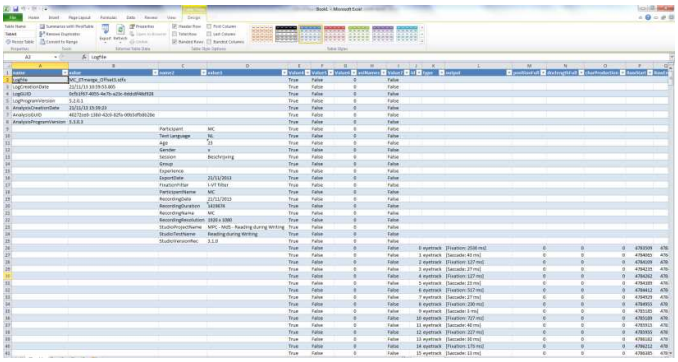
Output	Position	DocLength	Character Production	StartTime	StartClock	EndTime	EndClock	ActionTime	PauseTime	PauseLocation	X	Y	[FIX] Event Indices	[FIX] #GazeEvents	[FIX] Gaze Event Duration
LSHIFT	0	1	2	16989	00:00:16	12200	00:00:12	172	172	EYETRACK			37	1	140
T	0	1	2	18003	00:00:18	18050	00:00:18	78	983	WITHIN WORDS					
h	1	2	2	18674	00:00:18	18720	00:00:18	46	671	WITHIN WORDS					
r	2	3	3	18861	00:00:18	18892	00:00:18	31	187	WITHIN WORDS					
e	3	4	4	18861	00:00:18	18892	00:00:18	31	0	WITHIN WORDS					
BACK	4	5	5	21326	00:00:21	21388	00:00:21	62	2465	EYETRACK			38, 39, 40, 41, 42	5	883
BACK	3	4	6	21482	00:00:21	21544	00:00:21	62	156	REVISION					
e	2	3	6	21731	00:00:21	21809	00:00:21	78	249	WITHIN WORDS					
r	3	4	6	21965	00:00:21	22043	00:00:22	78	234	WITHIN WORDS					
e	4	5	7	22355	00:00:22	22402	00:00:22	47	390	WITHIN WORDS					
SPACE	5	6	8	22605	00:00:22	22698	00:00:22	93	250	AFTER WORDS					
i	6	7	9	22995	00:00:22	23088	00:00:23	93	390	BEFORE WORDS					
s	7	8	10	23307	00:00:23	23369	00:00:23	62	312	WITHIN WORDS					
SPACE	8	9	11	23463	00:00:23	23541	00:00:23	78	156	AFTER WORDS					
a	9	10	12	23697	00:00:23	23759	00:00:23	62	234	BEFORE WORDS					
SPACE	10	11	13	23837	00:00:23	23915	00:00:23	78	140	AFTER WORDS					
m	11	12	14	24430	00:00:24	24492	00:00:24	62	593	BEFORE WORDS					
a	12	13	15	24555	00:00:24	24664	00:00:24	109	125	WITHIN WORDS					
n	13	14	16	24929	00:00:24	24976	00:00:24	47	374	WITHIN WORDS					
SPACE	14	15	17	25163	00:00:25	25241	00:00:25	78	234	AFTER WORDS					

For every input action (letter, function, mouse click or movement, voice input, eye tracking) the session information is stored together with an identification of the input, the time stamps (in clock time and in seconds), the action time, the pausing time that followed it, and - for a mouse operation - the xy-value of the screen position. Note that the x and y values of Inputlog (for mouseclicks) correspond to the x and y values in Tobii (fixations and saccades).

Importing the merged data in an Excel file (see also import)



1. Open a new Excel file
2. Select the **data** menu
3. Select **from other sources**
4. Select **from XML data import**



5. Browse for the .xml General Analysis file you created from the merged idfx. You can find this in the output folder within the session folder.
6. Click **ok** to continue
7. Select cel **A1**
8. Click **ok** to finish
9. You now have your output in the excel file. The Colums A-I and rows 2-25 are session identifiers. The variables from column J onwards are described here below.
10. You can use the given variables or write formulas to create new variables to analyse the data.

	id	type	output	position	doclength	charProduction	RawStart	RawEnd	startTime	startClock	endTime	endClock	action Time	pause Time	pauseLocation	pauseLocationFull	FIX_gaze Event Type_nr	FIX_gaze Event Duration	fixationIndex	FIX_offscreen Time	FIX_nrOf Samples	FIX_nrOf Valid Samples
2	268	keyboard	T	0	1	2	3051036	3051083	18003	00:00:18.003	18050	00:00:18.050	78	983	1	WITHIN WORDS						
3	269	keyboard	h	1	2	2	3051707	3051753	18674	00:00:18.674	18720	00:00:18.720	46	671	1	WITHIN WORDS						
4	270	keyboard	r	2	3	3	3051894	3051925	18861	00:00:18.861	18892	00:00:18.892	31	187	1	WITHIN WORDS						
5	271	keyboard	e	3	4	4	3051894	3051925	18861	00:00:18.861	18892	00:00:18.892	31	0	1	WITHIN WORDS						
6	292	keyboard	BACK	4	5	5	3054359	3054421	21326	00:00:21.326	21388	00:00:21.388	62	2465	13	EYETRACK	5	883	38, 39, 40, 41, 42	403	265	146
7	293	keyboard	BACK	3	4	6	3054515	3054577	21482	00:00:21.482	21544	00:00:21.544	62	156	11	REVISION						
8	294	keyboard	e	2	3	6	3054764	3054842	21731	00:00:21.731	21809	00:00:21.809	78	249	1	WITHIN WORDS						
9	295	keyboard	r	3	4	6	3054998	3055076	21965	00:00:21.965	22043	00:00:22.043	78	234	1	WITHIN WORDS						
10	296	keyboard	e	4	5	7	3055388	3055435	22355	00:00:22.355	22402	00:00:22.402	47	390	1	WITHIN WORDS						
11	297	keyboard	SPACE	5	6	8	3055638	3055731	22605	00:00:22.605	22698	00:00:22.698	93	250	3	AFTER WORDS						
12	298	keyboard	i	6	7	9	3056028	3056121	22995	00:00:22.995	23088	00:00:23.088	93	390	2	BEFORE WORDS						
13	299	keyboard	s	7	8	10	3056340	3056402	23307	00:00:23.307	23369	00:00:23.369	62	312	1	WITHIN WORDS						
14	300	keyboard	SPACE	8	9	11	3056496	3056574	23463	00:00:23.463	23541	00:00:23.541	78	156	3	AFTER WORDS						
15	301	keyboard	a	9	10	12	3056730	3056792	23697	00:00:23.697	23759	00:00:23.759	62	234	2	BEFORE WORDS						
16	302	keyboard	SPACE	10	11	13	3056870	3056948	23837	00:00:23.837	23915	00:00:23.915	78	140	3	AFTER WORDS						
17	303	keyboard	m	11	12	14	3057463	3057525	24430	00:00:24.430	24492	00:00:24.492	62	593	2	BEFORE WORDS						
18	304	keyboard	a	12	13	15	3057588	3057697	24555	00:00:24.555	24664	00:00:24.664	109	125	1	WITHIN WORDS						
19	305	keyboard	n	13	14	16	3057962	3058009	24929	00:00:24.929	24976	00:00:24.976	47	374	1	WITHIN WORDS						
20	306	keyboard	SPACE	14	15	17	3058196	3058274	25163	00:00:25.163	25241	00:00:25.241	78	234	3	AFTER WORDS						
21	307	keyboard	s	15	16	18	3058836	3058930	25803	00:00:25.803	25897	00:00:25.897	94	640	2	BEFORE WORDS						
22	308	keyboard	l	16	17	19	3059148	3059195	26115	00:00:26.115	26162	00:00:26.162	47	312	1	WITHIN WORDS						
23	309	keyboard	e	17	18	20	3059351	3059413	26318	00:00:26.318	26380	00:00:26.380	62	203	1	WITHIN WORDS						
24	310	keyboard	a	18	19	21	3059585	3059725	26552	00:00:26.552	26692	00:00:26.692	140	234	1	WITHIN WORDS						
25	311	keyboard	p	19	20	22	3060458	3060521	27425	00:00:27.425	27488	00:00:27.488	63	873	1	WITHIN WORDS						
26	312	keyboard	e	20	21	23	3061706	3061769	28673	00:00:28.673	28736	00:00:28.736	63	1248	1	WITHIN WORDS						
27	313	keyboard	e	21	22	24	3061862	3061925	28829	00:00:28.829	28892	00:00:28.892	63	156	1	WITHIN WORDS						
28	327	keyboard	BACK	22	23	25	3064608	3064655	31575	00:00:31.575	31622	00:00:31.622	47	2746	13	EYETRACK	2	307	43, 44	0	92	92
29	328	keyboard	BACK	21	22	26	3064764	3064811	31731	00:00:31.731	31778	00:00:31.778	47	156	11	REVISION						
30	329	keyboard	i	20	21	26	3065388	3065450	32355	00:00:32.355	32417	00:00:32.417	62	624	1	WITHIN WORDS						
31	330	keyboard	n	21	22	26	3065591	3065669	32558	00:00:32.558	32636	00:00:32.636	78	203	1	WITHIN WORDS						
32	331	keyboard	g	22	23	27	3065981	3066043	32948	00:00:32.948	33010	00:00:33.010	62	390	1	WITHIN WORDS						
33	332	keyboard	SPACE	23	24	28	3066121	3066199	33088	00:00:33.088	33166	00:00:33.166	78	140	3	AFTER WORDS						

Note: the variables in grey are only shown when you import the data from the merged General Analysis (GA) in an Excel file. They are not represented in the GA quick view in Internet Explorer. These grey items are merely included for your information and to be thorough. Moreover, the order and variable names may slightly differ between the quick view and the imported data in Excel. This file is based on the order and variable names in the Excel file.

ID	Unique number of each event (consecutively) Note: Not all IDs from the IDFX are represented in the general analyses: e.g., mousemovements can be merged.
event type	All possible events that are logged: **keyboard** **mouse** **speech** **focus** **insert** **replacement** **eyetrack** Note: Focus, insert and replacement events do not have time stamps.
output	The output shows various types of output, related to the input **keyboard & speech** Keystroke and dictated segment **mouse** movements and clicks (with related XY values) **focus** window name **insert** text inserted from other source or other location in same document **replacement** selection of text that can either be pasted at a different location, or that is replaced afterwards by new text (See Illustration "Inserts - replacements.xlsx") **eyetrack** fixations, saccades, unclassified
position (Full)	Cursor position (starts at zero and starts counting while writing). Consecutive writing shows contiguous number. Any interruption in the contiguous counting shows a backwards movement and is an indication of a revision. Note: an in the generel analyse you can toggle this view.
doc length (Full)	Length of the document (starts at one) Deletions etc. are substracted from the document length. It is the product length at each moment. Note: you can toggle this view
charproduction	Characters produced (all keystrokes)
rawstart	StartTime of Tobii eye event

rawend	EndTime of Tobii eye event
starttime	Time of key in: in milliseconds
startclock	Time of key in: in clock time
end time	Time of key up: in milliseconds
endclock	Time of key up: in clock time
actiontime	Keyboard: Time between key in and key up: in milliseconds Speech: Time between beginning of dictated segment and end of dictated segment Eyetrack:
pausetime	Time between key in and key in: in milliseconds Speech: Time between beginning of dictated segment and end of dictated segment Eyetrack:
pauselocation	The location of the pauses are subdivided into 10 categories. **within words** — each pause within words **before words** — each pause before a word **after words** — each pause after a word Note: in the pause analyses after word and before word pauses are grouped together. **between sentences** — each pause before a sentence **before sentences** — each pause after a sentence **after sentences** — Note: in the pause analyses after sentence and before sentence pauses are grouped together. **between paragraphs** — each pause before a paragraph **before paragraphs** — each pause after a paragraph **after paragraphs** — Note: in the pause analyses after paragraph and before paragrap pauses are grouped together. **miscelaneous** **initial pause** **end pause** **transition pause** **revision pause** **unkown**
pauselocationfull	Label of the pause categorie
et_endtime	End time of Tobii eye event
et_fulltimestamp	Real clock time of eye event
et_starttime	Start time of Tobii eye event
fixationindex	Represents the order in which a fixation event was recorded. It has a unique

	number of each (consecutive) fixation event. This index is an auto-increment number starting with 1 (first gaze event detected). zie tobii - based op Filter (zie FixationFilter in de header information) **Session Identification** Participant [redacted] Text Language EN Age 84 Gender male Group p2 ExportDate 26/02/2014 FixationFilter I-VT filter ParticipantName [redacted] RecordingDate 25/02/2014 RecordingDuration 905976 RecordingName [redacted] RecordingResolution 1920 x 1080 StudioProjectName ML - MPC - ALZh StudioTestName FWO alzheimer StudioVersionRec 3.1.0
saccadeIndex	Represents the order in which a saccade event was recorded. It has a unique number of each (consecutive) saccade event. This index is an auto-increment number starting with 1 (first gaze event detected).
gazeeventtype	Type of eye movement event classified by the fixation Filter settings applied during the gaze data export (i.e., fixation; saccade; unclassified).
gazeeventduration	Duration of an eye movement event in milliseconds (start time of a gaze event, until start of the Next gaze event)
AverageGazePntX_ ADCSpx	Average gaze point X (mcspx) of Tobii calculated as a horizontal distance from the left upper corner of the screen. A score of 0 concerns the most left upper corner of the screen. Note: Calculated on samples with a validity below 2 (2 and higher is calculated as off-screen time 2)
AverageGazePntY_ ADCSpx	Average gaze point X (mcspx) of Tobii calculated as a vertical distance from the left upper corner of the screen. A score of 0 concerns the most left upper corner of the screen. Note: Calculated on samples with a validity below 2 (2 and higher is calculated as off-screen time 2) Explanation of **validity** by Tobii The validity code takes one of five values for each eye ranging from 0 to 4, with the following interpretation:

	0 – The eye tracker is certain that the data for this eye is correct. There is no risk of confusing data from the other eye. 1 – The eye tracker has only recorded one eye, and has made some assumptions and estimations regarding which is the left and which is the right eye. However, it is still very likely that the assumptions made are correct. The validity code for the other eye is in this case always set to 3. 2 – The eye tracker has only recorded one eye, and has no way of determining which is the left eye and which is the right eye. The validity code for both eyes is set to 2. 3 – The eye tracker is fairly confident that the actual gaze data belongs to the other eye. The other eye will always have validity code 1. 4 – The gaze data is missing or definitely belongs to the other eye. Validity codes should be used for data Filtering to remove data points that are obviously incorrect. If you export the raw data file, we recommend removing all data points with a validity code of 2 or higher. (manual - www.tobii.com)
AveragePupilLeft	Average of the estimated size (in millimeters) of the left eye pupil of **valid samples.** Higher estimated size indicates higher cognitive load.
AveragePupilRight	Average of the estimated size (in millimeters) of the right eye pupil of **valid samples.** Higher estimated size indicates higher cognitive load.
AverageValidityLeft	Indicates the average confidence level of **all samples taken together** that the left eye has been correctly identified. The values rage from 0 (high confidence) to 4 (eye not found) (i.e. the higher the value, the lower the validity). (See also matrix below)
AverageValidityRight	Indicates the average confidence level of **all samples** that the right eye has been correctly identified. The values rage from 0 (high confidence) to 4 (eye not found) (i.e. the higher the value, the lower the validity).
OffscreenTime	Total duration of all samples with a validity score of 2 and higher.
NrOfSamples	Total number of samples (lowest level of data collection, before Filters have been adapted) in gaze-event
NrOfValidSamples	Total number of valid samples (with a validity score of 2 and lower) in gaze-event
MinGazePointX_ ADCSpx	Horizontal coordinate of the averaged left and right eye gaze point on the screen (Min = 0; left upper corner of the screen).
MinGazePointX_ MCSpx	Horizontal coordinate of the averaged left and right eye gaze point on the media element (Min = 0; left upper corner of the media element). Column is empty if: • Fixation is outside media • Media is covered • No media is displayed

MinGazePointY_ ADCSpx	Vertical coordinate of the averaged left and right eye gaze point on the screen (Min = 0; left upper corner of the screen).
MinGazePointY_ MCSpx	Vertical coordinate of the averaged left and right eye gaze point on the media element (Min = 0; left upper corner of the media element). Column is empty if: • Fixation is outside media • Media is covered • No media is displayed
MaxGazePointX _ ADCSpx	
MaxGazePointX_ MCSpx	
MaxGazePointY_ ADCSpx	
MaxGazePointY_ MCSpx	
MaxDistanceX	Maximum absolute distance between two gazepoints on horizontal axis (left to right). Based on ADCSpx. **Note:** This concerns the distance within one fixation and not between fixations. In order to calculate the distance between two consecutive fixations, the variable MinGazePointX_ ADCSpx could be used. This variable also allows to study regressions during reading. It might however be relevant to set a threshold for what could be considered a regression (for example a regression should be at least 6 pixels backwards on the X-axis (6 pixels corresponds to an average of one character when using Calibri 14pt.
MaxDistanceY	Maximum absolute distance between two gazepoints on vertical axis (top to bottom)
DistanceX	Distance on the horizontal axis from left to right (from the first sample event to the last sample event) = positive value Distance from right to left = negative value Calculated as the End GazePntX ADCSpx minus the Start GazePntX ADCSpx
DistanceY	Distance on the vertical axis from top to bottom (from the first sample event to the last sample event) = positive value From bottom to top = negative value Calculated as the End GazePntY ADCSpx minus the Start GazePntY ADCSpx
StartGazePointX_ ADCSpx	The first gazepoint on the horizontal axis of a new gaze
StartGazePointY_ ADCSpx	The first gazepoint on the vertical axis of a new gaze
EndGazePointX_	The last gazepoint on the horizontal axis of a gaze

ADCSpx	
EndGazePointY_ ADCSpx	The last gazepoint on the vertical axis of a gaze
CumAbsDistanceX	Sum of absolute distance values on the horizontal axis including all valid samples (negative values are transformed to positive values).
CumAbsDistanceX_ Left	Sum of absolute distance value of movements towards left
CumAbsDistanceX_ Right	Sum of absolute distance value of movements towards right
CumAbsDistanceY	Sum of absolute distance values on the vertical axis including all valid samples (negative values are transformed to positive values).
CumAbsDistanceY_ Down	Sum of absolute distance value of movements towards bottom
CumAbsDistanceY_ Up	Sum of absolute distance value of movements towards top
DistanceLeft_Max	Maximum distance of the left eye to the screen.
DistanceLeft_Max	Minimum distance of the left eye to the screen
DistanceRight_Max	Maximum distance of the right eye to the screen.
DistanceRight_Min	Minimum distance of the right eye to the screen
EyePosLeftX_Max	Maximum horizontal coordinate of the 3D position of the left eye measured in millimeters
EyePosLeftX_Min	Minimum horizontal coordinate of the 3D position of the left eye measured in millimeters
EyePosLeftY_Max	Maximum vertical coordinate of the 3D position of the left eye measured in millimeters
EyePosLeftY_Min	Minimum vertical coordinate of the 3D position of the left eye measured in millimeters
EyePosLeftZ_Max	Distance/depth coordinate of the 3D position of the left eye measured in millimeters
EyePosLeftZ_Min	
EyePosRightX_Max	Horizontal coordinate of the 3D position of the right eye measured in millimeters
EyePosRightX_Min	
EyePosRightY_Max	Vertical coordinate of the 3D position of the right eye measured in millimeters
EyePosRightY_Min	
EyePosRightZ_Max	Distance/depth coordinate of the 3D position of the right eye measured in

	millimeters
EyePosRightZ_Min	
StudioEvent	Type of media or manual logging event (e.g., start, end and manual logging description).
StudioEventValue	Title of the studio event
KeyboardEvents	Keyboard event: displays the information of which key was pressed on the keyboard. (comparable to Inputlog keyboard events)
KeyboardNumerOf Events	Number of key events
MediaName	Name of the media/stimuli element from the Tobii Studio test timeline.
MediaPosX_ADCSpx	Horizontal coordinate (measured in pixels) of the left edge of the eye tracked media.
MediaPosY_ADCSpx	Vertical coordinate (measured in pixels) of the top edge of the eye tracked media.
MediaHeight	Vertical size (e.g., height measured in pixels) of the eye tracked media.
MediaWidth	Horizontal size (e.g., width measured in pixels) of the eye tracked media.
MouseEvents	Mouse click type during gaze-event: Left button and Right button
MouseNumberOf Events	Number of mouse events during gaze-events
x	Location of the mouse on the screen's x-axis
y	Location of the mouse on the screen's y-axis
Position	Cursor position (starts at zero) Cursor position (starts at zero and starts counting while writing). Consecutive writing shows contiguous number. Any interruption in the contiguous counting shows a backwards movement and is an indication of a revision. This is the original output of position: it can be toggled.
Doclength	Length of the document (starts at one) Deletions etc. are substracted from the document length. It is the product length at each moment. This is the original output of DocLength: it can be toggled.

Output (labels and description)

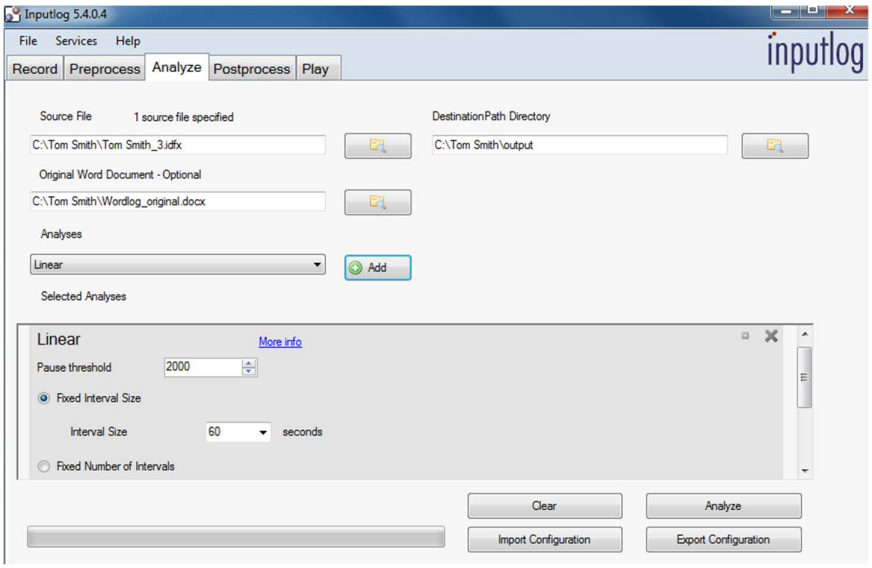
Additional output of the eye tracking information

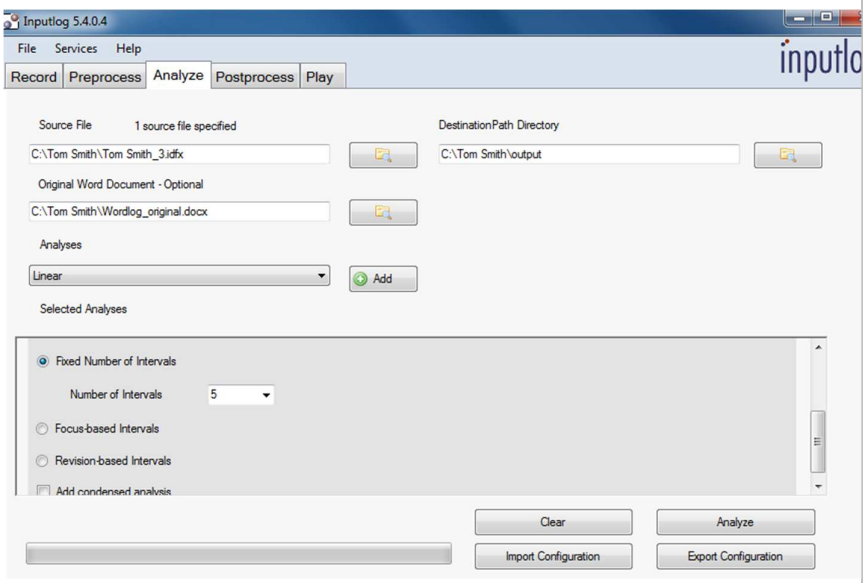
label	**action**

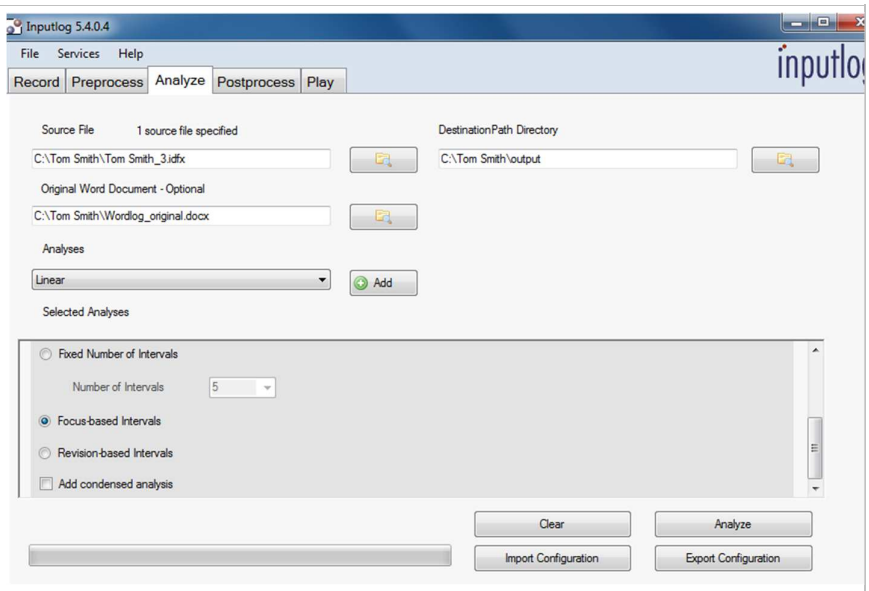
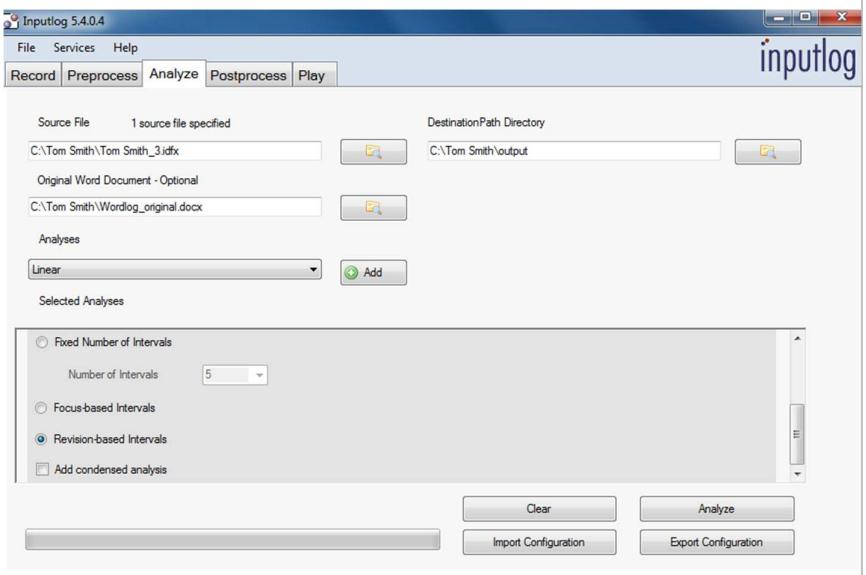
to be added	

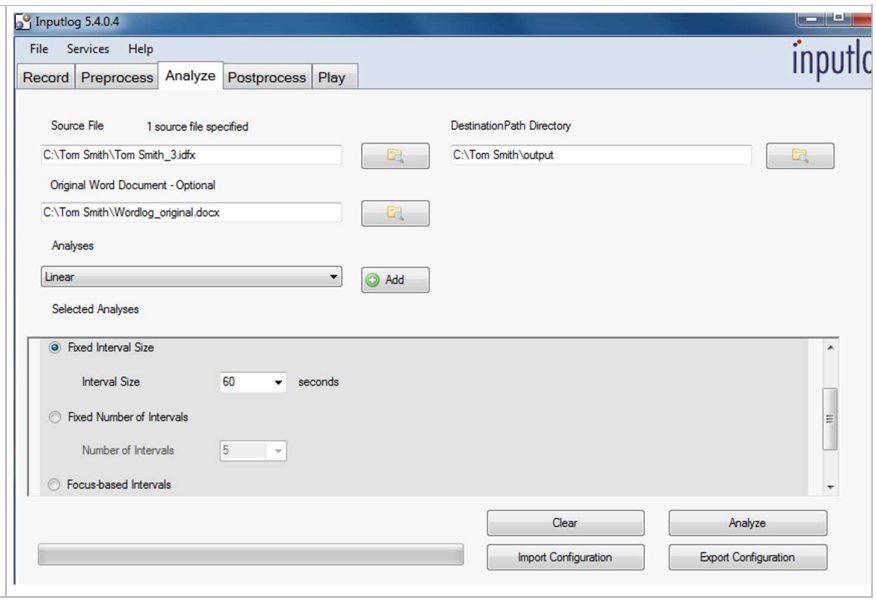
Linear

The linear analysis provides you with an XML file with a basic log file of the whole writing session in which all the lines represent an input action. The linear logging file has a pause threshold which you can choose and five other options you can select.



fixed interval size	XML file with the complete linear production of the text including the mouse movements and other activities of the writing process, divided into a number of periods of a fixed length. You can set a predefined interval size with 20, 30, 60 or 120 seconds, or you can select any other interval size.
fixed number of intervals	You can choose the number of intervals from 1 to 10. 
focus-based intervals	The output will be presented based on the sources of the writing session.

	
revision-based intervals	The linear output will be presented based on the revisions. 
condensed analysis	The output is a compact version of the produced text without the revisions and other keyboard information.


Inputlog 5.4.0.4
File
Services
Help
Record
Preprocess
Analyze
Postprocess
Play
Source File
1 source file specified
C:\Tom Smith\Tom Smith_3.idfx
DestinationPath Directory
C:\Tom Smith\output
Original Word Document - Optional
C:\Tom Smith\Wordlog_original.docx
Analyses
Linear
Add
Selected Analyses
Fixed Interval Size
Interval Size
60
seconds
Fixed Number of Intervals
Number of Intervals
5
Focus-based Intervals
Clear
Analyze
Import Configuration
Export Configuration

Output (labels and description)

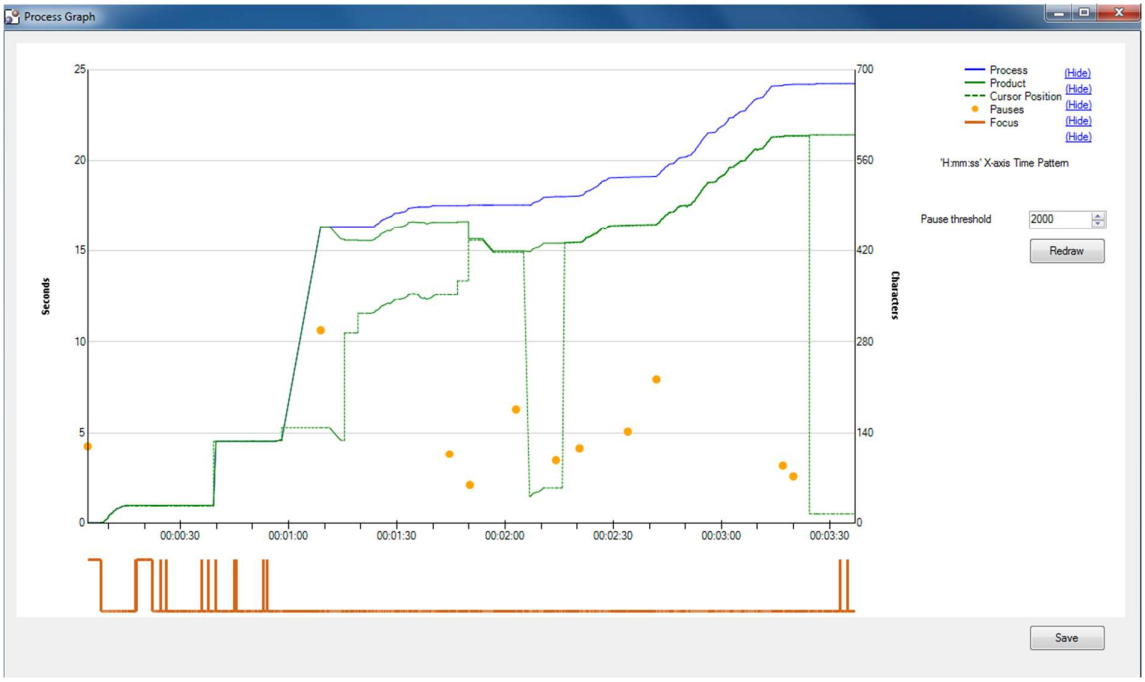
label	action
BACK	backspace
DELETE	delete
RSHIFT / LSHIFT	Shift button: capital letter is visible in output
CAPS LOCK	capslock: capitalization is visible in output
LCTRL / RCTRL	control + key
LALT	alt + key
UP	up: arrow up
DOWN	down: arrow down
LEFT	left: arrow to left
RIGHT	right: arrow to right
TAB	tab: indentation
ENTER	enter: Next line
HOME	home: beginning of document
END	end: end of document
PAGE UP	page up
PAGE DOWN	page down
INS	insert + character
Left Click	mouse click left
Movement	mouse movement (+ the start and end value of the xy-axis)
Scroll	scroll with the mouse scroll wheel & the Windows scrollbar + the start and end value of the xy-axis)
Right Click	mouse click right

Process Graph

variable	description
process	Total text production (or **Character Production**) at each instance of the writing process
pauses	Aggregated pauses
product	The document length of the text at each instance of the writing process
focus	The program (or focus) the writer uses is shown below the time-line (in orange): 1. lower line: word document 2. top line: other sources The vertical lines indicate where the writer switches between the sources.
position	Cursor (caret) Position of the writer at each instance of the writing process (indicating revisions)
pause threshold (ms)	Only pauses exceeding the Pause threshold will be shown in the Process Graph
save as	You can save the file in various formats PNG BMP EMF JPG TIFF

After you have generated the process graph you can interact with the figure (via Hide and Show). The process graph is also automatically saved as a picture.

The process graph of Tom Smith is as follows.



The progression graph plots the time-based progression of a writing process (x-axis) against the number of characters produced (right y-axis). The left y-axis shows the length of the pauses (orange dots).

For a description see Leijten, M., & Van Waes, L. (2013). Keystroke Logging in Writing Research: Using Inputlog to Analyze and Visualize Writing Processes. Written Communication, 30(3), 35. doi: 10.1177/0741088313491692).

Summary

The summary logging file has five main sections. You can choose the Pause Treshold as a parameter for these analyses.

process information	General information about the writing process: e.g., number of words produced during the writing process.
product information	General information about the writing product: e.g., number of words in the final product
product/process	The relation between product and process information
process time	Information related to writing and pausing times.
writing mode	Information related to the writing modes (especially interesting when using Speech Recognition).

The next table shows an explanation of all variables per section.

Process information

Characters	
Total	Total number of characters produced (both typed and copied) in the writing process
Total copied	Total number of characters copied within the Wordlog.docx and copied from other sources.
Total typed (incl. spaces)	Total number of characters typed including spaces within the Wordlog.docx and in other sources.
Per minute (incl. spaces)	Total number of characters typed including spaces within the Wordlog.docx and in other sources / length of the writing process in minutes
Total typed (excl. spaces)	Total number of characters typed excluding spaces within the Wordlog.docx and in other sources.
Per minute (excl. spaces)	Total number of characters typed excluding spaces within the Wordlog.docx and in other sources divided by the length of the writing process in minutes
Words	
Total	Total number of words produced during the writing process
Per minute	Total number of words produced during the writing process divided by the length of the writing process in minutes
Mean Word Length	Mean length of the words produced
St. Dev. Word length	Standard deviation of the length of the words produced
Sentences	
Total	Total number of sentences produced during the writing process

Mean Characters/sentence	Total number of characters produced during the writing process divided by the number of sentences
St. Dev. Characters/sentence	Standard deviation of the number of characters produced per sentence
Mean Words/sentence	Total number of words produced during the writing process divided by the number of sentences
St. Dev. Words/sentence	Standard deviation of the number of words produced per sentence
Paragraphs	
Total	Total number of paragraphs produced during the writing process
Mean Characters/paragraph	Total number of characters produced during the writing process divided by the number of paragraphs
St. Dev. Characters/ paragraphs	Standard deviation of the number of characters produced per paragraph
Mean Words/ paragraphs	Total number of words produced during the writing process divided by the number of paragraphs
St. Dev. Words/ paragraphs	Standard deviation of the number of words produced per paragraph

Note: If you export this information to Excel/SPSS the labels are transformed so each variable has a unique name: e.g., Characters total, Words total.

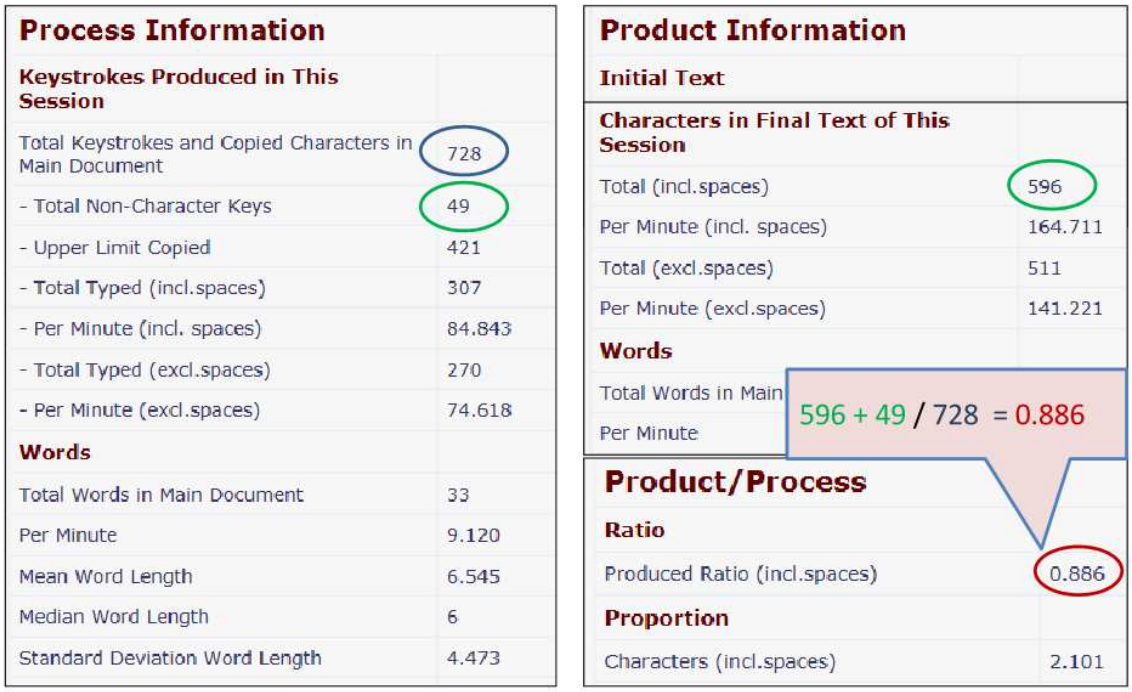
Product information

Characters	
Total (incl. spaces)	Total number of characters (including spaces) in the final text
Per minute (incl. spaces)	Total number of characters (including spaces) in the final text divided by the length of the writing process in minutes
Total (excl. spaces)	Total number of characters (excluding spaces) in the final text
Per minute (excl. spaces)	Total number of characters (excluding spaces) in the final text divided by the length of the writing process in minutes
Words	
Total	Total number of words in the final text
Per minute	Total number of words in the final text divided by the length of the writing process in minutes
Paragraphs	
Total	Total number of paragraphs in the final text

Lines	
Total	Total number of lines in the final text
Pages	
Total	Total number of pages in the final text

Product/process

Ratio	
Produced ratio (incl. spaces)	Sum of total number of characters in the final text (incl. spaces) and the total number of non-character keys, divided by the total number of characters **produced** during the writing process Writers generally produce more words during the process of writing than what appears in the final text. The Ratio calculation takes this difference into account. Remark: ▪ f this number is 1, no deletion has taken place. The lower the number, the bigger the difference between process and final product thus indicating greater amounts of deleted text. ▪ The non-character keys (like Enter) are added to the product calculation so as to align the character count in the product and the process calculation.
Proportion	
Characters (incl. spaces)	Total number of characters in the final text (including spaces) divided by the total number of characters **typed** during the writing process. Remark: In the calculation of the proportion the amount of copied text is not taken into account. Therefore, the proportion can be greater than 1; the larger the number, the more text that has been copied.
Characters (excl. spaces)	Total number of characters in the final text (excluding spaces) divided by the total number of characters typed during the writing process
Words	Total number of words in the final text divided by the total number of words during the writing process



Process Information

Keystrokes Produced in This Session	
Total Keystrokes and Copied Characters in Main Document	728
- Total Non-Character Keys	49
- Upper Limit Copied	421
- Total Typed (incl.spaces)	307
- Per Minute (incl. spaces)	84.843
- Total Typed (excl.spaces)	270
- Per Minute (excl.spaces)	74.618
Words	
Total Words in Main Document	33
Per Minute	9.120
Mean Word Length	6.545
Median Word Length	6
Standard Deviation Word Length	4.473

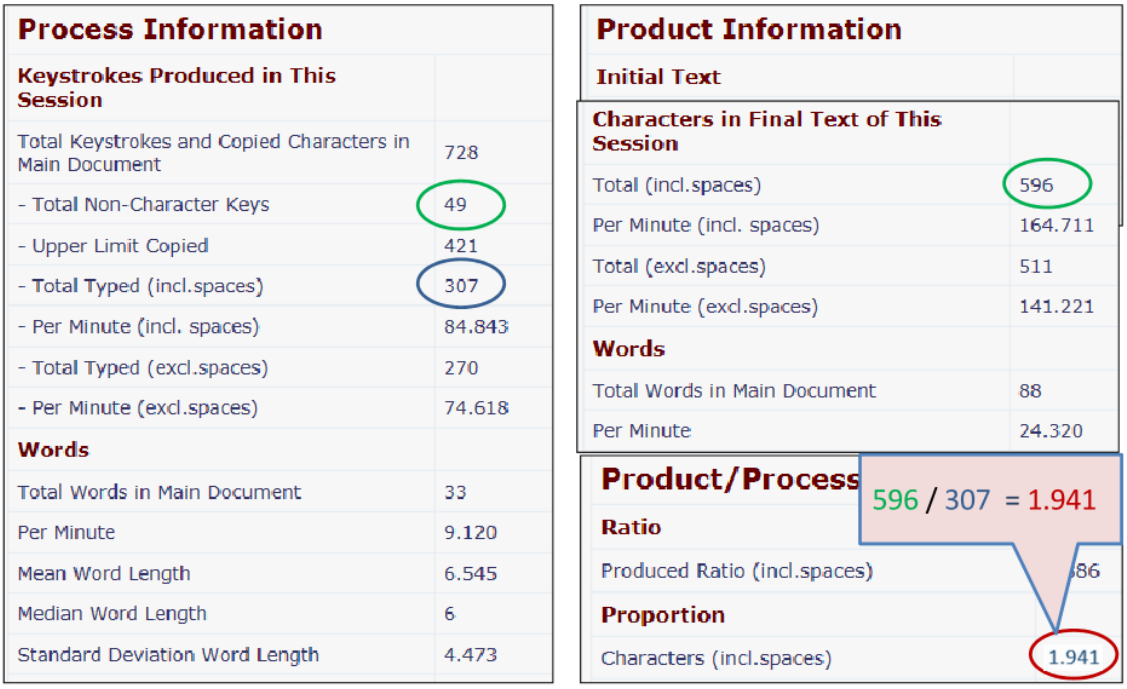
Product Information

Initial Text	
Characters in Final Text of This Session	
Total (incl.spaces)	596
Per Minute (incl. spaces)	164.711
Total (excl.spaces)	511
Per Minute (excl.spaces)	141.221
Words	
Total Words in Main	
Per Minute	

596 + 49 / 728 = 0.886

Product/Process

Ratio	
Produced Ratio (incl.spaces)	0.886
Proportion	
Characters (incl.spaces)	2.101



Process Information

Keystrokes Produced in This Session	
Total Keystrokes and Copied Characters in Main Document	728
- Total Non-Character Keys	49
- Upper Limit Copied	421
- Total Typed (incl.spaces)	307
- Per Minute (incl. spaces)	84.843
- Total Typed (excl.spaces)	270
- Per Minute (excl.spaces)	74.618
Words	
Total Words in Main Document	33
Per Minute	9.120
Mean Word Length	6.545
Median Word Length	6
Standard Deviation Word Length	4.473

Product Information

Initial Text	
Characters in Final Text of This Session	
Total (incl.spaces)	596
Per Minute (incl. spaces)	164.711
Total (excl.spaces)	511
Per Minute (excl.spaces)	141.221
Words	
Total Words in Main Document	88
Per Minute	24.320

596 / 307 = 1.941

Product/Process

Ratio	
Produced Ratio (incl.spaces)	86
Proportion	
Characters (incl.spaces)	1.941

What is an M-Burst? (See PhD Leijten p. 187 - Figure 8)

A M-burst is defined as a series of actions within the same writing mode, i.c. keystroke input that is not interrupted by another writing mode (mouse, speech).

Pause

The pause logging file has three main sections.

General information	General information about the pausing behavior during the writing process: e.g., number of pauses, total pause time.
Pause location	General information about the location of the pauses (within words, between words, between sentences, between paragraphs, etc.)
Summary per interval/ Summary per period	Summarized information per chosen parameter: - divided in a fixed number of intervals (e.g., 5 intervals, 10 intervals) - divided in periods of a fixed length (e.g., 60 seconds, 5 minutes)

The pause locations are subdivided in the following categories. You can find the definitions of words, sentences and paragraph in the glossary.

Within words	Each pause within a word.
after words	Each pause after a word.
before words	Each pause before a word.
between words	A consecutive combination of a pause labeled 'after words' and 'before words' is calculated as 'between words' Note: the pauses are related to the chosen threshold in ms. The number of before word and after word pauses do not automatically add up to the number of between words. E.g., if you choose a pause threshold of 1000ms the pauses of 748 (SPACE) and 453 (I) will not be taken into account. In the analyses of between word pauses this pause will be counted: 748+453=1201ms.

output	start Time	end Time	action Time	pause Time	pauseLocationFull
SPACE	16099	16240	141	93	AFTER WORDS
a	16833	17020	187	734	BEFORE WORDS
n	22605	22792	187	5772	WITHIN WORDS
SPACE	23353	23541	188	748	AFTER WORDS
i	23806	23915	109	453	BEFORE WORDS
l	24180	24258	78	374	WITHIN WORDS
l	24352	24399	47	172	WITHIN WORDS
u	24617	24726	109	265	WITHIN WORDS
s	24835	24991	156	218	WITHIN WORDS
t	24991	25101	110	156	WITHIN WORDS
r	25397	25553	156	406	WITHIN WORDS
a	25600	25740	140	203	WITHIN WORDS
t	25709	25849	140	109	WITHIN WORDS

	i	25849	26005	156	140	WITHIN WORDS
	o	25943	26177	234	94	WITHIN WORDS
	n	26083	26255	172	140	WITHIN WORDS
after sentences	Each pause after a sentence.					
before sentences	Each pause before a sentence.					
between sentences	A consecutive combination of a pause labeled 'after sentence' and 'before sentence' is calculated as 'between sentences'					
after paragrahs	Each pause after a paragraph.					
before paragrahs	Each pause before a paragraph.					
between paragrahs	A consecutive combination of a pause labeled 'after paragraph and 'before paragraph' is calculated as 'between paragrahs'					
initial pause	to be added (these rules are being reprogrammed)					
end pause	to be added					
unknown pauses	to be added					
***	to be added					
***						

Fixed Interval Size

Fixed Number of Intervals

Revision

To define revisions we have developed an algorithm and a set of rules. The revision analysis first of all defines critical process events in the writing process that can be linked to revision and then evaluates these instances by comparing the operations in the isolated writing episode to the revision rules in the algorithm. Inputlog successively analyses the beginning of the revision, the selection of the text involved in the revision, or the positioning of the cursor, the (possible) deletion/insertion of the text and the end of the revision.

Inputlog offers you two possibilities to analyse the participant's revision behaviour:

1. S-Notation: a formal notation system to represent the successive text editing actions;
2. Revision Matrix: a linear representation in which the revisions in a log file are listed, together with some basic time and position based characteristics.

S-notation

The S-notation is a computer-based method for tracing the writing process. It is first fully described by Py Kollberg and Kerstin Eklundh from the Swedish Royal Institute of Technology in 1997 (http://www.nada.kth.se/iplab/trace-it/S-notation.html). They based their description on an informal notation for revisions in handwriting published 10 years earlier by Matsuhashi (1987). Insertions, deletions, and breaks are numbered according to the order of their occurrence in the writing process. Therefore, the S-notation provides a complete record of the writer's revisions. It allows describing writing as a decomposed linear process. One can see at what point in the text a decision was made to delete an earlier production or to insert new characters or phrases and to see where these revisions occur. By numbering the revisions, the S-notation also shows the order in which the revisions occur in the writing process. The S-notation uses the following symbols:

Symbol	Representation	Explanation
Pipe or 'vertical tab'	$\|_i$	The pipe symbol indicates a break in the text, i.e., a point in the text process where the writer deviated from linear text production. It appears at the position of the last action preceding the revision. Every break is labeled with a subscript sequential number (i), to indicate the order in which the breaks occured.
Curled brackets	{inserted text}^{i}	Curled brackets indicate an insertion in the text that occured in the text produced so far. Every insertion is labeled with a superscript sequential number (i) that refers to the break from which the insertion originated.
Square brackets	[deleted text]^i	Square brackets surround a deletion, either at the character, word or at a higher level. Every deletion is labeled with a superscript sequential number (i) that refers to the break from which the deletion originated.

The following is an example of an S-notation:

Questions·of·science,·s[x]¹|₁cience·and·[progress]²|₃{evolution}³.·End.|₂·

And the text produced in this example reads as follows:

Questions of science, science and evolution. End.

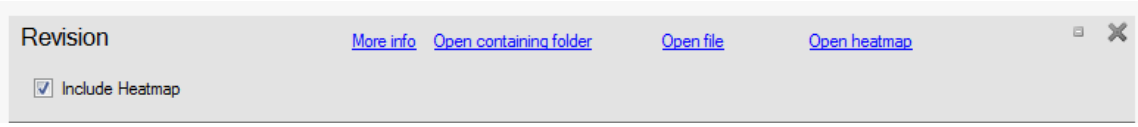
Revisions and breaks are numbered according to the order of their occurrence in the writing process.

- *Revision 1: correction of typing error, viz deletion of the single character 'x'*
  When typing the word 'science' the writer made a typo in the second letter: he hits the 'x' in stead of the 'c' (adjacent keys). The typo was immediately corrected (square brackets with index 1 left and right of the pipe line).

- *Revision 2: deletion of the word 'progress'*
  At the end of the text produced (see, |2 at the end of the line) the writer decided to delete the word 'progress' (see, [progress]2).
- *revision 3: insertion of the word 'evolution'*
  The third revision immediately followed the deletion of the word 'progress'. the writer inserted the word 'evolution', (see {evolution}3).

Revision Matrix

A second revision analysis produces a so called 'Revision matrix'. This revision matrix is a representation in which the revisions are listed sequentially, together with some basic time and position based characteristics.



variable	description
Include Heatmap	A heatmap is a representation of the final document in which the text areas at which a revision occured (either an insertion or a deletion) is color coded.

Matrix

After you have generated the revision analysis, a revision matrix is generated. The matrix looks as follows:

#Revision	Type	Content	Edits	Start	End	Duration	BeginPos	EndPos	Length	Chars	Chars without space	Words
0	Normal Production	Questions·of·science,·sx	32	11511	11525	00:00:13.806	0	24	24	20	17	4
1	Delete	x	1	11526	11526	00:00:00.062	24	23	1	1	1	0
0	Normal Production	cience·and·progress.·End	26	11527	11538	00:00:11.060	23	47	24	20	17	4
2	Delete	progress	6	11527	11538	00:00:11.060	34	42	8	8	8	1
3	Insert	evolution	21	11543	11546	00:00:02.746	34	43	9	9	9	1

variable	description
Revision	Every revision is numbered sequentially. These numbers correspond with the numbering system used in the indices of the S-Notation. The production of new text at the end of the text produced so far is labeled as 'normal production', and is coded as a 0-revision.
Type	Three types of text production are identified: ▪ normal production: new text produced at the end of the text produced so far; ▪ deletions: text that is deleted in the text produced so far, either immediately at the point of utterance, or delayed; ▪ insertions: text that is inserted in the text produced so far.
Content	Text that is produced, deleted or inserted at that specific stage of the process. *Remark: when characters are deleted with the Backspace key, they occur in the order in which they are deleted. This means that if you want to 'read' what is deleted, you should read the text from right to left.*

Edits	The number of actions that the writer needed to complete the action. This also includes shift key, cursor movements, mouse clicks etc.
Start	Start time of the action (s).
End	End time of the action (s).
Duration	Length of the action (hh:mm:ss:ms).
BeginPos	Position of the first character of the action.
EndPos	Position of the last character of the action. *Remark: the begin and end position enable you also to identify substitutions, e.g., the consecutive deletion and insertion (2 - 3) could be considered as a substitution because the position data reveal that these operation occured at the same place in the text.*
Length	
Chars	Number of chararacters produced.
Chars without spaces	Number of chararacters produced, excluding spaces.
Words	Number of words produced. *Remark: this count of words also includes incompleted words (e.g., 'sx' is recognized as a word in the first action).*

Heatmap

The generated heatmap looks as follows:

The heatmap is a proxy representation of the final MS-Word document in which the locations in the text that are characterized by revisions (either deletions or insertions, are highlighted. The intensity of the colour corresponds to the number of revisions that are located in the marked area: the colors range from yellow (relative few revisions) over green to dark red (lots of revisions).

Read more on revision analysis

Van Horenbeeck, E., Pauwaert, T., Van Waes, L., & Leijten, M. (2012). S-notation: S-notation markup rules (Technical Description) (D. o. Management, Trans.) (pp. 7). Antwerp: University of Antwerp.

Source

The source analysis is divided into two main categories. This analysis is based on the Focus events in the General Logging File.

element	description
window statistics	Statistics belonging to each Window in your logging session
window transition statistics	Statistics belonging to each transition or switch between the Windows used in your logging session.

This is an example of an original source analysis (based on the original idfx-file).

Window Statistics

Window Title	Total Time (s)	Total Time (relative)	Total Keystrokes	Total Keystrokes (relative)
Wordlog.docx - Microsoft Word	178.807	0.84	350	0.938
TASKBAR	5.648	0.027	0	0
Naamloos - Google Chrome	2.527	0.012	0	0
Google - Google Chrome	4.197	0.02	9	0.024
inputlog - Google Search - Google Chrome	1.419	0.007	0	0
Inputlog - Google Chrome	10.047	0.047	6	0.016
Program Manager	0.593	0.003	0	0
Features of Inputlog.docx - Microsoft Word	7.379	0.035	8	0.021
Inputlog 5.1.0.6	2.231	0.01	0	0
Total	212.85		373	

Window Transition Statistics

From Window Title	To Window Title	Count
Wordlog.docx - Microsoft Word	TASKBAR	4
TASKBAR	Naamloos - Google Chrome	1
TASKBAR	Wordlog.docx - Microsoft Word	2
TASKBAR	Program Manager	1
TASKBAR	Inputlog 5.1.0.6	1
Naamloos - Google Chrome	Google - Google Chrome	1
Google - Google Chrome	inputlog - Google Search - Google Chrome	1
inputlog - Google Search - Google Chrome	Inputlog - Google Chrome	1
Inputlog - Google Chrome	Wordlog.docx - Microsoft Word	1
Program Manager	Features of Inputlog.docx - Microsoft Word	1
Features of Inputlog.docx - Microsoft Word	TASKBAR	1
Total		15

variable	description
window title	title of the window: similar to Focus event of the General Logging File

total time (s)	total time spent in a window/source in seconds.milliseconds
total time (relative)	relative time spent in a window/source represented as a ratio (0-1)
total keystrokes	total number of keystrokes types in each window/source
total keystrokes (relative)	relative amount of characters typed in each window/source represented as a ratio (0-1)
from window title	representation of the window titles that transitions have started from
to window title	representation of the window titles that transitions went to
count	number of swiches between the represented window titles

This is an example of a recoded source analysis (based on a recoded idfx-file).

Window Statistics

Window Title	Total Time (s)	Total Time (relative)	Total Keystrokes	Total Keystrokes (relative)
Main document	186.14	0.875	350	0.938
Sources	26.708	0.125	23	0.062
Total	212.85		373	

Window Transition Statistics

From Window Title	To Window Title	Count
Main document	Sources	2
Sources	Main document	2
Total		4

variable	**description**
window title	title of the source: as chosen via the pre-process module recode The window title now can be called 'sources' and this might comprise 5 seperate window titles.
total time (s)	total sum of time spent in a recoded source in seconds.milliseconds
total time (relative)	relative time spent in recoded source represented as a ratio (0-1)
total keystrokes	total number of keystrokes typed in each recoded source
total keystrokes (relative)	relative amount of characters typed in each recoded source represented as a ratio (0-1)
from window title	representation of the recoded source that transitions have started from
to window title	representation of the recoded source that transitions went to
count	number of switches between the recoded sources

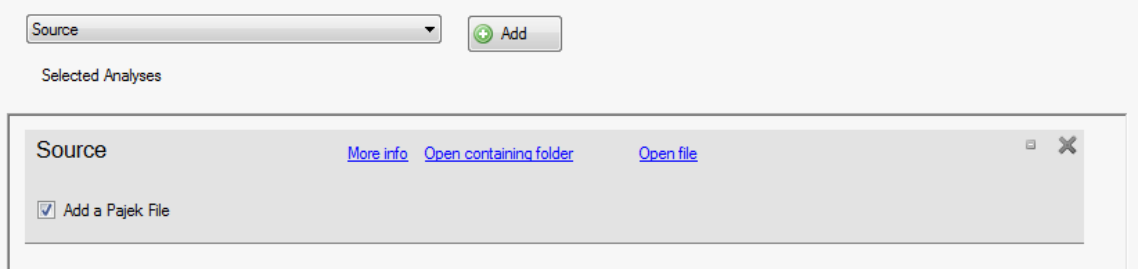
Pajek File

In the Source analysis you have the option to generate a Pajek file, which is a format you can use to generate a network diagram representing the relative proportion of time spent in a source, and the interaction between the sources accessed.

To generate a network diagram:

Step 1: Create a .net file

1. select the 'Source' analysis
2. add it to the list of requested analyses.
3. select the 'Add a Pajek File' option



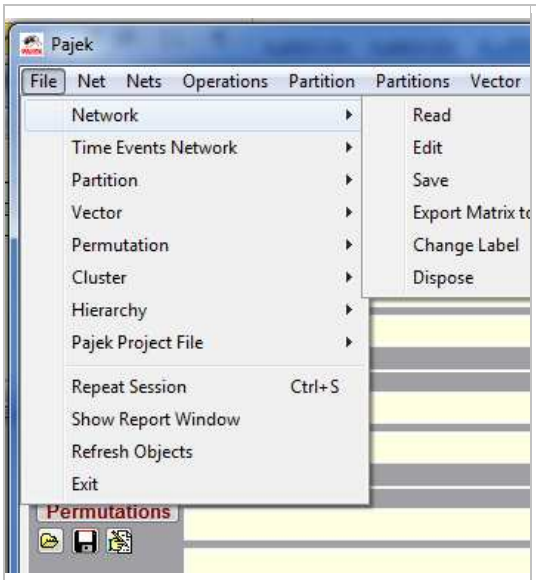
4. start the analysis by clicking the 'analyze' button
5. go to output folder by clicking 'Open containing folder'
   A net file - which has a **.net** extension - is available there. If you want to read it, right click the .net file and select 'Open with' > Notepad. The file is structured as follows:
   a. ***Vertices part***: The vertices represent the different sources accessed during the logging session. In the example below, there were three sources identified, resp. the Main Word Document, Twitter and an online Dictionary. For every source certain characteristics are added:
   (a) ellipse: the x/y_fact-score refer to the relative proportion of the source, i.e. size of the ellipse;
   (b) color: the ic/bc refer to the column of the ellipse
   (c) line: numbers represent the line weights
   b. ***Arcs:*** The arcs represent the arrows between the vertices representing the interaction between sources. The code in the .net file reads a s follows:
   (a) source origin id: the first number refers to the number of the vertix where an action starts from, e.g., Word document
   (b) source destination id: the second number refers to the number of the vertix the writer has switched to, e.g., Twitter
   c. (c) number of interactions: the third number represents the number of times a writer switched from the origin source to the destination source, e.g., two times from Word to Twitter in this case.

```
*Vertices 3

1 "Main Word Document" 0.0 0.0 0.5 ellipse x_fact 45 y_fact 45
ic Fuchsia bc Fuchsia lr 6 fos 10

2 "Twitter" 0.0 0.0 0.5 ellipse x_fact 48 y_fact 48 ic
OliveGreen bc OliveGreen lr 6 fos 10
```

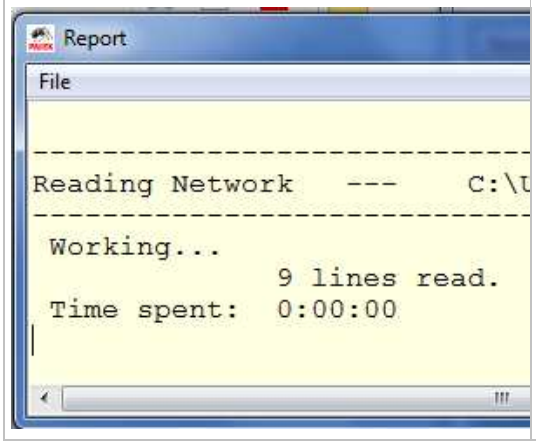
```
3 "Dictionary" 0.0 0.0 0.5 ellipse x_fact 7 y_fact 7 ic
CornflowerBlue bc CornflowerBlue lr 6 fos 10


*Arcs

1 2 2 w 1

1 3 1 w 1

2 1 2 w 1

3 1 1 w 1
```

Step 2: Open the .net file in Pajek

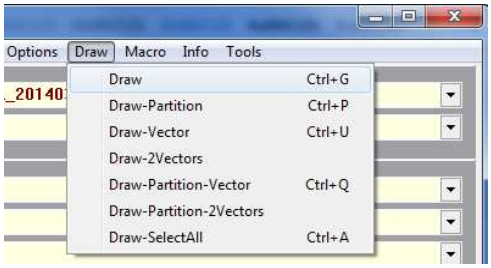
Pajek is an open source program that can be downloaded from http://pajek.imfm.si



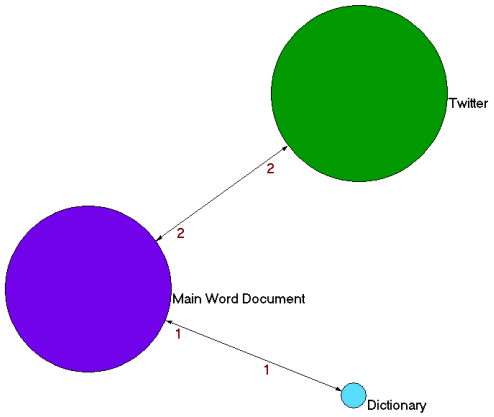
1. Open Pajek
2. Select File < Network < Read



3. Select and open the .net file created by Inputlog (output folder)



4. Select Draw > Draw in the top menu



5. In the Draw window you can now alter the network representation (position, labeling, colors etc.)

Remark: Use Node XL instead of Pajek
You can also opt to use the NodeXL template for Excel. This template allows you to import Pajek .net-files in Excel and draw network diagrams within this program.
"NodeXL is a free, open-source template for Microsoft® Excel® 2007, 2010 and 2013 that makes it easy to explore network graphs. With NodeXL, you can enter a network edge list in a worksheet, click a button and see your graph, all in the familiar environment of the Excel window."
http://nodexl.codeplex.com/
http://research.microsoft.com/en-us/projects/nodexl/

Copy task

The copy task analysis provides a carefully designed set of perspectives to explore and characterize a participant's motoric keyboard interaction. The default copy task - already available for a set of languages - is designed in such a way that it is possible to adequately address these characteristics, and use them in further analyses (e.g., as a co-variate in variance analyses, or in regression and multilevel analyses).

Characteristics	The copy task analysis addresses the following bigram characteristics: ▪ Frequency ▪ Adjacency ▪ Repetition ▪ Hand combination
Bigram Frequency	In the copy task word and sentence combinations are designed in such a way that it is possible to isolate - language specific - high and low frequency bigrams in the analysis. The frequency is based on large corpora (e.g., Subtlex or CELEX). ▪ High frequency: 30 % highest bigram percentiles ▪ Low frequency: 50 % lowest bigram percentiles
Bigram adjacency	When two keys are next to each other on the keyboard (left, right, up, down) they are labelled as 'adjacent'. this labelling is keyboard specific. Therefore, the identification of the keyboard lay-out at the start of each session is crucial.
Repetitive bigrams	When a bigram is produced by pressing the same key twice, we consider it as a repetitive bigram.
Bigram hand combination	To produce a bigram, the following hand combinations could be used, taking into account the order: ▪ Left - Right hand combination ▪ Right - Left hand combination ▪ Left - Left hand combination ▪ Right - Right hand combination Remark: The midzone of the keyboard is not included in the analysis, as - dependent on the typing expertise - the hand that is used to type these keys could vary (more often). 

Reported measures

For every sub analysis within the copy task analysis, the following descriptive measures are reported:

- Count (targeted): number of characters typed that correspond to a character that was prompted in one of the respective copy task components
- Count (not targeted): number of characters typed that do not correspond to a character that was prompted in one of the respective copy task components
- Mean interkey interval (IKI): arithmetic mean (in milliseconds) of the targeted bigrams within the scope of the respective analyses
- Standard deviation: standard deviation of the targeted bigrams within the scope of the respective analyses
- Median: median (in milliseconds) of the targeted bigrams within the scope of the respective analyses
- Log mean trimmed: 95% log converted trimmed geometric mean of the targeted bigrams within the scope of the respective analyses
- Coefficient of variation: calculated coefficient of variation (based on mean and stdev) 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, viz. 60 000 (ms) / mean IKI (ms)
- Absolute CPM: time based calculation of (theoretical number of) characters per minute based on an extrapolation of the time on task within a copy task component (e.g. for the default sentence copy task, participants are required to type for 45 s. In that case the time between the first and the last key, i.e. approximation of 45 s is used to calculate a time based extrapolation of the CPM).

Session information

The copy task analysis starts with some meta information about the participant and the session. This report is based on the questions that are presented at the beginning and end of the copy session: particpant information, hard and software used, handedness score, language and education.

Copytask Analysis

Meta Information	
Logfile	[redacted]_Dutch_30-03-2018.idfx
Log Creation	30/03/18 10:03:23.5
Log GUID	596568e6-5875-4c9c-9683-cac17ac6e38d
Logging Program Version Number	6.1.2.0
Analysis Creation	01/04/18 11:36:38
Analysis GUID	9bbc7119-230d-4565-915c-15135501861b
Analysis Program Version Number	7.1.0.53
Session Identification	
Participant	[redacted]
Text Language	NL
Age	23
Gender	V (vrouw)
Keyboard	AZERTY
Handedness Score	87.5
Computer	desktop
Keyboard Familiarity	vertrouwd
Browser	Google Chrome
Dominant Languages	Engels, Nederlands
Disorder	False
Education	bachelor
Repetition	False

Report overview

The copy task analysis provides a detailed analysis split up in the following sections:

- Correctness score
- Synthesis of InterKey intervals (IKI)
- Individual Component analysis
- Aggregated characteristics

Correctness

When execuring the copy tasks, participants strive to find an optimal trade-off between speed and correctness (cf. also basic instruction). The copy task analysis opens with a summary of correctness scores, comparing the targetted bigrams (in the prompted texts) with the characters actually typed.
The analysis first reports the correctness scores for each component seperately and then an aggregated synthesis is reported, both overall and for the group of selected components (see asterisks in component analysis).

Example:

Correctness Scores

Correctness per component

Component	Count (targetted)	Count (not targetted)	Correct
Typing Speed	151	0	100.0%
Zin	91	7	92.9%
Woorden 1*	130	3	97.7%
Woorden 2*	128	15	89.5%
Woorden 3*	130	4	97.0%
Woorden 4	110	4	96.5%
Medeklinkers	19	2	90.5%

Statistics

	Overall	Selected Components	
Aggr Targetted	759	388	
Aggr Not Targetted	35	22	
Aggr Correctness	95.6%	94.6%	
Mean	94.9%	94.8%	
Stdev	3.7	3.7	
Median	96.5%	97.0%	
Min	89.5%	89.5%	
Max	100.0%	97.7%	

Synthesis

The 'Overall Synthesis of the InterKey Intervals (IKI)' in the output summarizes the overall performance of the copy task.
Data analysis is represented at three levels:

- Targeted bigrams: selects all bigrams that are targeted by one of the prompts in the corresponding copy task
- High frequency bigrams: selects all high frequency (HF) bigrams targeted by one of the prompts in the corresponding copy task
- Selected components bigrams: limits the analysis of targeted bigrams to those occurring in the selected components of the corresponding copy task (see documentation on 'Copy task creator'). For the default copy tasks these selected components refer to the three HF bigram word combination tasks.

Example:

Overall Synthesis of the InterKey Intervals (IKI)

	Count (targetted)	Count (not targetted)	Mean IKI	StdDev	Median	LogMean (trimmed)	Coef. of Variance	CPM
Targetted Bigrams	784	29	141.1	105.1	120	119.7	58.1 %	425
High Frequency Bigrams	609	15	130.2	91.4	112	112.6	54.5 %	461
Selected Components' Bigrams	393	15	128.7	93.9	110	111.5	53.8 %	466

Component analysis

In the component analysis the following elements are reported:

- Overall analysis: analysis of all the produced characters produced in each copy task component
- Trial cut > 2: analysis of all the produced characters produced in each repetitive component task, excluding those in the first two trials
- Time filtered 10%: analysis of all the produced characters produced in each component task, excluding the first 10% time period of each task execution
- Trial report: analysis of all the produced characters produced in each trial within a component task.

Example:

Components

Overall

	Count (targetted)	Count (not targetted)	Mean IKI	StdDev	Median	LogMean (trimmed)	Coef. of Variance	CPM	Absolute CPM
Tapping task	121	1	121.3	40.1	128	113.9	38.2 %	495	488
Sentence	142	3	123.0	63.8	120	110.4	48.7 %	488	284
Words 1	130	2	122.6	86.0	105	105.9	56.0 %	489	307
Words 2	132	12	128.3	96.8	116	113.7	48.0 %	468	248
Words 3	131	1	135.2	97.9	104	115.0	57.1 %	444	282
Words 4	109	8	186.2	126.6	152	150.8	72.9 %	322	165
Consonants	19	2	402.0	231.0	375	348.8	58.4 %	149	94

Characteristics

In the characteristics analysis the following elements are reported:

- Frequency: overall analysis of all intra bigram intervals for resp. high and low frequent bigrams (HF and LF)
  Definition of high and low frequent bigrams: see above/link.
- Frequency (Trial Cut >2): overall analysis of all intra bigram intervals for resp. high and low frequent bigrams (HF and LF), excluding those produced in the first two trials of each repetitive component
- Frequency (Only repetitions & Trial Cut >2): overall analysis of all intra bigram intervals for resp. high and low frequent bigrams (HF and LF) in repetitive copy task components, excluding those produced in the first two trials of each of those components

Example:

Characteristics

Frequency

	Count (targetted)	Count (not targetted)	Mean IKI	StdDev	Median	LogMean (trimmed)	Coef. of Variance	CPM
HF	609	15	130.2	91.4	112	112.6	54.5 %	461
LF	33	5	370.0	185.8	337	337.2	43.4 %	162

Frequency (Trial Cut >2)

	Count (targetted)	Count (not targetted)	Mean IKI	StdDev	Median	LogMean (trimmed)	Coef. of Variance	CPM
HF	419	8	128.4	74.9	112	112.9	53.1 %	467
LF	27	3	376.2	200.3	320	338.8	46.4 %	159

- Hand combination : overall analysis of all intra bigram intervals for the different hand combinations needed to produce a bigram, resp. left-right hand combination (LR), right-left hand combination (RL), left-left hand combination (LL) and right-right hand combination (RR).
  Definition of high and low frequent bigrams: see above/link.
- Hand combination Frequency (Trial Cut >2): overall analysis of all intra bigram intervals for the different hand combinations needed to produce a bigram, excluding the first two trials of the repetitive components.
- Adjacency: overall analysis of all intra bigram intervals for non-adjacent (False) and adjacent (true) keys.
- Adjacency (trial cut >2): overall analysis of all intra bigram intervals for non-adjacent (False) and adjacent (True) keys, excluding the first two trials of repetitive components.
- Repetition: overall analysis of all intra bigram intervals for non-repetitive (False) and repetitive (True) keys.
- Adjacency (trial cut >2): overall analysis of all intra bigram intervals for non-repetitive (False) and repetitive (True) keys, excluding the first two trials of repetitive components.

Linguistic

When you select the Linguistic analysis, the program identifies the logging file and prepares a so-called W-Notation, which is a time and revision enriched word aggregated process file at the word level. This file is sent to the web service and the user is prompted to start the log-in procedure (personal account). At any time the user can monitor the progress of the analysis via the available web service. An automatic email alert service is provided for batch processing.

The Linguistic analysis has been funded via FWO and has been a collaboration between the University of Antwerp and University College Ghent. You can find additional information on the project website

Final Product(from document)

User document not available.

S-Notation

[q]^{1}|$_{1}$Th[re]^{2}|$_{2}$ere·is·a·man·sleap[ee]^{3}|$_{3}$ing·in·an·easy·chair.·He·may·have·drunk·a·little·too·much·br[e]^{4}|$_{4}$andy.·His·glass·is·emply.·The·little·girl,·perhaps·his·daughter.·is·trying·to·tellhim·about·an·impending·disaster.·The·cat[·]^{5}|$_{5}$·is·fishing·in·the·fishing·bowl,·looking·for·a·quick·meal·but·has·nudged·the·books·on·the·shelf·over·the·mans·hea[s]^{6}|$_{6}$d·and·books·are·now·on·thei·[e]^{7}|$_{7}$way·to·hit·him·and·wake·him·up.·Th[r]^{8}|$_{8}$er·are·less·dramatic·[h]^{9}|$_{9}$th[o]^{10}|$_{10}$ings·to·describe·as·weel.·There·are·[a]^{11}|$_{11}$dols·and·a·teddy·bear·on·the·f[o]^{12}|$_{12}$loor·and·a·stereo·in·the·cabinet·but·these·are·mot·part·of·the·real·dram[ma]^{13}|$_{13}$a.·

W-Notation

_1_[q]_1_^_1_Th_2_[re]_2_^_2_ere·is·a·man·sleap_3_[ee]_3_^_3_ing·in·an·easy·chair.·He·may·have·drunk·a·little·too·much·br_4_[e]_4_^_4_andy.·His·glass·is·emply.·The·little·girl,·perhaps·his·daughter.·is·trying·to·tellhim·about·an·impending·disaster.·The·cat_5_[·]_5_^_5_·is·fishing·in·the·fishing·bowl,·looking·for·a·quick·meal·but·has·nudged·the·books·on·the·shelf·over·the·mans·hea_6_[s]_6_^_6_d·and·books·are·now·on·thei·_7_[e]_7_^_7_way·to·hit·him·and·wake·him·up.·Th_8_[r]_8_^_8_er·are·less·dramatic·_9_[h]_9_^_9_th_10_[o]_10_^_10_ings·to·describe·as·weel.·There·are·_11_[a]_11_^_11_dols·and·a·teddy·bear·on·the·f_12_[o]_12_^_12_loor·and·a·stereo·in·the·cabinet·but·these·are·mot·part·of·the·real·dram_13_[ma]_13_^_13_a.·

Reconstructed Text

There is a man sleaping in an easy chair . He may have drunk a little too much brandy . His glass is emply . The little girl , perhaps his daughter . is trying to tellhim about an impending disaster . The cat is fishing in the fishing bowl , looking for a quick meal but has nudged the books on the shelf over the mans head and books are now on thei way to hit him and wake him . up . Ther are less dramatic things to describe as weel . There are dols and a teddy bear on the floor and a stereo in the cabinet but these are mot part of the real drama .

Inserts in context

There·is·a·man·sleaping·in·an·easy·chair.·He·may·have·drunk·a·little·too·much·brandy.·His·glass·is·emply.·The·little·girl,·perhaps·his·daughter.·is·trying·to·tellhim·about·an·impending·disaster.·The·cat·is·fishing·in·the·fishing·bowl,·looking·for·a·quick·meal·but·has·nudged·the·books·on·the·shelf·over·the·mans·head·and·books·are·now·on·thei·way·to·hit·him·and·wake·him·up.·Ther·are·less·dramatic·things·to·describe·as·weel.·There·are·dols·and·a·teddy·bear·on·the·floor·and·a·stereo·in·the·cabinet·but·these·are·mot·part·of·the·real·drama.·

Deletions in context

_1_[q]_1_Th_2_[re]_2_ere·is·a·man·sleap_3_[ee]_3_ing·in·an·easy·chair.·He·may·have·drunk·a·little·too·much·br_4_[e]_4_andy.·His·glass·is·emply.·The·little·girl,·perhaps·his·daughter.·is·trying·to·tellhim·about·an·impending·disaster.·The·cat_5_[·]_5_·is·fishing·in·the·fishing·bowl,·looking·for·a·quick·meal·but·has·nudged·the·books·on·the·shelf·over·the·mans·hea_6_[s]_6_d·and·books·are·now·on·thei·_7_[e]_7_way·to·hit·him·and·wake·him·up.·Th_8_[r]_8_er·are·less·dramatic·_9_[h]_9_th_10_[o]_10_ings·to·describe·as·weel.·There·are·_11_[a]_11_dols·and·a·teddy·bear·on·the·f_12_[o]_12_loor·and·a·stereo·in·the·cabinet·but·these·are·mot·part·of·the·real·dram_13_[ma]_13_a.·

element	description
final product (from document)	The final text of the Word document produced. Note: If you run the linguistic analyses with the idfx on its orignal location the final Word document is inserted here. But if you run the linguistic analyses in batch this field will remain empty.
S-notation	The S-notation is a complete record of the text including revisions as they occured in the writing process.
W-notation	The W-Notation is similar to the S-Notation. However, it uses different symbols that are easier to recognize by a computer program. Example: - Final production: "A new text with a deletion." - Revisions made during the writing process: A _1_{new}_1_ text ^_2_ with a _2_[small]_2_ deletion. ^_1_ Symbols used: - The break, expressed by a caret (^). A break is the moment in the text where the author interrupts the normal production to insert in or to delete from the text produced so far. In the example there are two breaks: first at the end of the sentence and secondly after the word 'text'. - The insertion, marked with curled braces ({}) around text to indicate an addition into the normal production. In the example 'new' was inserted between 'A' and 'text'. - The deletion, marked with square brackets ([]) around the text indicate what was removed from the writing. The word 'small' has been deleted. In order to identify the modifications each symbol gets a digit surrounded by underscores. We see the second break marked as: ^_2_ , the insertion as: _1_{new}_1_ and the deletion as: _2_[small]_2_
Reconstructed text	This is the recontructed final product of the text including spaces before each punctuation symbol.
Inserts in context	The inserts in contexts shows the reconstructed text in which insertions are added as they occured in the writing process
Deletions in context	The deletion in contexts shows the reconstructed text in which deletions are shown as they occured in the writing process

The logged process data are enriched with different kinds of linguistic information: part-of-speech tags, lemmata, chunk boundaries, named entity information, syllable boundaries and word frequency. The set-up of the extension to Inputlog is largely language-independent. As proof-of-concept, the extension has been developed for English and Dutch.

Revisions	S-Notation	#Chars	Token	PoSA	PoSB	PoS-Prob	Lemma	Lemma-Prob	ChunkA	ChunkB	NE	NE-Prob	LogFreq	RelFreq	Syll
1-I 2-I	[q]Th[re]ere·	8	There	EX	-	1.00	there	NaN	O	-	O	0.99	69	0.117	there
	is·	2	is	VBZ	-	1.00	be	0.70	B	VP	O	1.00	77	0.786	is
	a·	1	a	DT	-	1.00	a	NaN	B	NP	O	1.00	77	1.516	a
	man·	3	man	NN	-	0.99	man	NaN	I	NP	O	1.00	69	0.030	man
3-I	sleap[ee]ing·	10	sleaping	VBG	-	0.88	sleap	0.86	B	VP	O	1.00	NaN	NaN	slea-ping
	in·	2	in	IN	-	0.97	in	NaN	B	PP	O	1.00	77	1.414	in
	an·	2	an	DT	-	1.00	an	NaN	B	NP	O	1.00	77	0.254	an
	easy·	4	easy	JJ	-	1.00	easy	NaN	I	NP	O	1.00	69	0.021	ea-sy
	chair·	5	chair	NN	-	1.00	chair	NaN	I	NP	O	1.00	61	0.007	chair
	.	1	.	.	-	1.00	.	NaN	O	-	O	1.00	NaN	NaN	=

Left part of the variables.

variable	**description**
Revisions	Number of revisions
S-notation	Each separate word and insertion based on the S-notation. Word level revisions are included in the construction of a word. (e.g., as such within word typing errors are still parsed correclty: only word level revisions and larger revisions are taken into account as separate words. Remark: the S-notation was parsed and three types of data were extracted from the S-notation: word level revisions — the word-level revisions can be extracted from the S-notation by retaining all words with word-internal square or curly brackets deleted fragments — the deleted fragments can be extracted from the S-notation by retaining only the words and phrases that are surrounded by word-external square brackets final writing product — the final product data can be obtained by deleting everything in between square brackets from the S-notation
#Chars	Number of characters of the insertion
Token	An individual occurrence of a linguistic unit in writing
PoSA	part-of-speech tagging is the process of marking up a word in a text (corpus) as corresponding to a particular part of speech, based on both its definition, as well as its context. In the process data the relationship with adjacent and related words in a phrase, sentence, or paragraph are taken into account. Simplified Part-of-Speech tagset. Full list available on-line. **Alphabetical list of part-of-speech tags used in the CGN-dataset:**

Tag	Meaning	Examples
ADJ	adjective	new, good, high, special, big, local
ADV	adverb	really, already, still, early, now
CNJ	conjunction	and, or, but, if, while, although
DET	determiner	the, a, some, most, every, no
EX	existential	there, there's
FW	foreign word	dolce, ersatz, esprit, quo, maitre
MOD	modal verb	will, can, would, may, must, should
N	noun	year, home, costs, time, education
NP	proper noun	Alison, Africa, April, Washington
NUM	number	twenty-four, fourth, 1991, 14:24
PRO	pronoun	he, their, her, its, my, I, us
P	preposition	on, of, at, with, by, into, under
TO	the word to	to
UH	interjection	ah, bang, ha, whee, hmpf, oops
V	verb	is, has, get, do, make, see, run
VD	past tense	said, took, told, made, asked
VG	present participle	making, going, playing, working
VN	past participle	given, taken, begun, sung
WH	wh determiner	who, which, when, what, where, how

Alphabetical list of part-of-speech tags used in the Penn Treebank Project:

Number	Tag	Description
1.	CC	Coordinating conjunction
2.	CD	Cardinal number
3.	DT	Determiner
4.	EX	Existential there
5.	FW	Foreign word
6.	IN	Preposition or subordinating conjunction
7.	JJ	Adjective
8.	JJR	Adjective, comparative
9.	JJS	Adjective, superlative
10.	LS	List item marker
11.	MD	Modal
12.	NN	Noun, singular or mass
13.	NNS	Noun, plural
14.	NNP	Proper noun, singular
15.	NNPS	Proper noun, plural
16.	PDT	Predeterminer
17.	POS	Possessive ending
18.	PRP	Personal pronoun
19.	PRP$	Possessive pronoun
20.	RB	Adverb
21.	RBR	Adverb, comparative
22.	RBS	Adverb, superlative
23.	RP	Particle
24.	SYM	Symbol

	25. TO to 26. UH Interjection 27. VB Verb, base form 28. VBD Verb, past tense 29. VBG Verb, gerund or present participle 30. VBN Verb, past participle 31. VBP Verb, non-3rd person singular present 32. VBZ Verb, 3rd person singular present 33. WDT Wh-determiner 34. WP Wh-pronoun 35. WP$ Possessive wh-pronoun 36. WRB Wh-adverb
PosB	Further elaboration of the features of the tags. Full list available on-line.
PosProb	The probability that the Part of Speech tagging is correct (0-100 %) Note: the probabilities are based on the used databases.
Lemma	For each orthographic token, the base form (lemma) is given. For verbs, the base form is the infinitive; for most other words, this base form is the stem
Lemma-prob	The probability that the Lemma tagging is correct (0-100 %) Note: the probabilities are based on the used databases.
ChunkA	The LT3 chunkers are rule-based and contain a small set of constituency and distituency rules. consituency rules — part-of-speech tag sequences that can occur within a constituent (such as preposition + noun) distituency rules — 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. In the IOB-tagging scheme, each token belongs to one of the following three types: I Inside O Outside B Begin The B- en I-tags are followed by the chunk type, e.g., B-VP, I-VP.
ChunkB	Second part of the chunking Examples of chunks: NP chunk: an individual noun phrase should be formed whenever the chunker finds an optional determiner (DT) followed by any number of adjectives (JJ) and then a noun (NN). VP chunk: verb phrase.

	The O chunk tag is used for tokens which are not part of any chunk Other types are PP, ADJP and ADVP chunks.
NE	The Named Entity identifies sequences of words in a text that belong to predefined NER categories such as names of persons, products or locations (e.g., Tom Smith, Inputlog, Antwerp)
NEProb	The probability that the Named Entity is correct (0-100 %)
LogFreq	The Logscore of the absolute frequency of a word in relation to the database of Web1T Google corpus
RelFreq	The relative frequencies of a word in relation to the database of Web1T Google corpus
Syll	Syllabification was approached as a classification task: a large instance base of syllabified data is presented to a classification algorithm, which automatically learns from it the patterns needed to syllabify unseen data.

Syll	Start WordID	End WordID	Start WordTime	End WordTime	BfrWord-2	BfrWord-1	Word Prod	Within WordPause	AftWord+1
there	0	6	9033	22605	0	0	13369	13322	250
is	6	9	22995	23463	250	390	374	312	156
a	9	11	23697	23837	156	234	62	234	140
man	11	15	24430	25163	140	593	546	499	234
slea-ping	15	24	25803	33088	234	640	7207	7145	140
in	24	27	33400	33712	140	312	234	156	156
an	27	30	33915	34601	156	203	515	452	234
ea-sy	30	35	34898	36754	234	297	1201	1139	717
chair	35	42	37799	42401	717	1045	4322	4259	343
=	-	-	-	-	-	-	-	-	-

Right part of the variables

StartWordID	The ID related to the beginning of the word Remark: this ID is generated specifically for this purpose, there is no relation with the ID's of the General logging file
EndWordID	The ID related to the end of the word Remark: this ID is generated specifically for this purpose, there is no relation with the ID's of the General logging file
StartWordTime	The time in milliseconds related to the beginning of the word: start time of the beginning of the word
EndWordTime	The time in milliseconds related to the end of the word: start time of the first event after the word
BfrWord-2	The Before Word Pause -2: connected to the spacebar

BfrWord-1	The Before Word Pause -1: directly connected to the word
WordProd	The word production time measured from the start time of the word until the end time of the final character of the word Note: if there occur revisions in a word, the word production time can be very long. If this is noise to your data, than you can opt to filter the words that contain revisions.
WithinWordPause	The sum of the pauses within the word
AftWord+1	The after word pause +1: directly connected to the word

The databases used for the linguistic analysis are:

Variable	**Dutch**	**English**
Part of speech tagging	Corpus Gesproken Nederlands (316 tags)	Penn Treebank (45 tags)
Lemmatizer	Celex Nederlands	Celex English
Named Entity	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.	CONLL-2003 shared task data and discerns four NER categories: person, organization, location and miscellaneous
Chunking	LT3	LT3
Word Frequency	Word frequency information for English and Dutch is retrieved from frequency lists derived from the Web1T Google corpus, which is available from LDC. 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).	
Syllibification	The syllabification tools were trained on Celex using Timbl as classification algorithm.	

Remark: The part-of-speech tagger, the lemmatizer and the named entity recognizer are trained with CRF++ , an open source implementation of Conditional Random Fields (Lafferty, McCallum, & Pereira, 2001), which is a machine learning technique suited for labelling sequential data.

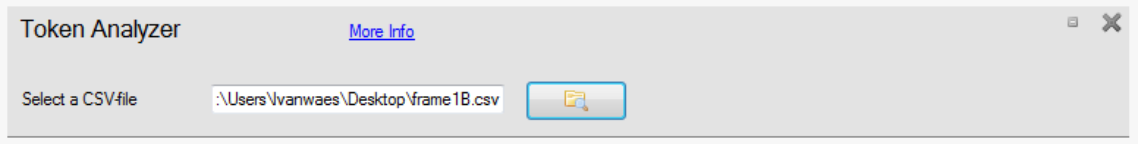
Read more on linguistic analysis

Leijten, M., Macken, L., Hoste, V., Van Horenbeeck, E., & Van Waes, L. (2012). From Character to Word Level: Enabling the Linguistic Analyses of Inputlog Process Data. Paper presented at the European Association for Computational Linguistics, EACL - Computational Linguistics and Writing (CL&W 2012): Linguistic and Cognitive Aspects of Document Creation and Document Engineering, Avignon.

Macken, L., Hoste, V., Leijten, M., & Van Waes, L. (2012). From keystrokes to annotated process data: Enriching the output of Inputlog with linguistic information. Paper presented at the Eight International Conference on Language Resources and Evaluation (LREC'12), Istanbul, Turkey.

Token analyzer

The Token Analysis allows the researcher to analyse the orthographical process of specific words. The program will search for words in the logging file that match, in full or in part, with a list of given target words (words with the correct orthography).



Select a CSV-file	Select the comma separated text (CSV) file that contains a list of target words (and the surrounding words) for which you want the pause information (see below for a detailed instruction about the syntax of this CSV-file).

Example

Consider a spelling test (e.g., a dictation) to assess the correct writing of the homophones (words sharing the same pronunciation) *laid*, *led* and *let*. The following three sentences will be used:

1. After his death, he was *laid* in a splendid tomb.
2. These circumstances *led* to believe in his innocence.
3. You can't hold it, *let* it go.

In this case there are three words that the program should find in the logging file. A separate file (.csv) is provided that holds those target words. In order to facilitate the lookup the targets are surrounded with the word that precedes it and with two words that follow it. As such:

Example of a CSV-file	was; *laid*; in; a circumstances; *led*; to; believe it; *let*; it; go

Each target word stands on its own line together with the words preceding and following it, separated with a semicolon (;). Do not include punctuation marks. A simple text editor can be used to prepare this file, no formatting is needed. The file should be saved not as a *.txt but with the extension *.*csv* (comma separated values).

After taking the spelling test, the Token Analyzer analysis will process the results. At this point the proper csv file should be selected by pressing the file selection button.

The output

The main results window shows the final text as it was produced from the process string by deleting the deletions and inserting the insertions. The s-notation represents the writing process with all its revisions. The next part of the report shows a detailed analysis of every tested word with a line ID, the user identification, the target word, the word as it was finally produced, the process data of the word (s-notation), and the revisions (if any) relevant to the test word. For instance: line 1 has two revisions for the word 'laid': one revision was delayed (3-D) and one was immediate (4-I). An

immediate revision happens between words, a delayed revision has one or more words between the break and the actual deletion or insertion.

The remaining part of the line is filled with the digraphs (or bigrams) of the word, abbreviated as Digr_1, Digr_2, ..., with the appropriate pause time for that digraph. For analysis purposes an additional digraph is added with the letter that precedes the test word.

Reconstructed Text

After·his·death·he·was·laid·in·a·splendid·tomb·These·circumstances·led·to·believe·in·his·innoence··You·can't·hold·it,·let·if·go····

S-Notation

After·his·death·he·was·l[ee]^{3}|$_{4}${ai}^{4}d·in·a·sp[[e]^{1}|$_{1}$m]^{2}|$_{2}$lendid·tomb|$_{3}$·These·circumstances·led·to·believe·in·his·inno[n{}^{5}|$_{6}$c]^{6}ence|$_{5}$··You·can't·hold·it,·let·if·go····

ID	User	Target	Produced	S-Notation	Revisions	Digr_1	Pause_1	Digr_2	Pause_2	Digr_3
1	E	laid	laid·	s·l[ee]{ai}d·	3-D 4-I	s·	141	·l	2137	le
2	E	led	led·	s·led·		s·	609	·l	1965	le
3	E	let	let·	t,·let·		t,	280	,·	203	·l

Additional Information

The Token Analysis file has additional information proper to this task: the name of the csv file, the number of target words in that file and the number of words that (more or less) match the target. Target words with no match are listed too.

Token Analysis File

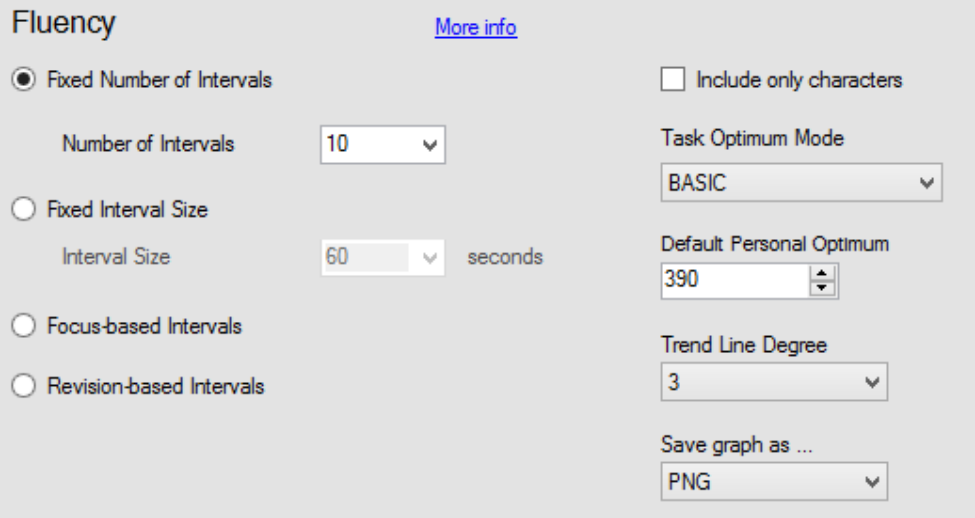
Meta Information	
Logfile	E_1.idfx
Log Creation	04/04/14 17:16:35.228
Log GUID	258bab2f-af83-465f-8626-ecf408b2834a
Logging Program Version Number	6.0.0.1
Analysis Creation	04/04/14 17:24:16
Analysis GUID	6f1f428f-ea03-4429-99b6-ff4f8d6b852e
Analysis Program Version Number	6.0.0.1
Session Identification	
Participant	E
Text Language	NL
Target file information	
CSV file used:	spell.csv
Total Targets in the file:	3
Matched Targets:	3
Missed Targets:	All targets found.
Damerau-Levenshtein Distance:	1

The Damerau-Levenshtein Distance is a metric defining the difference between two strings (target and production), given by counting the minimum number of operations needed to transform one string into the other, where an operation is defined as an insertion, deletion, or substitution of a single character, or a transposition of two adjacent characters. It is used to judge if a target and a produced word match.

Fluency

Personal Optimum Storing

By default, Inputlog stores personal optima for each author. If no optimum has been stored yet for the author, a default value is picked. Whenever a fluency analysis for the author yields a task optimum, personal optimum, a message box offers to store this task optimum as the new personal optimum. Personal optimum history is stored in file Workspace\Settings\PersonalOptima.txt. Note when (manually) editing the file, all author names should be converted to lowercase.



Interval settings	Fluency analysis will only process intervals of length ≥ 10s
Include only characters	By default, the fluency analysis counts every keypress. This checkbox allows the user to restrict the fluency analysis to "productive keystrokes" keypresses which generate or remove actual characters.
task optimum mode	The fluency analysis comes with two methods for computing the task optimum. 1. Basic: Task optimum is computed as a rolling average with a window of 3 times 10s = 30s. 2. Interval Dependent: Rolling average window size equals minimal interval length x (the length of a single interval, when all intervals have the same length). The window is split into 2-6 parts, depending on interval length via an inverse factorial relationship. For example, for an interval size of 3 minutes, the rolling average window will be 4 times 45s = 180s. This method ensures that there is always a real possibility of achieving the task optimum in a single interval, regardless of session length.
default personal optimum	Allows the user to change the default personal optimum (used when no optimum has been stored for an author) on a per-analysis basis.
trend line degree	Degree of the polynomial fit for the trendline (note: this should be < number of intervals to achieve a trendline effect).

save graph as	You can save the file in various formats PNG BMP EMF JPG TIFF

Show Graph (after analysis)

single vs multiple idfx	In case multiple idfx files (or the same idfx file with different settings), with compatible intervals (same number or same length), are analyzed together or subsequently, the Fluency Analysis will build an interactive graph showing all of these plotted together. The single idfx graph will always be generated and saved to an image file for each idfx.
	By default, the fluency analysis counts every keystroke. This checkbox allows the user to restrict the fluency analysis to "productive keystrokes" keypresses which generate or remove actual characters.
single idfx	The single idfx graph shows a plot and a trendline for each optimum type. These can be toggled on and off. The zone in which the task optimum was achieved is marked as well. Exact y-values for the datapoints are displayed on mouseover.
multiple idfxs	The multiple idfx graph has a dropdown box for optimum type, to reduce cluttering. For each idfx file, the plot, trendline and Task Optimum Zone are displayed.

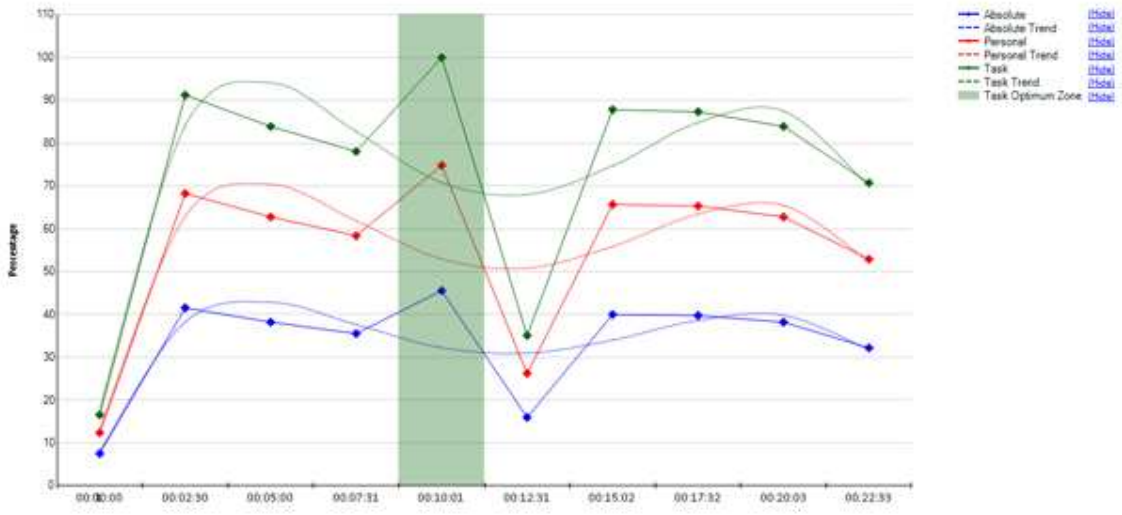


Figure 4. Single idfx

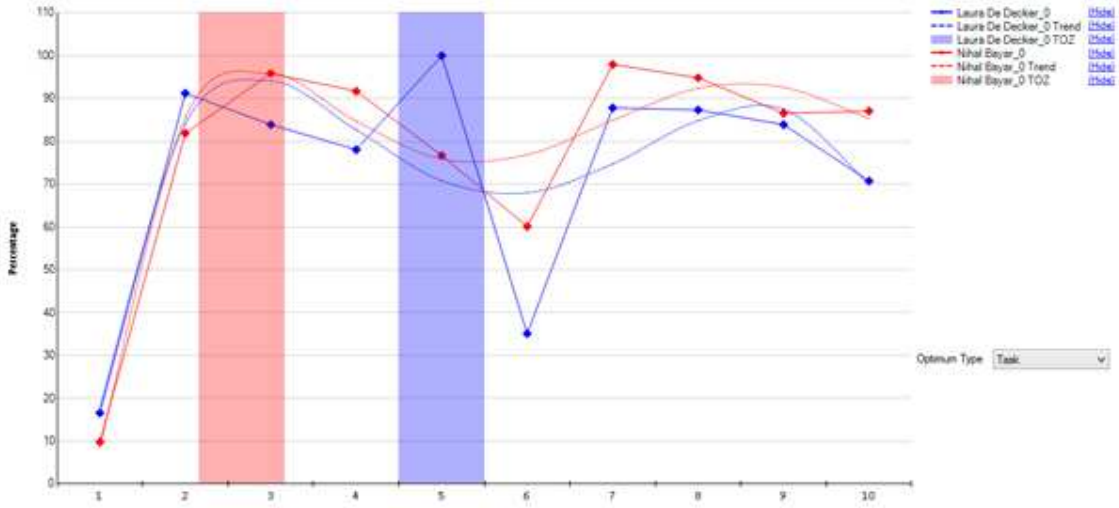


Figure 5. Multiple IDFX

Bigram analysis

The bigram analysis is a fine grained analysis that complements the overall pause analysis by providing more detailed information on a lower level. A bigram (or digram) is defined as a sequence of two consecutive keystrokes. In the analysis we mainly focus on bigrams that represent a combination of two (small) letters, e.g., 'ab' or 'th'. However, the analysis also provides information on the latency of other keystroke combinations, e.g., the transition between a character and a space, or between two numbers.

Measures

We report the latencies between keystrokes as follows:

Count	number of observations
Mean	mean length of the interkey pause
Std Dev	standard deviation of the interkey pause
Median	median length of the interkey pause
Minimum	minimum length of the interkey pause
Maximum	maximum length of the interkey pause
90% Interval Low	minimum pause length corrected for a 90% interval [remark 90% interval is defined as follows: mean - (1.645 * stdev)]
90% Interval High	maximum pause length corrected for a 90% interval [remark 90% interval is defined as follows: mean - (1.645 * stdev)]

Categories

The bigram analysis consists of four main sections:

1. Bigram categories
2. Bigram Speed
3. Bigram frequency
4. Lists
   - Alphabet Bigrams
   - Non-Alpha Bigrams

Bigram categories

In the first section, the pause related to the bigrams are categorized in 12 different categories, representing different types of bigrams in which one keystroke always represents a character:

CHAR_CHAR	pause between combination of two characters
CHAR_CHAR_SYLLABLE_BOUNDARY	pause between two characters that coincide with a syllable boundary
CHAR_SPACE	pause between a character and a space, usually the end of a word

SPACE_CHAR	pause between a space and a character, usually the beginning of wa word
CHAR_PUNCT_INTER	pause between a character and a punctuation mark, usually the end of a subordinate clause or a sentence
CHAR_PUNCT_INTRA	pause between a character and a punctuation mark within a word, e.g., a hyphen or a full stop in an abbreviation
CHAR_REV	pause between a character and a revision operation, e.g., a backspace
REV_CHAR	pause between a revision operation and a character, e.g., a prior deletion
SHIFT_CHAR	pause between a Shift key and a character, usually the beginning of a sentence or a proper name
CAP_CHAR	pause between a Capital letter and the Next character, usually the transition between the first and the second character of the first word of a sentence or within a proper name
NR_NR	pause between two numbers
UNSPECIFIED	any other key transition
Total	overall pause between any key combination

Bigram speed

In the second section, a selection is made of the five character bigrams that are produced, respectively the fastest and the slowest.

Bigram frequency

In the third section, we report the inter-bigram pauses based on frequency. We used the language specific frequency measures for bigrams as reported in the CELEX reports (English, Dutch, German; http://celex.mpi.nl/ - extra language specific lists can be added; remark: if the list is not available for a particular language the English Celex frequency list for bigrams is used). The language identification of a session is based on the reported language in the Session Identification of the log file.

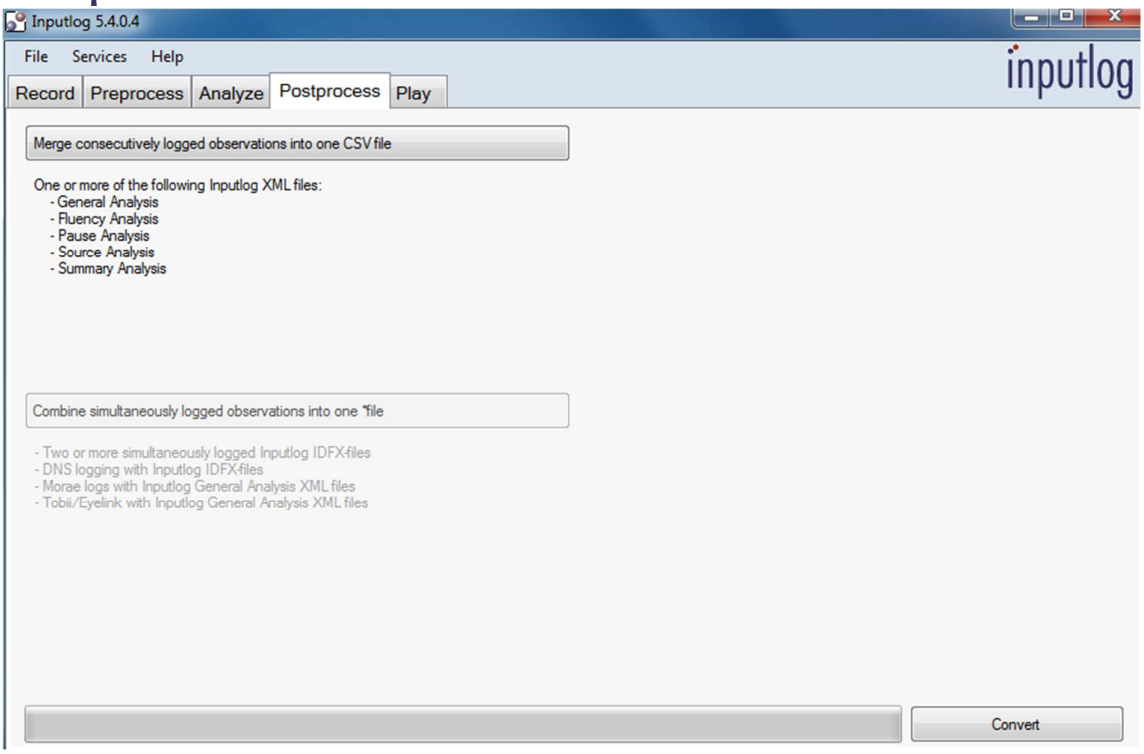
First the 5 bigrams with the highest frequency are reported, than those with the lowest frequency. Next the frequency is reported related to the number of occurrences of bigrams in the specific logging session.

Lists: Alphabetic and non-alphabetic bigrams

In the fourth section, two lists are shown:

1. Alphabet Bigrams: an alphabet list of all bigrams, ranging from aa to zz.
2. Non-alpha bigrams: a sorted list of all keystroke combinations in which either one or both of the keystrokes represent a non-alphabetic event.

Post-process

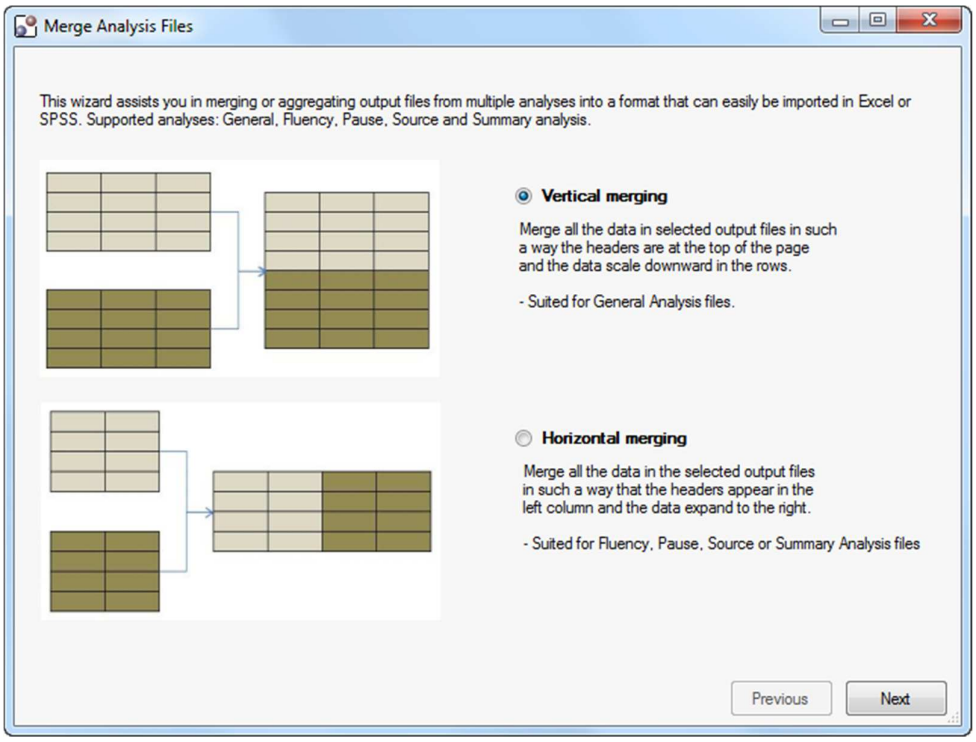


The ***Post-process*** module offers two options

option	description
merge consecutively logged observations into one CSV file	You can aggregate the output of various analyses into one file. This holds for: 1. General analysis 2. Fluency analysis 3. Pause analysis 4. Source analysis 5. Summary analysis
combine simultaneously logged observations into one *file	This feature needs to be implemented in a later stage.

Merge consecutively logged observations into one XML file

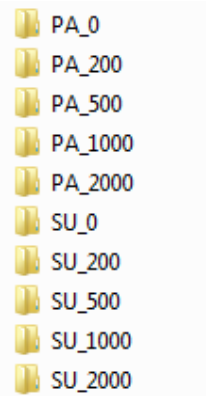
You can merge the output of various analyses into one file. You have two possibilities.

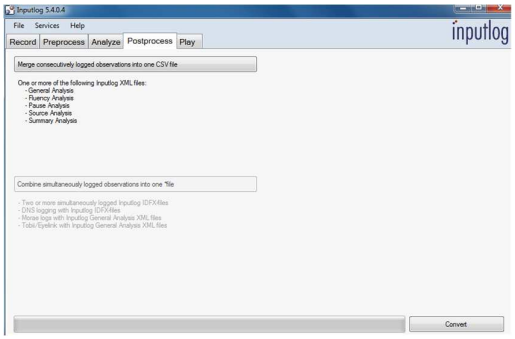
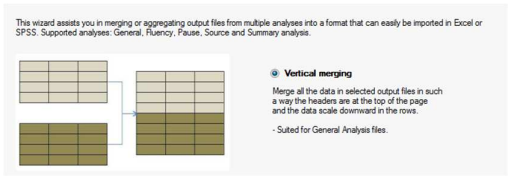
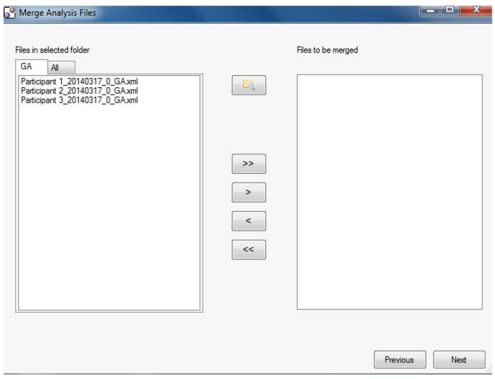


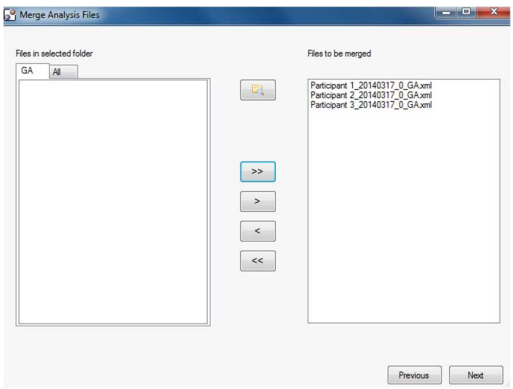
option	description
vertical merging	This option is suitable for the General analysis. You add data in a vertical way: the headers are at the top of the page and the data scale downward in the rows.
horizontal merging	This option is suitable for the following analyses: 1. Fluency analysis 2. Pause analysis 3. Source analysis 4. Summary analysis

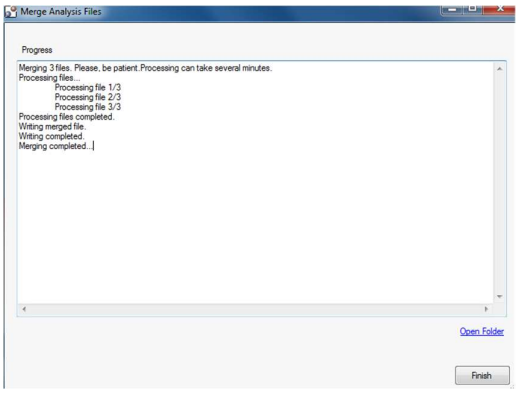
Vertical Merging

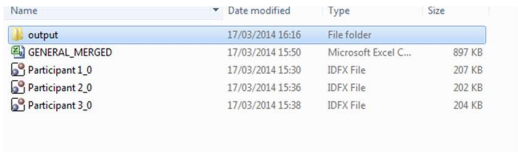
Tip: In order to merge various output files you can opt to save the outputs already in separate folders. You can find in each logging a file with 'output'. You can simple copy the general analysis to a file where you can put all the general outputs together. You can do this with all other analyses too.



	1. Click Merge consecutively logged observations into one CSV file
	2. Select Vertical merging 3. Click Next
	4. Browse to the General analysis files you would like to merge 5. You can see above that you are merging GA (General analysis). 6. Use the arrows to move the files to the right part under Files to be merged 7. Click Next



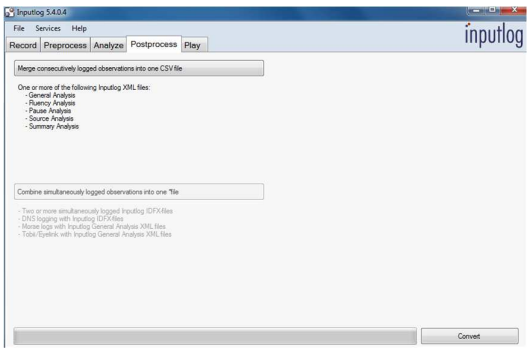




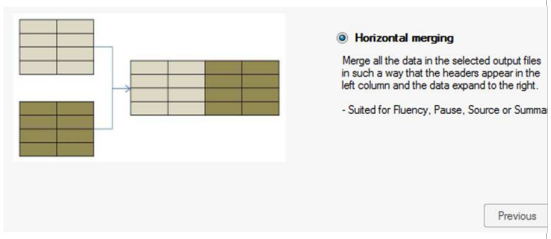
8. Click Open folder
   You can find here the CSV-file. It will be located in the same folder where the idfx are located.
9. Click Finish to proceed with a new merging process

Horizontal Merging

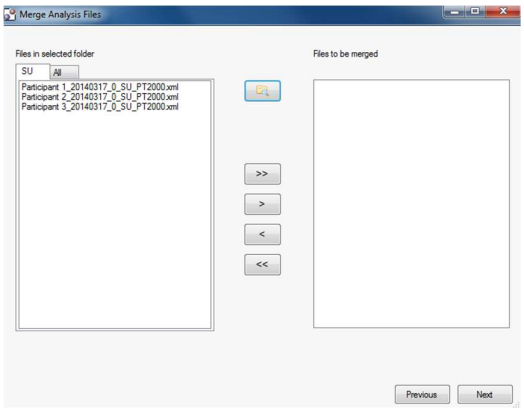
Tip: In order to merge various output files you can opt to save the outputs already in separate folders. You can find in each logging a file with 'output'. You can simple copy the general analysis to a file where you can put all the general outputs together. You can do this with all other analyses too.



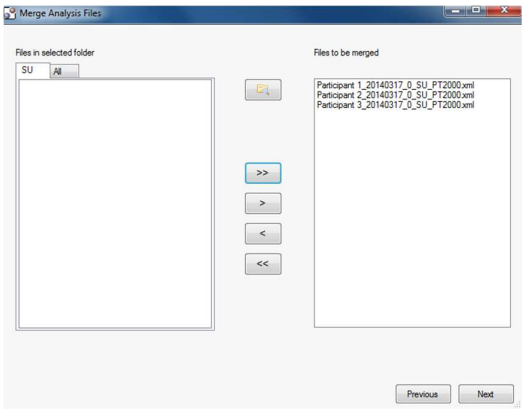
1. Click Merge consecutively logged observations into one CSV file

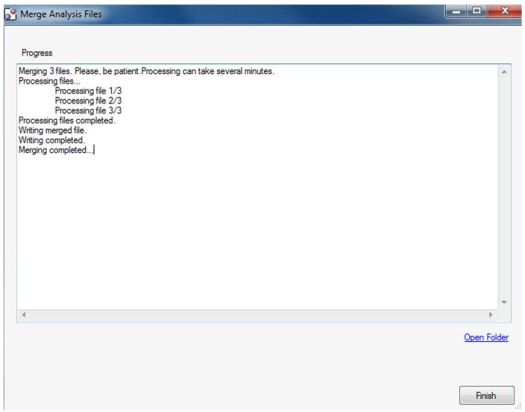


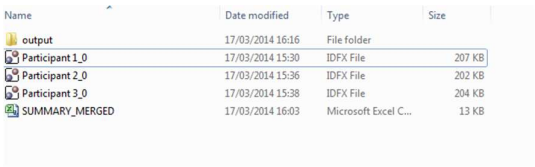
2. Select Horizontal merging
3. Click Next



4. Browse to the files you would like to merge
5. You will see that Inputlog detect which analysis you want to merge.
6. Use the arrows to move the files to the right part under Files to be merged
7. Click Next







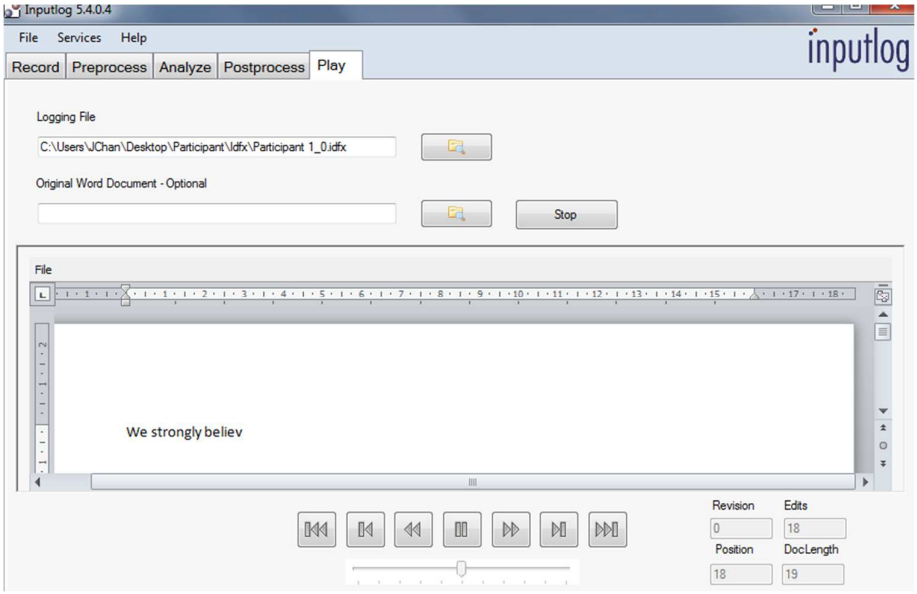
8. Click Open folder
   You can find the CSV-file here. It will be placed in the same folder where the idfx are located.
9. Click Finish to proceed with a new merging process

Combine simultaneously logged observations in to one *file

This feature needs to be implemented in a later stage.

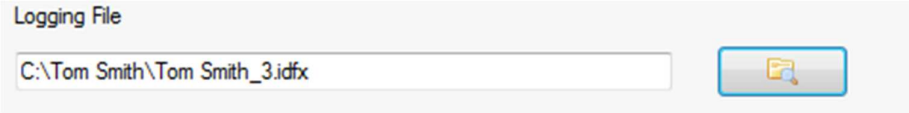
Play

In the tab Play you can see a reconstruction of the writing process. With the different buttons you can adapt the speed.



Logging File

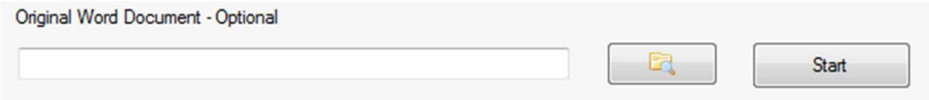
In the tab Logging File you select the file you would like to replay.



Important: The file has to be in the same Data Structure as the Word log document.

Original Word Document - Optional

When you want to replay a reconstruction with the Linguistic analysis, Inputlog automatically saves it into the same file where the original document is located.



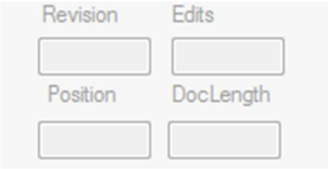
Play Speeds

You can adapt the speed with the following buttons.

	You will go to the beginning of the writing process.

	You will go one revision back in the writing process.
	You rewind the process in the speed setting that you have selected.
	You can replay the writing process in its natural pace (or any other pace that you would like).
	You can fastforward with this button.
	You will go one revision forward in the writing process.
	You will go directly to the end of the process.

In the right corner you can see the revisions and edits of the writing process. It shows in which positions the changes happen and where in the document length they are located.



Revision	Related number of revision of Revision Matrix
Edits	Related number of edit of Revision Matrix
Position	Related cursor position of General analyses
DocLength	Related DocLength of General analyses

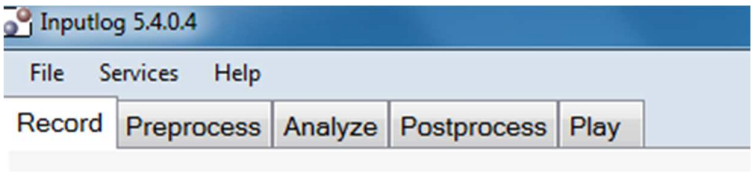
Toolbar

You have three tabs in the toolbar:

1. File
2. Services
3. Help

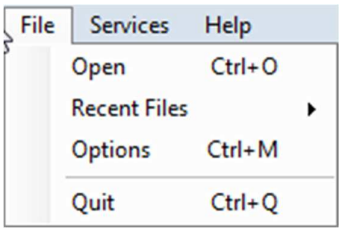
File

With this tab you can choose how to open the idfx file.



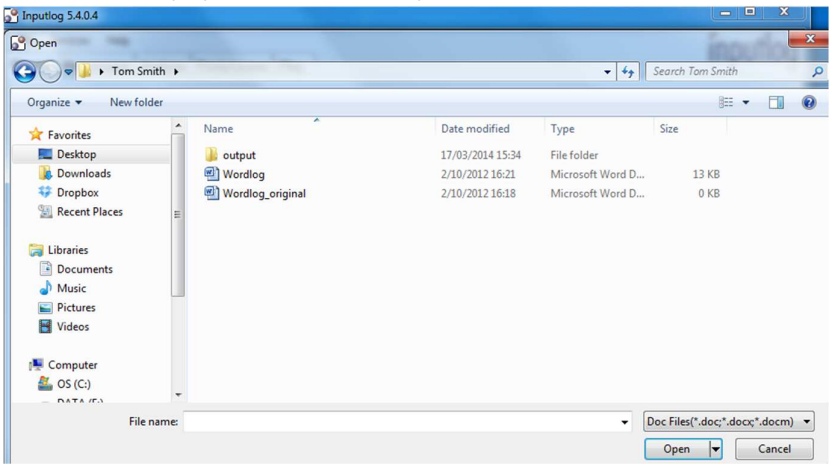
There are four options in the File-tab.

1. Open
2. Recent files
3. Options
4. Quit



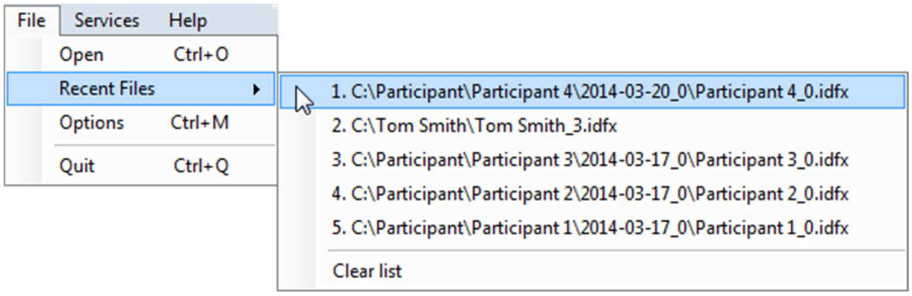
Open

You can directly open the idfx with'Open'.



Recent files

You can quickly go into the last files you have opened before.



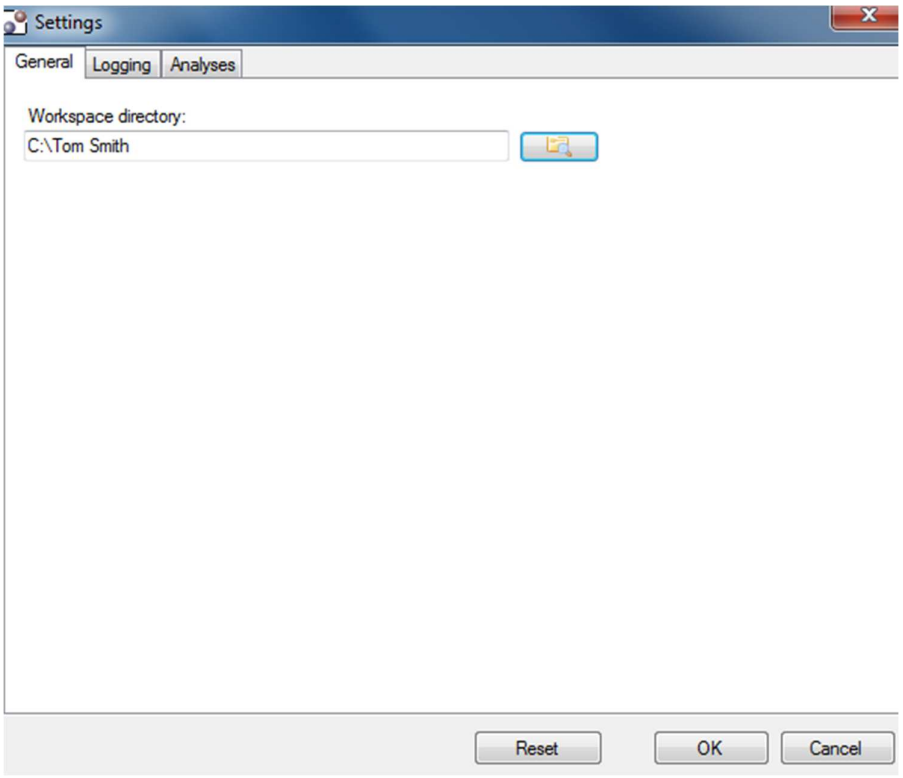
Options

You can change the settings. There are three options:

1. General settings
2. Logging settings
3. Analyses settings.

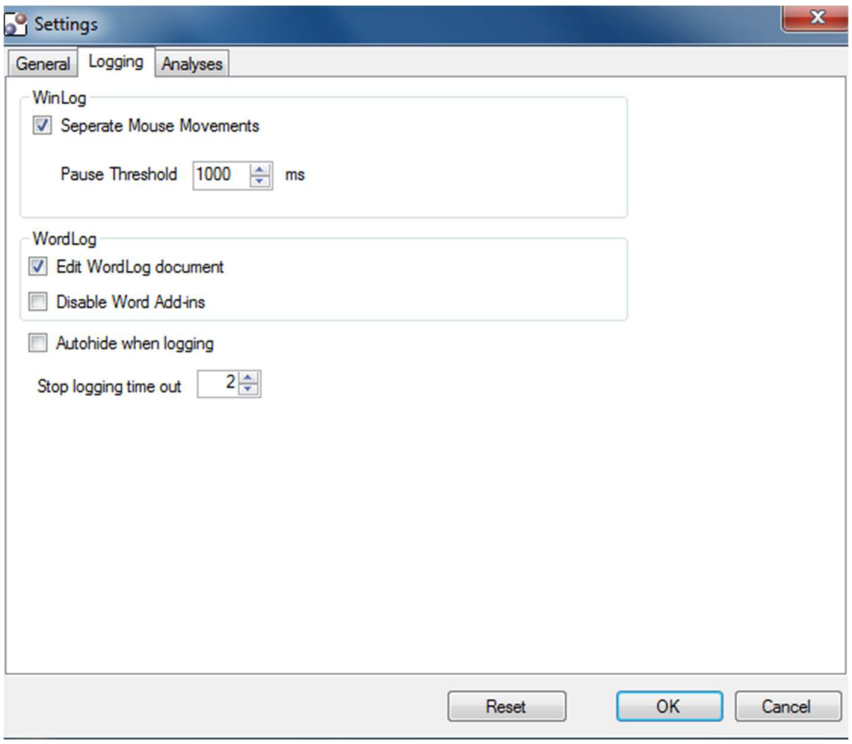
General

With this tab you can change the place where the loggings and outputs will be located. You can choose where all the data will be put in to. Change the workspace directory and continue with OK.



Logging

In the tab Logging you can define the Winlog and Wordlog settings.



Winlog settings

With this tab you can switch off the seperate mouse movements and change the pause threshold.

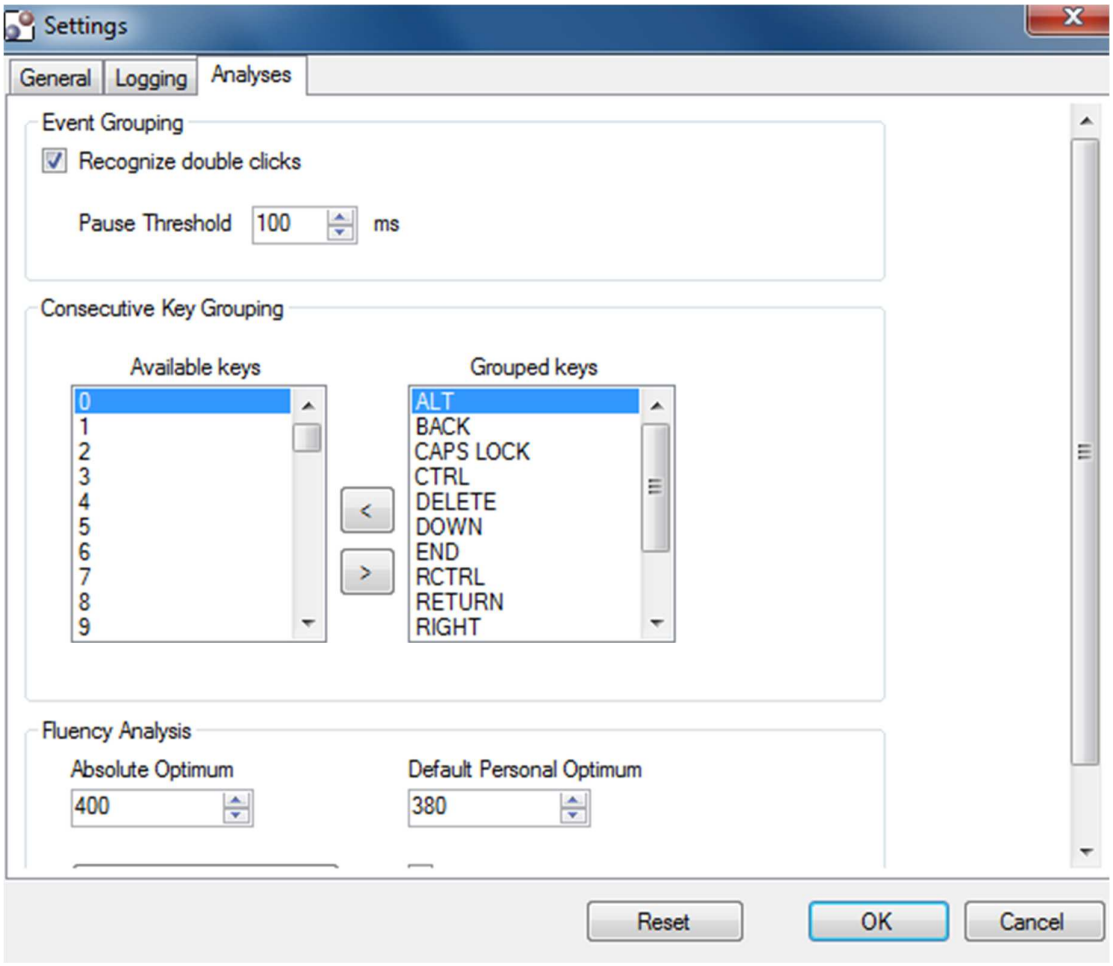
Wordlog settings

In this tab you can change different options:

1. Edit Wordlog document: this option is default and should always be selected. This option enables you to edit your document during the writing process. If this option is unselected NO information of your logging session in Word is saved. Only other Windows actions will be logged.
2. Disable Word Add-ins: you can disable add-ins of external programs. It happens that other programs cause a malfunction of Inputlog (e.g., Endnote). If you select this option these add-ins are automatically disabled.
3. Autohide when logging: you can hide the icon in the traybar of Inputlog while logging the writing process
4. Stop logging time out: time to wait before a stop is forced (not implemented)

Analyses

In the Analyses tab you can group events and define personal optima for the fluency analyses.



Event grouping

You can opt to automatically recognize double clicks as such. Remark: you need to calculate this threshold yourself.

Consecutive key grouping

You can opt to show keystrokes in the analyses as grouped keys. You have two options; e.g.,:

- LEFT SHIFT H for a capital H, or
- H for a capital H

In the second instance, the keys are grouped. This setting is advisable for most general writing process research. If you are interested in very detailed keystroke information you better not group the keys.

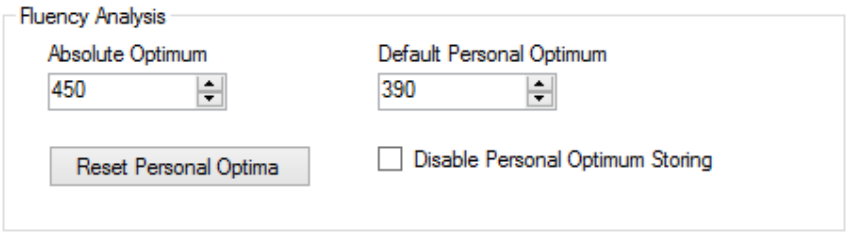
5	mouse	LEFT Click		
6	mouse	LEFT Click		
7	mouse	Movement		
8	keyboard	q	0	1
9	keyboard	BACK	1	2
10	mouse	Movement		
11	mouse	LEFT Click		
13	keyboard	LSHIFT	0	1
14	keyboard	LSHIFT	0	1
15	keyboard	LSHIFT	0	1
16	keyboard	LSHIFT	0	1
17	keyboard	LSHIFT	0	1
18	keyboard	LSHIFT	0	1
19	keyboard	LSHIFT	0	1
20	keyboard	LSHIFT	0	1
21	keyboard	LSHIFT	0	1
22	keyboard	LSHIFT	0	1
23	keyboard	LSHIFT	0	1
24	keyboard	LSHIFT	0	1
25	keyboard	LSHIFT	0	1
26	keyboard	LSHIFT	0	1
27	keyboard	LSHIFT	0	1
28	keyboard	LSHIFT	0	1

5	mouse	LEFT Click		
6	mouse	LEFT Click		
7	mouse	Movement		
8	keyboard	q	0	1
9	keyboard	BACK	1	2
10	mouse	Movement		
11	mouse	LEFT Click		
68	keyboard	T	0	1
69	keyboard	h	1	2
70	keyboard	r	2	3
71	keyboard	e	3	4
72	keyboard	BACK	4	5
73	keyboard	BACK	3	4
74	keyboard	e	2	3
75	keyboard	r	3	4
76	keyboard	e	4	5
77	keyboard	SPACE	5	6
78	keyboard	i	6	7
79	keyboard	s	7	8
80	keyboard	SPACE	8	9
81	keyboard	a	9	10
82	keyboard	SPACE	10	11

Via Reset you choose the default settings (by default the keys are grouped)

Fluency analyses

In the analyses settings for the fluency analyses you can indicate various types of optima.



1. **Absolute Optimum:** Used as absolute optimum in all fluency analyses
2. **Default Personal Optimum:** Used as personal optimum in analysis if no personal optimum is stored
3. **Reset Personal Optima:** Deletes all previously stored personal optima

Disable Personal Optimum Storing: Disable storage of personal optima. This does not disable reading of stored optima. To completely disable personal optimum history, you should also reset it or remove/rename personal optimum file.

Note: when analysing writing fluency at a lower, local level (for example sentence or paragraph level) it is important to control for continuous pressed keys (for example holding on to the BACK-key to delete an entire sentence). When one and the same key is pressed for a longer time, this could distort the image of the writer's fluency, because here pause times are set to 32ms for all events within that same keystroke. This would suggest highly fluent writing, while this is not necessarily the case. It is therefore important to control for this issue and, if needed, filter these events.

Services

Via Inputlog you can activate the server of the Linguistic analyses Inputlog. You can also browse to this page via: http://Inputlog.ua.ac.be/WebSite/Account/Register.

You can find further details on the

Copy task

The copy task analysis provides a carefully designed set of perspectives to explore and characterize a participant's motoric keyboard interaction. The default copy task - already available for a set of languages - is designed in such a way that it is possible to adequately address these characteristics, and use them in further analyses (e.g., as a co-variate in variance analyses, or in regression and multilevel analyses).

Characteristics	The copy task analysis addresses the following bigram characteristics: ▪ Frequency ▪ Adjacency ▪ Repetition ▪ Hand combination
Bigram Frequency	In the copy task word and sentence combinations are designed in such a way that it is possible to isolate - language specific - high and low frequency bigrams in the analysis. The frequency is based on large corpora (e.g., Subtlex or CELEX). ▪ High frequency: 30 % highest bigram percentiles ▪ Low frequency: 50 % lowest bigram percentiles
Bigram adjacency	When two keys are next to each other on the keyboard (left, right, up, down) they are labelled as 'adjacent'. this labelling is keyboard specific. Therefore, the identification of the keyboard lay-out at the start of each session is crucial.
Repetitive bigrams	When a bigram is produced by pressing the same key twice, we consider it as a repetitive bigram.
Bigram hand combination	To produce a bigram, the following hand combinations could be used, taking into account the order: ▪ Left - Right hand combination ▪ Right - Left hand combination ▪ Left - Left hand combination ▪ Right - Right hand combination Remark: The midzone of the keyboard is not included in the analysis, as - dependent on the typing expertise - the hand that is used to type these keys could vary (more often).

Reported measures

For every sub analysis within the copy task analysis, the following descriptive measures are reported:

- Count (targeted): number of characters typed that correspond to a character that was prompted in one of the respective copy task components
- Count (not targeted): number of characters typed that do not correspond to a character that was prompted in one of the respective copy task components
- Mean interkey interval (IKI): arithmetic mean (in milliseconds) of the targeted bigrams within the scope of the respective analyses
- Standard deviation: standard deviation of the targeted bigrams within the scope of the respective analyses
- Median: median (in milliseconds) of the targeted bigrams within the scope of the respective analyses
- Log mean trimmed: 95% log converted trimmed geometric mean of the targeted bigrams within the scope of the respective analyses
- Coefficient of variation: calculated coefficient of variation (based on mean and stdev) 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, viz. 60 000 (ms) / mean IKI (ms)
- Absolute CPM: time based calculation of (theoretical number of) characters per minute based on an extrapolation of the time on task within a copy task component (e.g. for the default sentence copy task, participants are required to type for 45 s. In that case the time between the first and the last key, i.e. approximation of 45 s is used to calculate a time based extrapolation of the CPM).

Session information

The copy task analysis starts with some meta information about the participant and the session. This report is based on the questions that are presented at the beginning and end of the copy session: particpant information, hard and software used, handedness score, language and education.

Report overview

The copy task analysis provides a detailed analysis split up in the following sections:

- Correctness score
- Synthesis of InterKey intervals (IKI)
- Individual Component analysis
- Aggregated characteristics

Correctness

When execuring the copy tasks, participants strive to find an optimal trade-off between speed and correctness (cf. also basic instruction). The copy task analysis opens with a summary of correctness scores, comparing the targetted bigrams (in the prompted texts) with the characters actually typed.
The analysis first reports the correctness scores for each component seperately and then an aggregated synthesis is reported, both overall and for the group of selected components (see asterisks in component analysis).

Example:

Synthesis

The 'Overall Synthesis of the InterKey Intervals (IKI)' in the output summarizes the overall performance of the copy task.
Data analysis is represented at three levels:

- Targeted bigrams: selects all bigrams that are targeted by one of the prompts in the corresponding copy task
- High frequency bigrams: selects all high frequency (HF) bigrams targeted by one of the prompts in the corresponding copy task

- Selected components bigrams: limits the analysis of targeted bigrams to those occurring in the selected components of the corresponding copy task (see documentation on 'Copy task creator'). For the default copy tasks these selected components refer to the three HF bigram word combination tasks.

Example:

Component analysis

In the component analysis the following elements are reported:

- Overall analysis: analysis of all the produced characters produced in each copy task component
- Trial cut > 2: analysis of all the produced characters produced in each repetitive component task, excluding those in the first two trials
- Time filtered 10%: analysis of all the produced characters produced in each component task, excluding the first 10% time period of each task execution
- Trial report: analysis of all the produced characters produced in each trial within a component task.

Example:

Characteristics

In the characteristics analysis the following elements are reported:

- Frequency: overall analysis of all intra bigram intervals for resp. high and low frequent bigrams (HF and LF)
  Definition of high and low frequent bigrams: see above/link.
- Frequency (Trial Cut >2): overall analysis of all intra bigram intervals for resp. high and low frequent bigrams (HF and LF), excluding those produced in the first two trials of each repetitive component
- Frequency (Only repetitions & Trial Cut >2): overall analysis of all intra bigram intervals for resp. high and low frequent bigrams (HF and LF) in repetitive copy task components, excluding those produced in the first two trials of each of those components

Example:

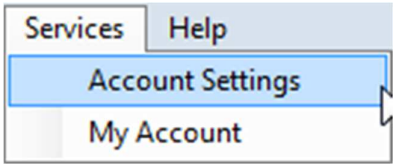
- Hand combination : overall analysis of all intra bigram intervals for the different hand combinations needed to produce a bigram, resp. left-right hand combination (LR), right-left hand combination (RL), left-left hand combination (LL) and right-right hand combination (RR).
  Definition of high and low frequent bigrams: see above/link.
- Hand combination Frequency (Trial Cut >2): overall analysis of all intra bigram intervals for the different hand combinations needed to produce a bigram, excluding the first two trials of the repetitive components.
- Adjacency: overall analysis of all intra bigram intervals for non-adjacent (False) and adjacent (true) keys.
- Adjacency (trial cut >2): overall analysis of all intra bigram intervals for non-adjacent (False) and adjacent (True) keys, excluding the first two trials of repetitive components.
- Repetition: overall analysis of all intra bigram intervals for non-repetitive (False) and repetitive (True) keys.
- Adjacency (trial cut >2): overall analysis of all intra bigram intervals for non-repetitive (False) and repetitive (True) keys, excluding the first two trials of repetitive components.

Linguistic analyses in research articles under the section Read More.

Account settings

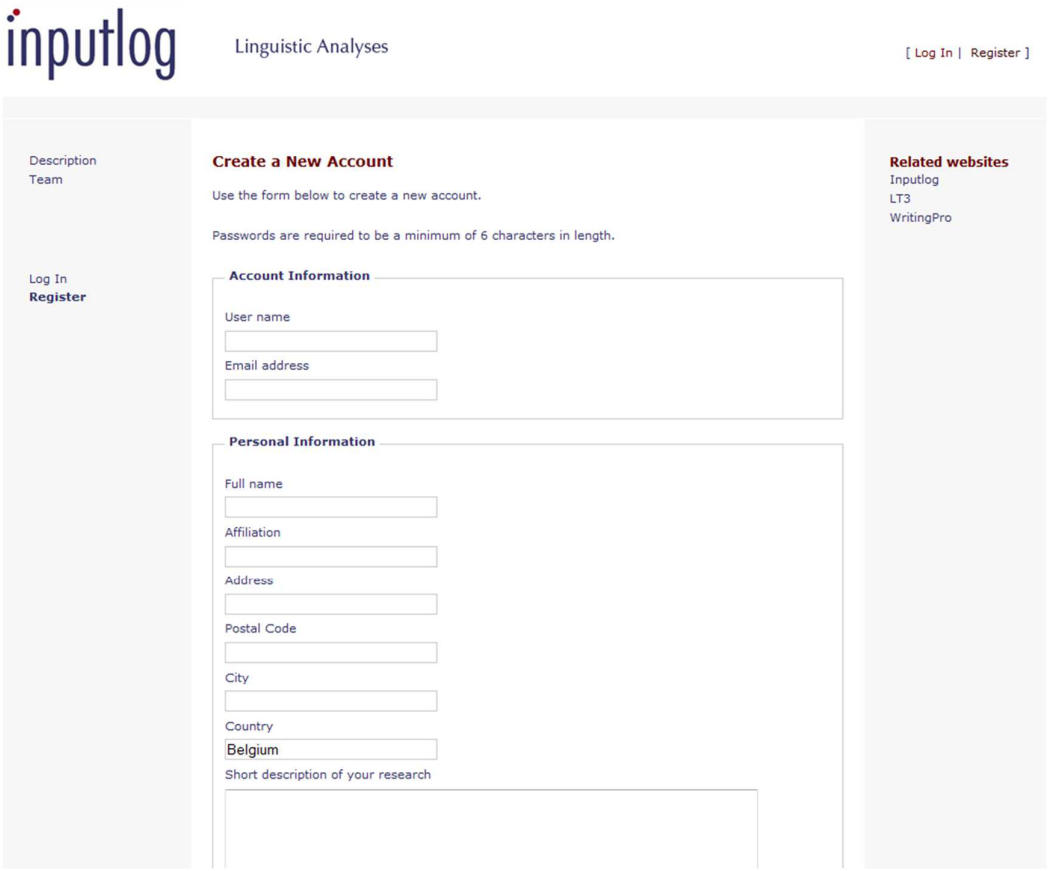
You can log in in two ways.

1. You can do it before you analyze the data. Therefore you go to 'Services' and click on 'Account settings'.
2. Or when you add the linguistic analysis, there will be a window automatically appears that ask you to log in before proceeding with the linguistic analysis.



My account

As mentioned before in the linguistic analysis' section, you have to registrate on http://Inputlog.ua.ac.be/WebSite/Account/Register, before you can use the linguistic analysis. You have to complete the registration form first and then you will receive a personal password in your mailbox. With the registration you create your own account where you can see all the analysis you've made with Inputlog. This step is only required with the linguistic analysis.



When the linguistic analysis is completed, you will receive an email that confirms your analysis has been completed.

Help

Inputlog help

You can find the latest version of the Help-file here in three formats:

1. WinHelp
2. Web help
3. Full PDF-version

Inputlog tour

You can find a short example of how you can record and analyze a logging session via Inputlog.

Inputlog on the web

You can find a link to the website of Inputlog.

Check for updates

Via this link you go to the download page of Inputlog. New versions are published here. You can also mail marielle.leijten@uantwerpen.be or luuk.vanwaes@uantwerpen to obtain more information about new functionalities. Not every in-between version is published on the Inputlog-website.

About Inputlog

Via this link you can learn more about the history and contributors in various fields of the program. Please note that Inputlog is a research instrument that is free to the research community. If you publish research in which Inputlog has been used please refer to:
Leijten, M., & Van Waes, L. (2013). Keystroke Logging in Writing Research: Using Inputlog to Analyze and Visualize Writing Processes. Written Communication 30(3), 358–392 | DOI: 10.1177/0741088313491692

Donate to Inputlog

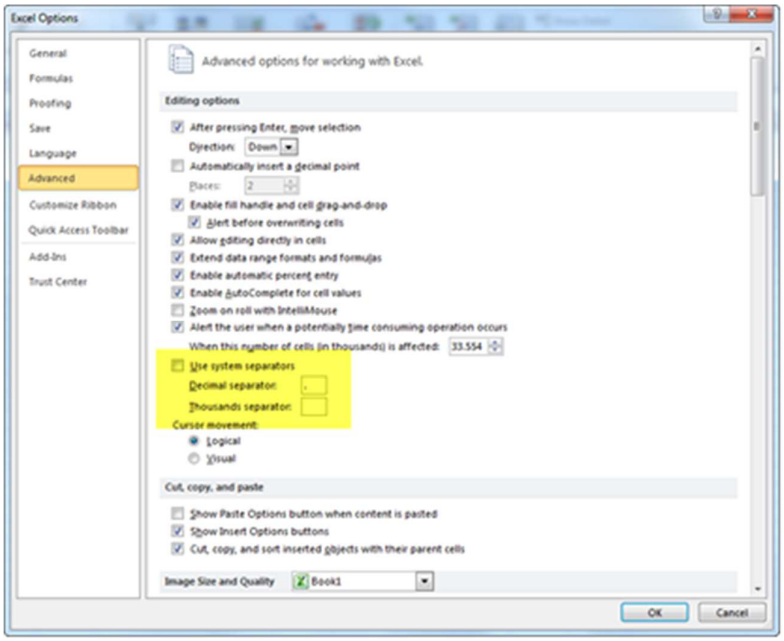
If you would like to support the development of Inputlog, you can contact marielle.leijten@uantwerpen.be or luuk.vanwaes@uantwerpen.be to discuss possibilities. You can find a list of researchers who invested into Inputlog on the website.

The investments of individual researchers range from € 200 to € 2000.

Import data in Excel

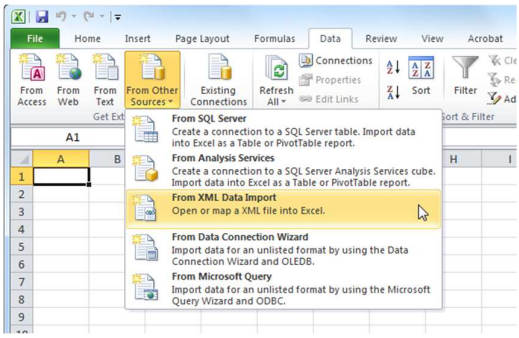
You can import data from the XML-files into Excel. Be aware that each Excel program uses its own settings on your computer. To import the data of Inputlog you need to make sure that Excel can read the data as numbers. For Inputlog data you use:

1. Decimal seperators: . (full stop)
2. Thousands seperators: _ (space)

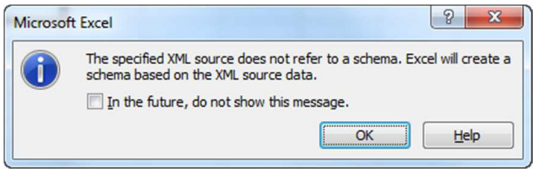


Import original XML-files in Excel

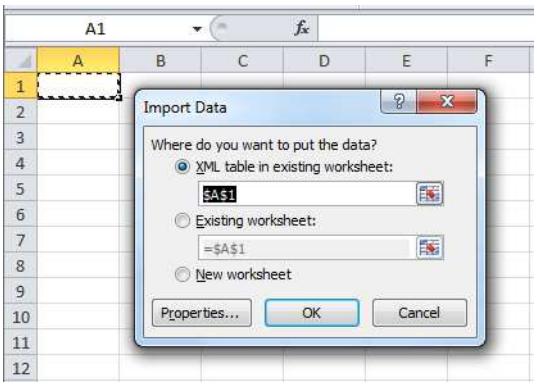
If you would like to import an original XML file into Excel you can use the data import function.



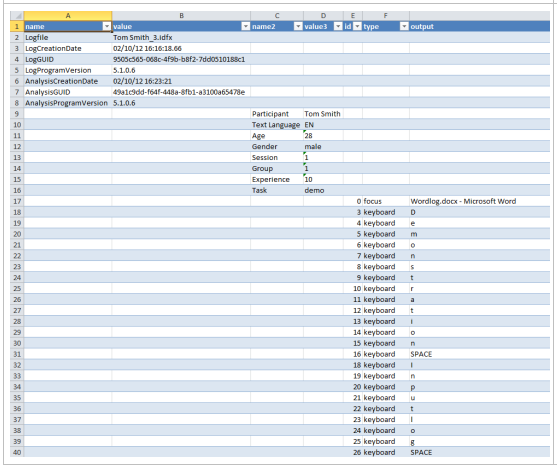
1. Select Data in the upper menu bar
2. Select From other Sources
3. Select from XML Data Import



4. Click Ok



5. Position your cursor in cell A1
6. Click Ok



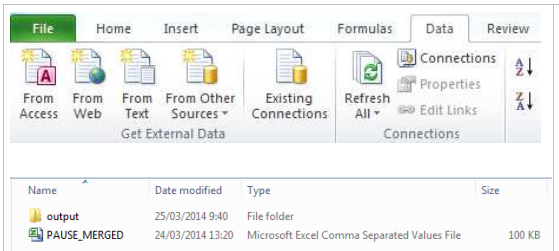
All the data are now inserted in the first worksheet of Excel

Tip: Copy the relevant data in a new worksheet (calculations) and keep the original data (original). Create new colums to the right of your data for your variables (participant, age, etc.)

More examples can be found on http://www.inputlog.net and www.writingpro.eu

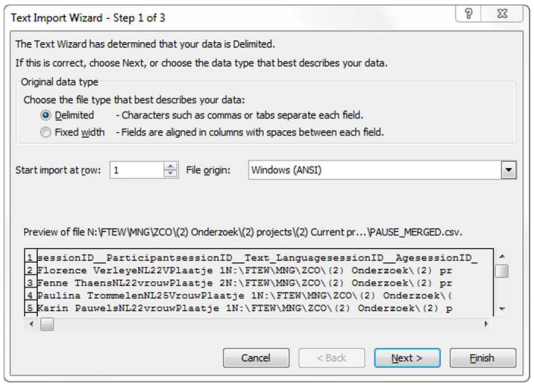
Import merged CSV-files in Excel

If you would like to import merged Inputlog data located in a CSV-file into Excel you can use the data import function.

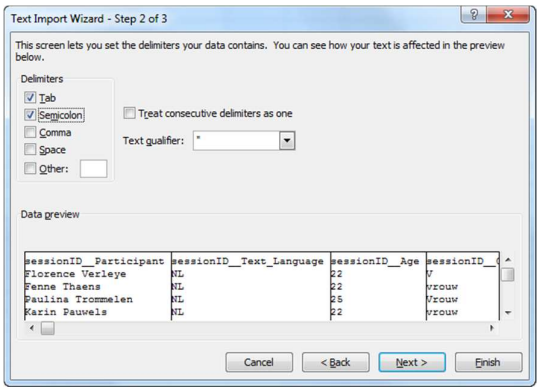


1. Select Data in the upper menu bar
2. Select From Text
3. Browse to *_merged.CSV file double-click the file

The Text Import Wizard automatically opens

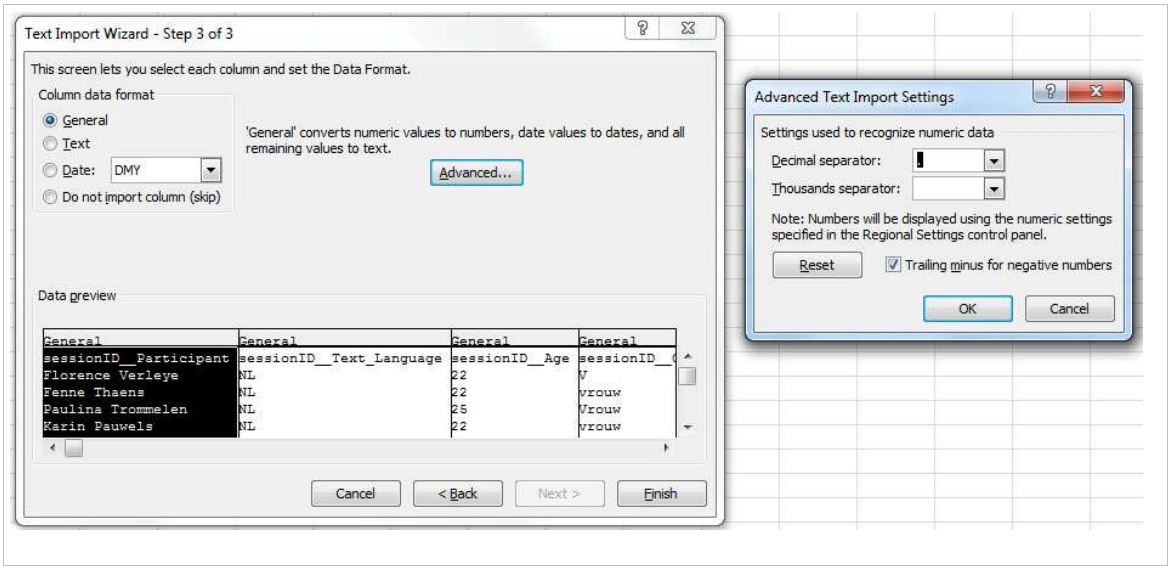


4. Select Delimited
5. Click Next



6. Select the Tab and Semicolon Delimiters
7. Choose " as Text Qualifier
8. Optional: Click Next if you haven't changed the settings of Excel to understand Inputlog data as numeric data.
9. Click Finish

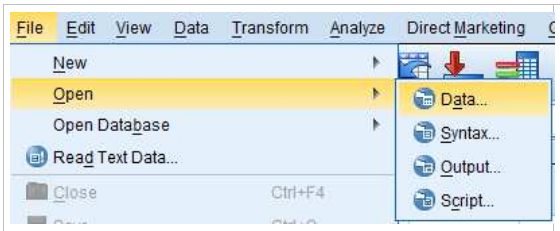
Advanced Text Import Settings:



Tip: You can also start this procedure by double-clicking the *_Merged.CSV-file. The data is imported instantcly as text information. Select column A and select under Data the option Text to Columns. The same Text Import Wizard will open as described before.

Import data in SPSS

If you would like to import merged datafiles you better work via Excel. In Excel you can easily combine various CSV-files of the summary, pause, source and fluency analyses of various thresholds. This file can be opened via SPSS.



1. Select File, Open and Data
2. Browse to the Excel file
3. Change Files of type into XLSX in order to open an Excel file.
4. Select the worksheet you would like to open
5. Select Read variable names from first row
6. Click ok

Glossary

To be added

Index