

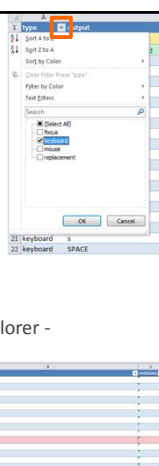
First impression of your data

You can filter data in order to search important occurrences:

- “Significant” pauses (yellow)
- Program switches (green)

You can code data into groups of information

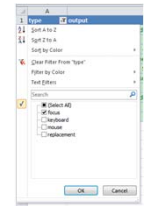
- Revisions (red - automatically)
- Multiple sources (MS Word versus Internet Explorer - manually)



type	output	programs (switches)
1	WordLog - Microsoft Word	main document
3	TASKBAR	internet
61	Mozilla Firefox Start Page - Mozilla Firefox	internet
95	colplay lyrics - Google zoeken - Mozilla Firefox	internet
101	colplay songteksten Songteksten.nl - Your Lyrics Source - Mozilla Firefox	internet
107	The Scientist Colplay songtekst Songteksten.nl - Your Lyrics Source - Mozilla Firefox	internet
116	TASKBAR	main document
118	WordLog - Microsoft Word	main document
140	TASKBAR	main document
144	Inputlog 5.0.1.6	inputlog

Filter focus events

- Filter in column A on focus events



type	output	programs (switches)
1	WordLog - Microsoft Word	main document
3	TASKBAR	internet
61	Mozilla Firefox Start Page - Mozilla Firefox	internet
95	colplay lyrics - Google zoeken - Mozilla Firefox	internet
101	colplay songteksten Songteksten.nl - Your Lyrics Source - Mozilla Firefox	internet
107	The Scientist Colplay songtekst Songteksten.nl - Your Lyrics Source - Mozilla Firefox	internet
116	TASKBAR	main document
118	WordLog - Microsoft Word	main document
140	TASKBAR	main document
144	Inputlog 5.0.1.6	inputlog

Copy values

- =IF(T3>0;T3;U2)

T	U
programs (switches)	programs (continued)
inputlog	inputlog
main document	main document
main document	main document
main document	main document

- Easily adaptable to other situations in which you work with switches)

Number of characters produced (in MS Word)

- Process measure needed to make chart in which total text production is important.

M	N	W
position	doclength	characters produced in MS Word
0	0	1
1	1	2
2	2	3
3	3	4
4	4	5
5	5	6
6	6	7
7	7	8
8	8	9
9	9	10
10	10	11
11	11	12
12	12	13
13	13	14
14	14	15
15	15	16
16	16	17
17	17	18
18	18	19
19	19	19

Summarize data

- Pivot table

type	output	startLock	endLock	endLockTime
1	inputlog	318	00:00:00.358	1182
2	movement	3746	00:00:01.344	3796
3	movement	3746	00:00:01.344	3796
4	focus	WordLog dock - Micro		
5	movement			
6	movement			
7	keyboard			
8	keyboard			
9	keyboard			
10	keyboard			
11	keyboard			
12	keyboard			
13	keyboard			
14	keyboard			
15	keyboard			
16	keyboard			
17	keyboard			
18	keyboard			
19	keyboard			
20	keyboard			

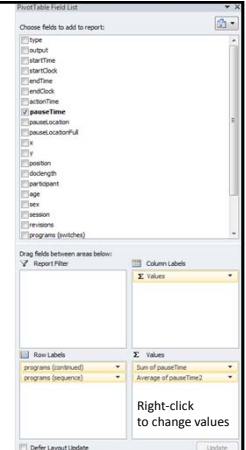


Dotted line indicates selection

Summarize data

- You can easily construct tables like:

Row Labels	Sum of pauseTime	Average of pauseTime2
inputlog	4852	1617
0	3822	3522
4	1030	515
internet	19173	349
2	19173	349
main document	26796	311
1	17614	297
3	8892	342
Grand Total	50731	352



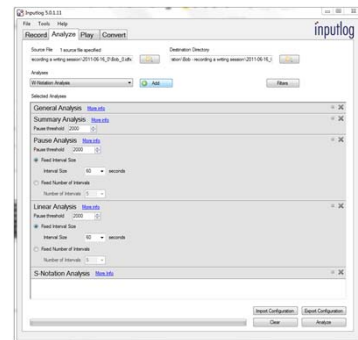
Right-click to change values

Understand output files (handout)

Based on "This is a sentence"-case

powerful simple savvy legendary bright
 introspective social nourishing vast
 energizing uplifting funny shared
 beautiful provocative insightful mad
 current dynamic engaging lasting
 elevating thoughtful poignant
 bold original detailed vivid giving

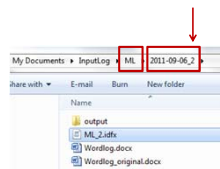
Output files



Data structure

Inputlog files are structured as follows

1. Folder participant ML
2. Folder date & consecutive number
3. Folder output
4. Folder output X¹
5. X.idfx
6. Wordlog.docx
7. Worldlog_original.docx



Data structure

The Inputlog output folder is structured as follows

1. Folder Images
2. Folder Scripts
3. Folder Style
4. All generated files

Name	Date modified	Type	Size
Images	6/09/2011 15:12	File folder	
Scripts	6/09/2011 15:12	File folder	
Style	6/09/2011 15:12	File folder	
ML_20110906_2_FA.xml	6/09/2011 15:12	XML File	3 KB
ML_20110906_2_GA.xml	6/09/2011 15:12	XML File	33 KB
ML_20110906_2_IA_FT2000_FL60.xml	6/09/2011 15:12	XML File	4 KB
ML_20110906_2_PA_FT2000_FL60.xml	6/09/2011 15:12	XML File	4 KB
ML_20110906_2_SA_FT2000.xml	6/09/2011 15:12	XML File	6 KB
ML_20110906_2_SNA.xml	6/09/2011 15:12	XML File	2 KB

Summary analysis

- Open "This is a sentence that" > "Explanation calculations summary analyses.xlsx"

Examples

- Number of characters & spaces
- Total pause time varies upon chosen threshold
- Clusters & segments



Summary analysis

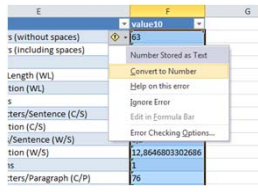
- Be aware that these automatic calculations might not be what you want to work with. E.g.
 - Active writing time is not the sum of all action times
 - Average pause time varies if you select a filter (e.g. All modes 24s, only Key 40s)

General advice

Check summary analysis with the purpose of your research question.
 Perform small test sessions to understand the output of the data.

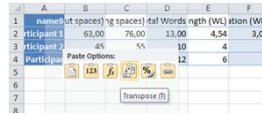
Calculate with summary analyses

- Convert text to numbers to be able to calculate with data



Prepare data for SPSS

- Add multiple participants to one dataset
- Select Variable names and values (column E:H)
- Copy & paste special (transpose)



Remark: be very careful in deciding which information you use for further analyses

Product information

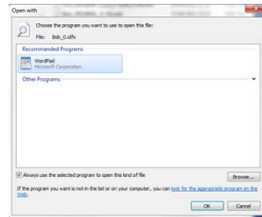
- Product information is retrieved from MS Word
- Process/product ratio

Product Information	
Words	
Total Characters (without spaces)	61
Total Characters (including spaces)	74
Total Far East Characters	0
Total Words	13
Paragraphs	
Total Paragraphs	0
Lines	
Total Lines	1
Pages	
Total Pages	1

Advanced data preparation¹

- Change data in XML-files
 - Change of letters (unknown character)
 - Change pause locations

open any general logging file by right-clicking on the xml-file



XML-data

- Example of general analysis (Bob)

```
<?xml version="1.0" encoding="utf-8"?>
<xml-stylesheet type="text/xsl" href="style/general_analysis.xsl"?>
<session
  <entry name="LogCreationDate" value="16/06/11 11:16:58.774" />
  <entry name="LogGUID" value="40ec4893-3507-4c9f-91da-081b0089fe8" />
  <entry name="AnalysisCreationDate" value="16/06/11 11:18:14" />
  <entry name="AnalysisGUID" value="d09f6612-80d4-4ae5-8586-bb231e1d5b31" />
  </meta>
  <sessionIdentification>
    <entry name="Participant" value="Bob" />
    <entry name="Age" value="33" />
    <entry name="Sex" value="1" />
    <entry name="Session" value="1" />
    <entry name="Group" value="" />
    <entry name="Experience" value="" />
  </sessionIdentification>
  <event>
    <type>Mouse</type>
    <output>Movements</output>
    <startTime>3324</startTime>
    <startClock>00:00:00.000</startClock>
    <endTime>4508</endTime>
    <endClock>00:00:04.508</endClock>
    <actionTime>896</actionTime>
    <pauseTime>3324</pauseTime>
    <pauseLocation%></pauseLocation>
    <pauseLocationFull>INITIAL PAUSE</pauseLocationFull>
    <id206/>
    </event>
  </event>
```

Advanced data preparation²

- Manipulate idfx
 - Remove beginning of file
 - Remove parts that you don't need in analysis
 - Split IDFX

(open file "Split_IDFX_files.xlsx")

IDFX-file

```
Output
[Movement][Move] [LCTRL][ENTER] [Movement][RSHFT][This is part [RSHFT]1: of an experiment
[LSHFT] [CTRL] [LCTRL][ENTER] [RSHFT][This is part [RSHFT]2: of an experiment[LSHFT].
[Movement][LCTRL][ENTER] [LCTRL][Click]
```

The example shows two parts of an experiment

- The data needs to be split after experiment 1
- LCTRL + ENTER is mark of breaking point in file
- Search in General logging file for output & position

General logging file	IDFX-file
43 keyboard LCTRL 33	829 <event type="keyboard" id="42">
	829 <part type="wordlog">
	830 <position>33</position>
	831 <documentLength>33</documentLength>
	832 <replay>True</replay>
	833 </part>
	834 <part type="winlog">
	835 <startTime>522765</startTime>
	836 <endTime>5228170</endTime>
	837 <key>VK_LCONTROL</key>
	838 <value />
	839 <keyboardstate />
	840 </part>
	841 </event>

IDFX-file

- Example of idfx-file (Bob)

```
<xml version="1.0" encoding="utf-8">
<log>
<meta>
<entry>
<key>_logCreationDate</key>
<value>16/06/11 11:16:58.774</value>
</entry>
<entry>
<key>_GUID</key>
<value>40ec449b-3507-4c8f-91da-081b908f6c</value>
</entry>
<entry>
<key>_logRelativeCreationDate</key>
<value>76218625</value>
</entry>
</meta>
<event type="mouse" id="0">
<part type="winlog">
<startTime>7622447</startTime>
<endTime>7622333</endTime>
<keyID>6</keyID>
<type>mouse</type>
</part>
</event>
...
<event type="keyboard" id="13">
<part type="wordlog">
<position>3</position>
<documentLength>3</documentLength>
<replay>True</replay>
</part>
<part type="winlog">
<startTime>7622905</startTime>
<endTime>7622896</endTime>
<keyID>0</keyID>
<value>0</value>
<keyboardstate />
</part>
</event>
</session>
```

IDFX-file

The structure of the manipulated IDFX-file needs to be as follows:

- <log>
- <metadata>
- <session identification>
- <events N>
- </events N>
- </log>
- If you delete information in the beginning of the events than you need to update the LogRelativeCreationDate

IDFX-file

- If you delete information in the beginning of the events than you need to update the LogRelativeCreationDate

```
<key>_logRelativeCreationDate</key>
<value>521359</value>

<event type="keyboard" id="43">
<part type="wordlog">
<position>33</position>
<documentLength>33</documentLength>
<replay>True</replay>
</part>
<part type="winlog">
<startTime>5227967</startTime>
<endTime>5228045</endTime>
<key>VK_RETURN</key>
<value>
</value>
<keyboardstate>
</keyboardstate>
</part>
</event>
```

Word count

- The word count can be used as input for fluency measures
- Examples of use:
 - Live subtitling
 - Multiple source analyses
 - L1 versus L2 writing
 - Translations
 - Thinking aloud protocols

Word count (live subtitling via speech recognition)

- Open document "Transcription Flemish Broadcasting Company - Extra Time - Word count.xlsx"

	E	F	G	H	I	J	K	L	M
Spoken Comment by interviewee	# characters (exact spaces)	# characters (exact spaces)	# characters (exact spaces & "')	# words		Dictated text by respondent (inputlog (DMS filtered))	# characters (exact spaces)	# characters (exact spaces)	# words
Cluid Brugge zou een normaal verspreiden	51	45	44	7		Cluid Brugge zou vandaag een normaal verspreiden	50	44	7
zou zelfs een persconferentie geven, maar ja dat is er niet van gekomen.	72	60	59	13		maar dat is er niet van gekomen	31	25	7
Er stonden de hele dag journalisten voor het Jan Breijdelstadion met camera's in aanloop	87	74	74	14		en stonden de hele dag journalisten voor het Jan Breijdelstadion	63	54	10
maar uiteindelijk geen woord van Cluid Brugge.	45	39	38	7		maar uiteindelijk geen woord van Cluid Brugge	44	38	7
Dit is Jan van Winie, de physical trainer	43	36	35	8		dit is Jan van Winie, de physical trainer van Cluid Brugge	55	49	11
Die stopt welgemiddeld naar huis, verder zie ik.	48	42	41	7			0	0	0

Word count calculations

- Character count including spaces: =LEN(TRIM(E2))
- Character count excluding spaces: =LEN(TRIM(SUBSTITUTE(E2;" ";""))))
- Character count excluding spaces and full stops =LEN(TRIM(SUBSTITUTE(E2;" ";"")))-IF(ISNUMBER(SEARCH(".",E2));1;0)
- Word count =IF(LEN(TRIM(E2))=0;0;LEN(TRIM(E2))-LEN(SUBSTITUTE(E2;" ";""))+1)

Word count (Multiple sources)

Masterthesis Evelyn Van Dyck

Word count (Multiple sources)

Word count (Multiple sources)

- Open file "Fluency – wordcount"

Type	Output	Position	DocLength	StartTime	StartBack	EndTime	EndBack	ActionTime	PauseTime	ProcessTime	RangeLength	X	Y
mouse	LEFT Click	5584	00:00:12	5571	00:00:12	5584	00:00:12	0	0	UNUSABLE PAUSE	848	102	
mouse	LEFT Click	5613	00:00:12	5600	00:00:12	5613	00:00:12	0	0	UNUSABLE PAUSE	848	102	
focus	TASKBAR	5638	00:00:12	5625	00:00:12	5638	00:00:12	0	0	UNUSABLE PAUSE	848	102	
mouse	Movement	5686	00:00:12	5673	00:00:12	5686	00:00:12	0	0	UNUSABLE PAUSE	848	102	
mouse	LEFT Click	5775	00:00:12	5762	00:00:12	5775	00:00:12	0	0	UNUSABLE PAUSE	848	102	
mouse	Movement	5793	00:00:12	5780	00:00:12	5793	00:00:12	0	0	UNUSABLE PAUSE	848	102	
keyboard	LCtrl + V	6075	00:01:00	6062	00:01:00	6075	00:01:00	0	0	UNUSABLE PAUSE	848	102	
mouse	Movement	6199	00:01:01	6186	00:01:01	6199	00:01:01	0	0	UNUSABLE PAUSE	848	102	

BEFORE [121] Inputlog is a logging tool that logs all types of input modes: keyboard, mouse & speech recognition. Researchers make frequent use of keystroke logging tools to describe online writing or translation processes in detail.

[ENG] [121]

Insert LEFT Click 6250 00:01:02.510 6264

mouse Movement 6269 00:01:02.619 6283

Linear analysis

Time	Position	DocLength	StartTime	StartBack	EndTime	EndBack	ActionTime	PauseTime	ProcessTime	RangeLength	X	Y
5793	00:00:57.939	5780	00:00:57.939	00:00:57.939	5793	00:00:57.939	0	0	UNUSABLE PAUSE	848	102	
5793	00:00:57.939	5780	00:00:57.939	00:00:57.939	5793	00:00:57.939	0	0	UNUSABLE PAUSE	848	102	

V-lookup

- =VLOOKUP(C2;'N-square'!\$A\$4:\$C\$8;3)

Worksheet basic

	A	B	C
1	Actiontime	PauseTime	Type of document
2	140	150	main document
3	140	800	main document
4	140	1000	main document

Worksheet N-square

	A	B	C	D
1	Type of sources			
2				
3	From	To	From	To
4	google	google	3	3
5	main document	main document	1	1
6	powerpoint	powerpoint	4	4
7	word		2	2
8	zzz		5	5

More information

Mariëlle Leijten, Flanders Research Foundation, Belgium
marielle.leijten@ua.ac.be ~ www.ua.ac.be/marielle.leijten

Luuk Van Waes, University of Antwerp, Belgium
luuk.vanwaes@ua.ac.be ~ www.ua.ac.be/luuk.vanwaes

www.writingpro.eu **writing PRO**
knowledge center for writing process research

www.inputlog.net **inputlog**

www.jowr.org **journal of WRITING RESEARCH**