

## Observing and analysing digital writing processes with Inputlog

Luuk Van Waes & Mariëlle Leijten

erful simple savvy legendary bright  
ntrospective social nourishing vast  
energizing uplifting funny shared  
autiful provocative insightful mad  
g current dynamic engaging lasting  
us elevating thoughtful poignant  
bold original detailed vivid giving

erful simple savvy legendary bright  
ntrospective social nourishing vast  
energizing uplifting funny shared  
autiful provocative insightful mad  
g current dynamic engaging lasting  
us elevating thoughtful poignant  
bold original detailed vivid giving

## Topics

- Pauses
- Thinking aloud protocols
- Linguistic corpora
- Fluency

## Introduction

erful simple savvy legendary bright  
ntrospective social nourishing vast  
energizing uplifting funny shared  
autiful provocative insightful mad  
g current dynamic engaging lasting  
us elevating thoughtful poignant  
bold original detailed vivid giving

Inputlog output:

- filtered output ▶ representation
- analyses ▶ rule based interpretation

Manual analyses: ▶ dependent on research objectives

## Pauses

erful simple savvy legendary bright  
ntrospective social nourishing vast  
energizing uplifting funny shared  
autiful provocative insightful mad  
g current dynamic engaging lasting  
us elevating thoughtful poignant  
bold original detailed vivid giving

## A pause is a pause is a pause...



## Pauses

erful simple savvy legendary bright  
ntrospective social nourishing vast  
energizing uplifting funny shared  
autiful provocative insightful mad  
g current dynamic engaging lasting  
us elevating thoughtful poignant  
bold original detailed vivid giving

- **Concept**  
Pauses are the basic concept in most cognitive writing studies: important indication of the writer's increased cognitive load.
- **Writing vs. Speech**  
Writer is less conscious about the process, more freedom to distribute his attention; reader's feedback is product based (Strömquist et al. 2006)
- **Matsuhashi**

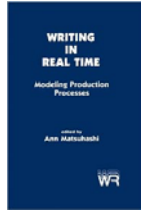
# Matsuhashi 1981

## Pausing and Planning: The Tempo of Written Discourse Production

ANN MATSUHASHI  
University of Illinois at Chicago Circle

### ABSTRACT

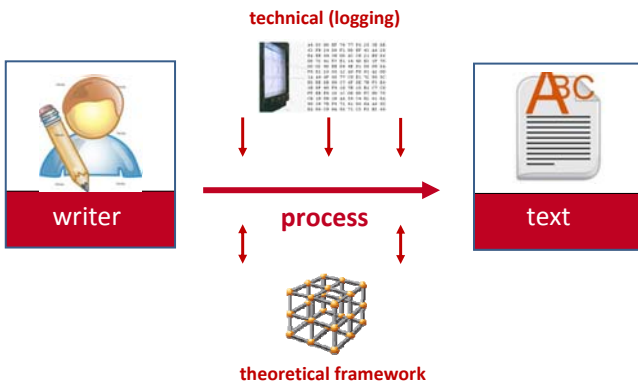
This study explores how discourse purpose influences text production. It assumes that pauses during writing offer clues to cognitive planning processes during written discourse production. Four skilled high school writers composed two compositions for each of three discourse purposes: to report, to persuade, and to generalise. The writing activity of each subject was videotaped and aimed to determine the lengths of pauses between words. The written compositions were analysed to identify eight language units: abstraction level, sentence role, paragraph, initial modifying structure, lexical cohesion in the object of an adjective phrase, syntax, and content words. Mean pause lengths were compared for the overall discourse categories and pairs to each of the eight language units. Significant differences were found for discourse purpose for two of the four writers, although a strong trend suggested that generalising and persuading were more time consuming than reporting for all writers. Significant differences were found for different abstraction levels, for T-units which do and do not begin paragraphs, and between non-cohesive simple subjects and those cohesive through the repetition of same word. The results suggest that purpose for writing influences planning for global as well as for local aspects of text production.



## Purpose

- Planning
- Evaluating (reading / skimming / identifying)
  - for revision
  - for planning
  - TPSF / analogue or digital sources
- Distraction

## Perspectives to pauses



## technical (logging)



## Pauses: technical

- accuracy
  - measurement error
  - filtering
- instrumental definitions
  - threshold
  - key in/out
  - aggregation
- characteristics
  - writing mode: keyboard mode & lay-out, mouse, speech

## Accuracy

- video registration: 25 pictures/sec ~40 ms
- keystroke logging: ~10 ms (processor dep.)



## Interkey intervals



### pausing or flight time

the latency between pressing the previous key and the actual key

$$\text{Pause}_i = P_i - P_{i-1}$$

### action or dwell time

latency between pressing a key and the release of the same key

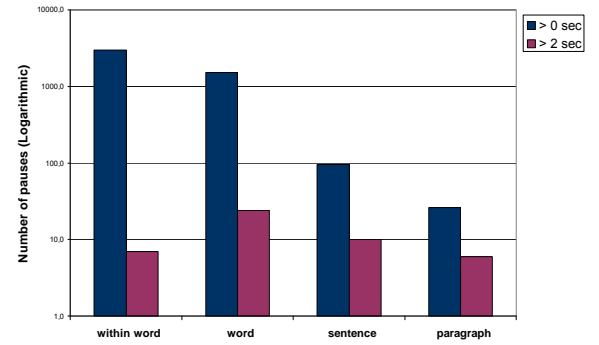
$$\text{Action}_i = P_i - R_i$$



## Threshold

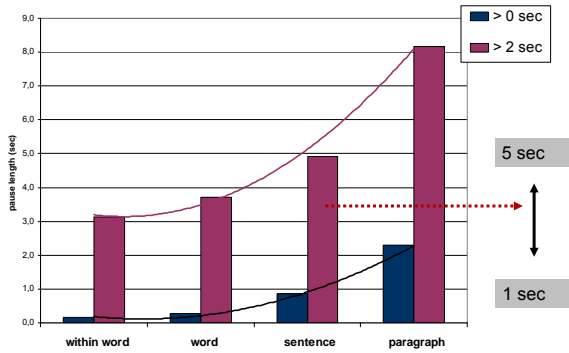
- speech
  - threshold: > 200 ms
- writing
  - each transition > 1 ms (*remark: key combinations*)
  - threshold: > 2000 or 3000 ms

## Number of pauses at different text levels

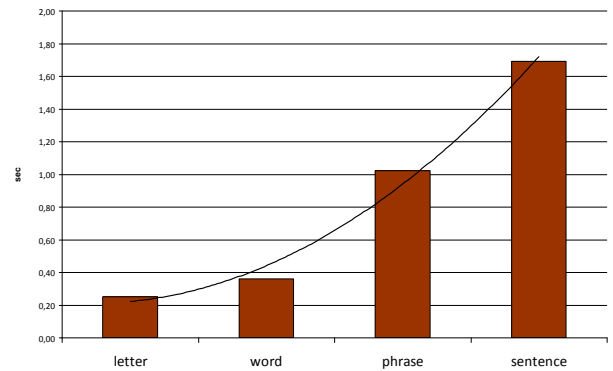


EJ Case

## Mean pause length at different intervals

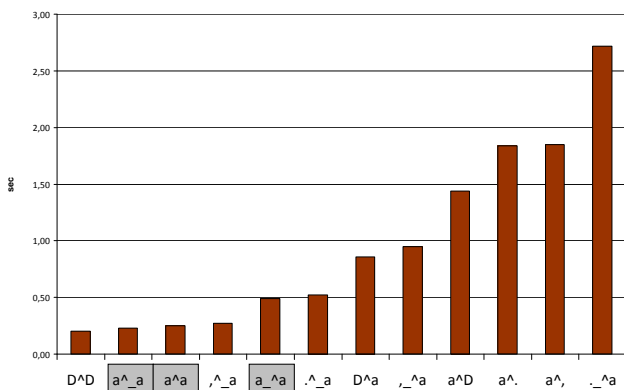


## Median Transition times

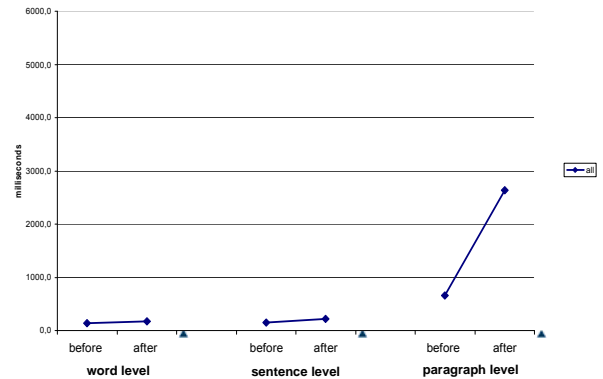


source: Wengelin 2006

## Median Transition times

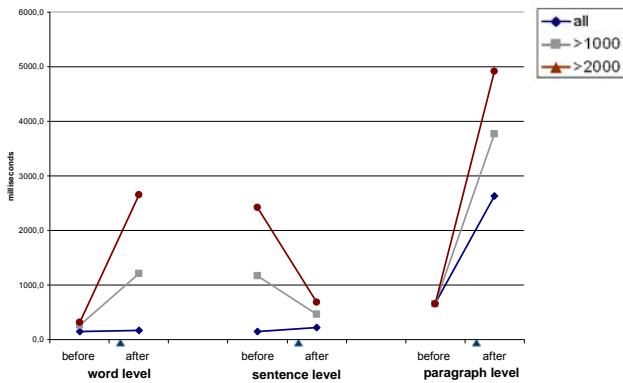


## Pauses before-after syntactic boundaries



EJ case

## Pauses before-after syntactic boundaries



## Defining pause characteristics

00:09:13

This raises the overall temperature of our world.

EJ case

Write an essay designed to convince people that petrol prices should be increased to reduce global warming.

normal

melting of the polar ice caps. Many people in the north of the world have already become "refugees of global warming", and have had to leave their homes due to extreme flooding. However, just because these major issues are not right on our doorstep does not mean that we are not feeling the effects. England's odd weather in recent years can also be blamed upon global warming, and as our summers become warmer and people start to die from heat exposure, our rains also increase and people homes become ruined. These events can be reduced if everyone does their bit to help the environment around them, this could include actions such as turning off the lights when leaving a room, reducing waste by recycling and not leaving electrical appliances plugged in when not in use. A big part of global warming can be blamed upon fuel consumption. When petrol is used the exhaust fumes travel into the atmosphere and damage the ozone layer. This has caused holes to appear and when the sun's rays travel through they are stronger, and can also become trapped by all of the debris in the

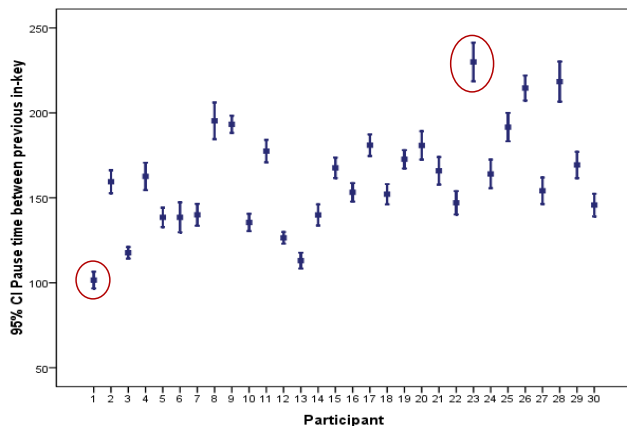
Task ID: 413  
 EVENT\_ID: 413  
 EVENT\_TIME: 326593  
 EVENT\_TYPE: INSERT  
 EVENT\_KEY: TEXT  
 EVENT\_KEY\_STATE: REGULATION  
 EVENT\_KEY\_STATE\_FROGATION  
 StoppingMode: NEXT  
 TEXT\_CHAR\_COUNT: 207  
 TEXT\_WORD\_COUNT: 35  
 TEXT\_SENTENCE\_COUNT: 13  
 TEXT\_PARAGRAPH\_COUNT: 1  
 KEY\_KEY:  
 PREV\_CHAR\_SEQUENTIAL\_Tow  
 PREV\_CHAR\_INTERFAL: 145  
 PREV\_CHAR\_CURSOR\_RC [D1,10]  
 PREV\_CHAR\_TIME: 326546  
 PREV\_CHAR\_KEY: +  
 NEXT\_CHAR\_SEQUENTIAL\_Tow  
 NEXT\_CHAR\_INTERFAL: 191  
 NEXT\_CHAR\_CURSOR\_RC [D1,12]  
 NEXT\_CHAR\_TIME: 326574  
 NEXT\_CHAR\_KEY:



## Pauses: writer

- personal characteristics  
e.g. age, disorders
- technical expertise  
e.g. touch typist vs. peck-and-hunt, left-right handed
- language expertise  
e.g. L1/L2
- genre expertise  
e.g. knowledge telling/transforming/crafting

## Differences between persons (n=20 | CI 95%)



## Criteria: intra word transition

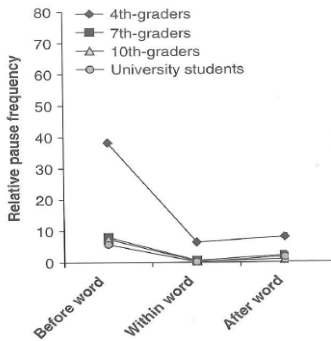
### writer variance

	fastest (ms)	slowest (ms)
Copy task	110	228
R&W corpus	247	488
Spencer corpus	181	568
Dyslexia	796	

Remark: bigram variance

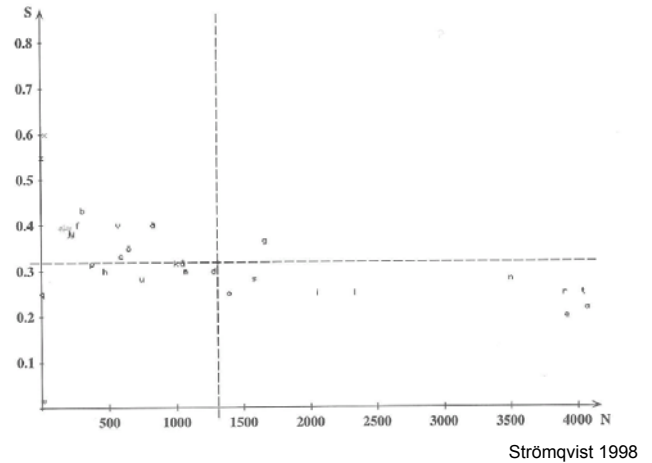
## Discriminating pauses

before word: most discriminating (early) developmental factor



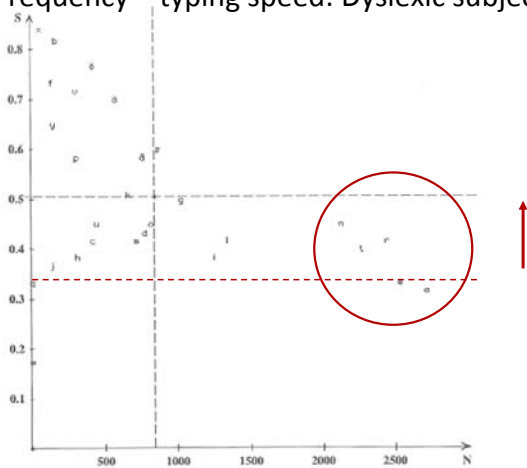
wengelin & strömquist 2004

## Frequency ~ typing speed: Control group



Strömquist 1998

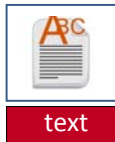
## Frequency ~ typing speed: Dyslexic subjects



## Typing skills

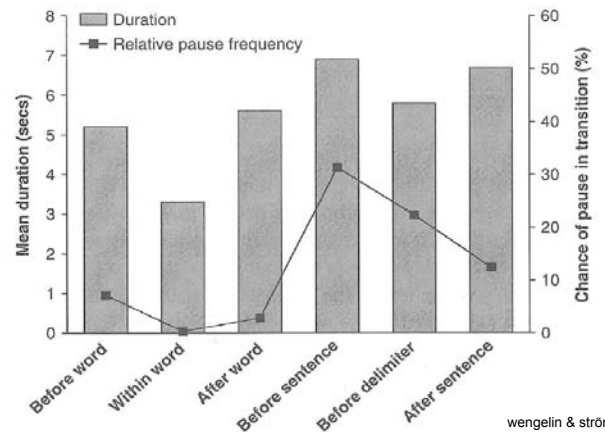
- **Fast typing** seems to co-occur
  - with much deleting
  - higher probability of mistyping
  - higher probability of immediate correction of typos
 Grabowski et al (2010): copy task
- **Personal preferences**
  - immediate correction of spelling errors
  - vs. mainly correcting at the end of the writing task
- **Touch typists** (cf. eye tracking)
  - shorter pauses
  - longer action times

## Pauses: text



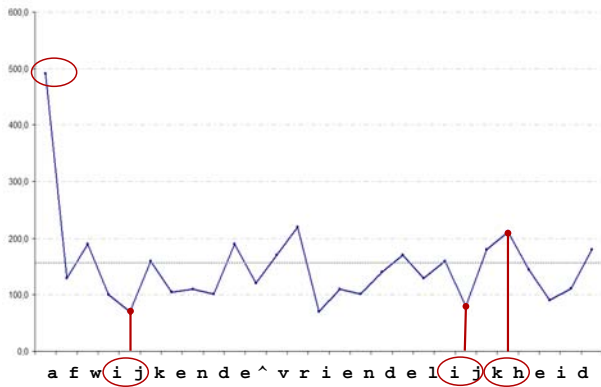
- level: textual and linguistic hierarchy
  - letter – word – clause - sentence - paragraph
  - e.g. bigram: frequency, syllable boundaries, morphological, correctness
  - framing devices (pragma-linguistics perspective)
  - pauses mainly located at “discoursally significant junctures”
  - cf. Spelman Miller (2006)
- task & genre
  - e.g. text length
  - e.g. genre characteristics

## Duration / Frequency

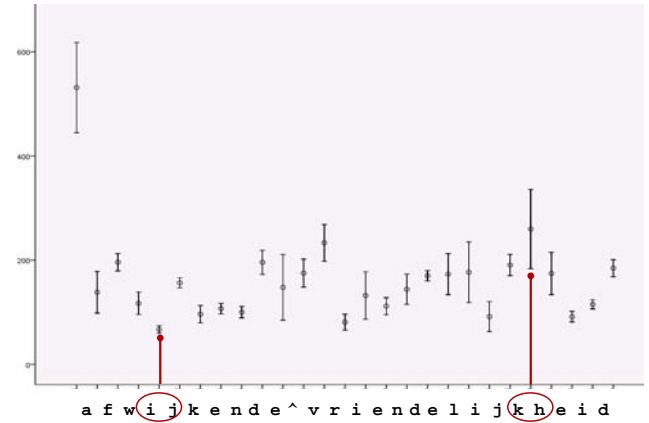


wengelin & strömquist 2004

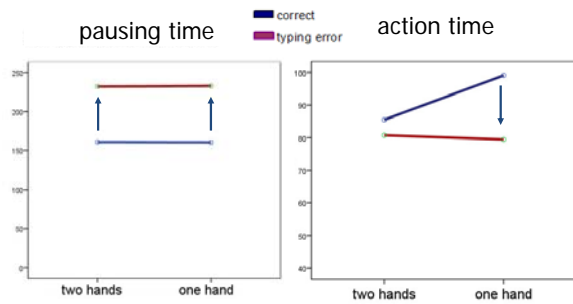
### Differences within digraphs (median | n=20)



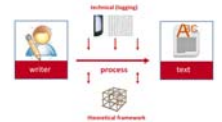
### Differences within digraphs (n=20 | CI 95%)



### Pausing and Action Time: Hand Comb.

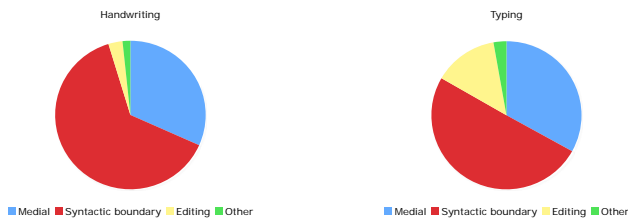


### Pauses: process



- characteristics
  - duration of process
  - fragmentation
  - complexity
  - subprocesses: e.g. planning vs. revision, distribution
  - collaboration
  - writing medium
- intertextuality
  - translation
  - internal and external sources
  - analogue and digital sources

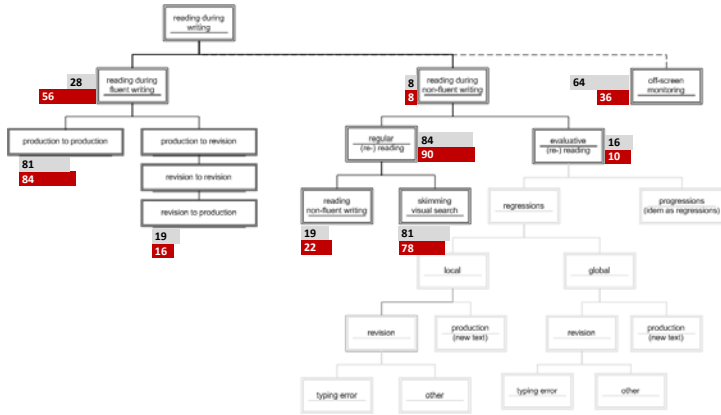
### Pause Location



### What happens during a pause?

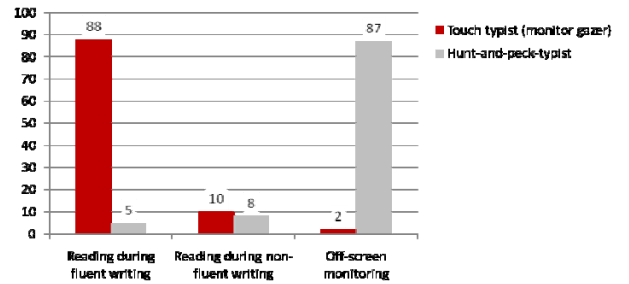
- interpretation of context
- combination with eye-tracking
- combination with thinking aloud protocols

## Reading during writing - scheme

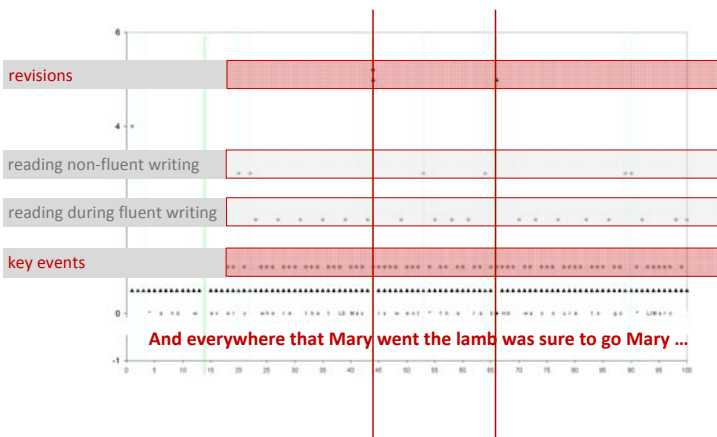


## Reading during writing

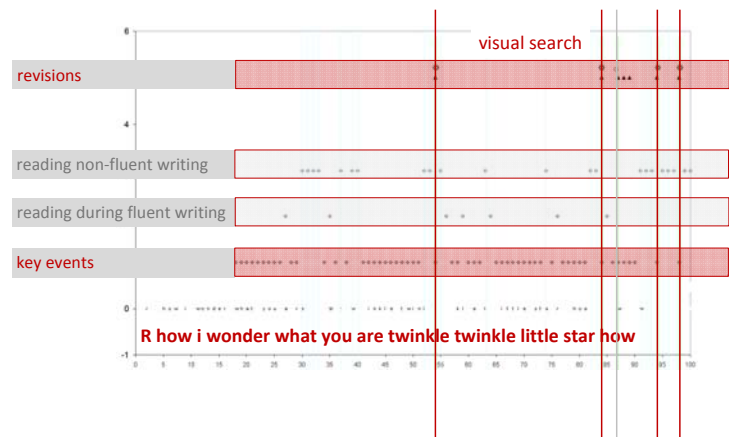
Process of two extremely different writers



## Reading during writing – monitor gazer



## Reading during writing – hunt-and-peck



## Keystroke logging and Thinking aloud protocols

rful simple savvy legendary bright  
 ntrospective social nourishing vast  
 energizing uplifting funny shared  
 eautiful provocative insightful mad  
 g current **dynamic** engaging lasting  
 us elevating thoughtful poignant  
 bold original detailed vivid giving

## Aggregating KSL data to word and sentence level

rful simple savvy legendary bright  
 ntrospective **social** nourishing vast  
 energizing uplifting funny **shared**  
 eautiful provocative insightful mad  
 g current dynamic engaging lasting  
 us elevating thoughtful poignant  
 bold original detailed vivid giving

## Fluency measures

erful simple savvy legendary bright  
ntrospective social nourishing vast  
energizing uplifting funny shared  
autiful provocative insightful mad  
g current dynamic engaging lasting  
us elevating thoughtful poignant  
bold original **detailed** vivid giving

## More information

Mariëlle Leijten, Flanders Research Foundation, Belgium  
marielle.leijten@ua.ac.be ~ [www.ua.ac.be/marielle.leijten](http://www.ua.ac.be/marielle.leijten)

Luuk Van Waes, University of Antwerp, Belgium  
luuk.vanwaes@ua.ac.be ~ [www.ua.ac.be/luuk.vanwaes](http://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**